Hendricks, M., Plantz, M. C., & Pritchard, K. J. (2008). Measuring outcomes of United Way–funded programs: Expectations and reality. In J. G. Carman & K. A. Fredericks (Eds.), *Nonprofits and evaluation. New Directions for Evaluation*, 119, 13–35.

2

Measuring Outcomes of United Way–Funded Programs: Expectations and Reality

Michael Hendricks, Margaret C. Plantz, Kathleen J. Pritchard

Abstract

In 1996, United Way of America (UWA) developed and began disseminating the most widely used approach to program outcome measurement in the non-profit sector. Today an estimated 450 local United Ways encourage approximately 19,000 local agencies they fund to measure outcomes. The authors first describe and then assess the strengths and limitations of the distinguishing features of the UWA approach, efforts to disseminate the approach, implementation by local United Ways, and actual outcome measurement by local agencies. The chapter ends with a description of United Way's relatively new emphasis on community impact and how that initiative relates to program outcome measurement. © Wiley Periodicals, Inc.

Note: The authors thank Don Buchholtz, Joanne Carman, Kimberly Fredericks, Ken Fyfe, Gail Harris, Harry Hatry, Melanie Hwalek, Linda O'Black, Beverly Prueter, Dawn Hanson Smart, Jeff Stys, and Roger Wood for their helpful comments on an earlier draft of this chapter.



States, is a federated system of more than 1,300 autonomous, local United Ways across the country. The United Way mission calls for each United Way (UW) to improve its community by collaborating with local stakeholders to forge a shared agenda of local problems and develop strategies to address the underlying causes of these problems. To help these strategies become reality, UWs mobilize people and organizations to invest time, skills, expertise, and money—just over \$4 billion across the system in 2006. The United Way of America (UWA) is the national leadership and support organization for this system.

At least three forces converged to spur UWs to begin requiring local agencies to measure the outcomes of UW-funded programs. First, local UWs wanted to ensure that they could both direct money to demonstrably effective programs and demonstrate to donors the results of their financial contributions to the UW. Starting in the 1980s, a small but growing number of UWs began local initiatives to measure program outcomes. By the mid-1990s, many UWs were encouraging UWA to develop a methodology and resources that all could use, sparing individual UWs from having to create their own approaches and materials.

Second, measuring outcomes reflects the logical evolution of performance measurement in the nonprofit sector in general. Nonprofit programs already were measuring several aspects of performance including inputs, activities, outputs, financial accountability, adherence to standards of quality in service delivery, participant-related measures, and client satisfaction (Plantz, Greenway, & Hendricks, 1997). UWs had been asking funded programs to report on these measures for many years. An obvious next step was to start measuring program outcomes, defined as benefits or changes for program participants (sometimes irreverently called the "So what?" question).

Third, the early 1990s saw other sectors of U.S. society enter a general era of results and accountability. The public sector, for example, was strongly influenced by the Government Performance and Results Act of 1993 (GPRA), which required each federal agency to define its desired results and measure progress annually. At the same time, the private sector was influenced by the quality improvement movement and a push for measurable goals and objectives, while the health care industry in particular was influenced by the outcome-oriented tenets of managed care.

Some skeptics have suggested another possible reason for the timing of UWA's interest in documenting effective outcomes: its potential public relations value. In 1992, the entire UW system was shaken by a scandal involving misuse of UWA funds by then-president William Aramony, who was forced to resign. Aramony was eventually convicted of fraudulent use of UWA funds and sent to prison. The fallout damaged UWA's credibility and fundraising efforts at local UWs and helped fuel a decline in trust in the nonprofit system in general. To prevent future abuses and rebuild trust,

UWA and a growing number of local UWs adopted organizationwide emphasis on accountability in the wake of this scandal. At UWA, the emerging interest in outcome measurement became one component of the push for greater accountability.

For whatever combination of reasons, in 1995 UWA initiated a highlevel and highly visible effort to help UWs document and improve the results of their investments in human services, including measurement of program outcomes. Estimates are that UWA may have spent nearly \$2.4 million on this effort between 1995 and 2000, including grants received from the Lilly Endowment, the W. K. Kellogg Foundation, and the Ewing Marion Kauffman Foundation. This chapter examines four aspects of the program outcome measurement component of that effort: (1) the approach UWA developed at the national level, (2) rolling out the approach, (3) how local UWs implement the approach, and (4) outcome measurement as actually practiced by local agencies funded by UWs. For each of these four, we first describe the aspect and then present our assessment, both pro and con. A final section touches on United Way's current emphasis on community impact and how program outcome measurement intersects with that focus.

The UWA Approach to Measuring Program Outcomes

After creating a new internal unit to manage this function, UWA established a 31-member Task Force on Impact that included several prominent evaluators: James Bell, David Cordray, Stacey Daniels, Harry Hatry, Astrid Merget, Ricardo Millett, Patricia Patrizi, Howard Rolston, James Sanders, Carol Weiss, and Joseph Wholey. The task force also included representatives from local UWs in the United States and Canada, national health and human service organizations, foundations, and the public sector. The first task was to develop an approach by which local UWs could ask programs they fund to (1) identify the benefits they expect clients to experience as a result of their services, (2) measure the extent to which clients actually achieve those results, (3) use the information to increase the programs' effectiveness, and (4) share news of that effectiveness with many audiences.

The task force drew from established evaluation practice and expertise, as well as the grassroots innovations and lessons of local UWs and others, to develop United Way of America's approach to measuring outcomes. The approach has several distinguishing features (see Table 2.1). Conceptually, it:

• Emphasizes outcomes as more meaningful measures of a program's value than outputs. This predated the position adopted by the U.S. Office of Management and Budget (OMB) to assess federal programs via its Program Assessment Rating Tool (PART): "Outcome measures are the most informative measures about performance, because they are the ultimate results of a program that benefit the public. Programs must try to translate existing

Table 2.1. Distinguishing Features of the United Way of America Approach to Measuring Program Outcomes

Conceptual features

Focus on outcomes as measures of effectiveness Quantitative measurement of outcomes Regular, systematic measurement Not evaluation as traditionally defined Program improvement as main objective Local measurement necessary

Practical features

Most steps done by in-house staff of UWs and agencies Avoids often-confusing terminology Logic model a key component Programs identify their own outcomes Supports using products of other national organizations Relatively long time horizon for implementation A round or two of measurement precedes setting targets

measures that focus on outputs into outcome measures . . . " (U.S. Office of Management and Budget, 2007, p. 9).

- Encourages agencies to develop ways to measure outcomes quantitatively. Although it recognizes that qualitative information can help both in designing meaningful measurement and in understanding and responding to the findings, and it acknowledges the limitations of quantitative data, the UWA approach nonetheless emphasizes numerical evidence of program performance.
- Values repeated measurement of outcomes at appropriate intervals. This interval will vary with the outcome; students' learning might be measured weekly, for example, while women's heart health might appropriately be measured only once a year. But the approach encourages accumulation of longitudinal data.
- Does not attempt to be evaluation as most evaluators know it, but a more modest effort simply to track outcomes. Therefore outcome measurement cannot support assertions of causality or explain why a certain level of outcome was achieved. For this reason, and even though some UWs call their efforts "outcome evaluation," UWA has consistently used the term outcome measurement.
- Emphasizes improving program effectiveness as the primary reason for measuring program outcomes, with external accountability as a secondary purpose. This order is sometimes difficult for both local UWs and agencies to adhere to consistently, but the approach advocates outcome measurement primarily as a tool to improve programs. In this regard, the UWA

- approach parallels the W. K. Kellogg Foundation's primary emphasis on improving programs (cf. W. K. Kellogg Foundation, 1998).
- Asserts that demonstration of effectiveness in national studies does not substitute for measurement of effectiveness in individual sites. That is, even though a particular set of services or delivery methods may have been rigorously shown to be effective under certain conditions, this fact does not automatically equate to program effectiveness at a given local site, which may operate under quite different conditions. In the UWA approach, the only way to confirm local effectiveness is to measure local effectiveness.

In addition to these distinguishing features at the *conceptual* level, the UWA approach also has several distinguishing features at the *practical* level. For example, the approach:

- Suggests that most aspects of outcome measurement can be implemented by in-house staff of local UWs and agencies, supplemented as needed by help from contracted evaluators or other outsiders. As we will see below, this belief may be more or less accurate, but it underlies many of the approach's practical suggestions.
- Avoids attempting to parse the often ill-defined differences among terms such as mission, goal, purpose, objective, result, outcome, and impact and focuses only on a limited number of terms related to outcomes. The fact that certain results must be achieved first in order to achieve later results is addressed by speaking of a specific sequence of outcomes—initial, intermediate, and longer-term—and by highlighting the "if-then" linkages between activities and the resulting outcomes at different levels.
- Advocates development of a program logic model as a valuable tool for discovering and displaying the links between activities and outcomes. For many outsiders, this may be the most visible symbol of the UWA approach; the four-part graphic of Inputs-Activities-Output-Outcomes has become familiar to many persons within and outside the UW system.
- Asserts the value for a program of identifying its own logic model, outcomes, and indicators, both to increase shared understanding of the program's theory of change and to assure the relevance and usefulness of outcome data to program managers and staff. Recognizing the potential for harm when such materials are imposed externally or plucked thoughtlessly from a list of options developed by people without program-relevant expertise, UWA's task force recommended against creating repositories of standardized materials from which local agencies can make selections. Other organizations, responding to their undeniable time-saving appeal, have developed such repositories (e.g., www.urban.org/center/cnp/projects/outcomeindicators.cfm).
- Encourages, recognizes, and refers UWs and local agencies to national health and human service organizations that have developed outcomes, indicators, measurement tools, analysis procedures, and reporting systems specifically for their substantive program areas (e.g., UWA 2003b).

- Urges practitioners and funders to expect a relatively long time horizon for developing, testing, and refining outcome measurement systems. In fact, UWA asserts that agencies will require two to four years from the time the idea of outcome measurement is first introduced to the point of having meaningful outcome data. To our knowledge, no other approach acknowledges this reality so explicitly.
- Advises against establishing performance targets until the program has collected enough outcome data to know what targets are reasonable. This contrasts with GPRA requirements to establish performance targets immediately, regardless of whether sufficient baseline data exist.

Our Assessment of the UWA Approach to Measuring Program Outcomes

In the interest of full disclosure, each of us is closely connected to the United Way efforts to measure program outcomes. Hendricks has consulted with UWA and continues to consult regularly with various UWs around the country; Plantz was instrumental in developing UWA's program outcome measurement efforts and continues some involvement in that work; Pritchard pioneered outcome measurement efforts in a local UW that became a system leader and was working at UWA when this chapter was written. Clearly we are not unbiased, outside observers. Nonetheless, despite these connections—or perhaps partly because of them—we feel uniquely positioned to identify some of the strengths and weaknesses of the UWA approach.

As Table 2.2 shows, we believe the UWA approach contains many strong elements:

- Its clear emphasis on outcomes is moving the entire nonprofit sector in this direction (Hendricks, 2002), as can also be seen in other chapters of this issue.
- Stressing program improvements, not external accountability, as the primary reason for measuring outcomes is beginning to change how local-level programs and their funders think about evaluative activities.
- Standardizing and simplifying the terminology helps everyone involved, and it would be a good model for the evaluation world in general to emulate. We spend far too much effort parsing the differences between unnecessarily confusing terms.
- Agency after agency has remarked on the benefits of logic model thinking. It has been enormously helpful for programs that too often have no clear, shared vision of what their program is trying to accomplish, or how.
- Public acknowledgment that implementing such a fundamental organizational change will take years, not months, relieves pressure to produce outcome data immediately and allows UWs and agencies to experiment with approaches tailored to their unique situation.

Table 2.2. Our Assessment of the UWA Approach to Measuring Program Outcomes

Strengths

Emphasis on outcomes
Program improvement as primary motivator
Simplified terminology
Logic model as important tool
Timeframe for implementation in years, not months
Makes some evaluation practices practical

Limitations

Underestimated challenges of outcome measurement for in-house agency and UW staff
Underestimated difficulties of imparting to in-house staff the evaluation skills needed to meet these challenges
Underestimated need for TA after initial training
Too little guidance on how agencies can use data
Too little guidance on using qualitative narratives to identify and illustrate quantitative outcomes
Underemphasized importance of regular progress
Insufficient links to quality tools and examples

• Finally, and importantly, the UWA approach makes application of some evaluation practices practical for local UWs and agencies. Prior to UWA's effort, most UWs and agencies had focused on counting activities and outputs (mostly clients served) while a few had experienced expensive, one-time, university- or consultant-completed "program evaluations," which too often produced neither a useful product nor a satisfying process. UWA's approach offers a doable process yielding information that programs can use to improve their results for clients.

On the other hand, we also believe that time has revealed several limitations to the approach.

- First, the approach underestimated not only the challenges agency staff face in implementing and using outcome measurement but also the challenges UW staff face in helping agencies accomplish those tasks. Either role would be challenging for even a highly experienced evaluator, and although some agency and UW staff had sufficient evaluation training and experience to fulfill these new roles, the majority of in-house staff found the new tasks to be outside their previous experience and expertise.
- Second, the approach underestimated the difficulties of imparting the needed evaluation skills to in-house staff of both agencies and UWs, especially those staff with little or no background in the discipline. As readers of this journal know quite well, evaluation training can be a lifelong journey. We mention

here some concerns about the specific strategy used to train in-house staff, but the task would have been difficult using any strategy. It is simply not easy for agency and UW staff to become skilled practitioners of outcome measurement.

- A third, related limitation is the failure to recognize that ongoing, task-specific, and content-specific technical assistance is essential to effective implementation, even if initial training is bulletproof. Training is necessary, but training alone is insufficient; ongoing technical assistance is also required.
- Fourth, UWA's task force focused more of its efforts on developing the front end of the process—UW implementation of outcome measurement among programs and agency collection of outcome data—than on the back end of the process, especially ways to analyze, interpret, and use those data. As a case in point, of the 170 pages in the UWA manual, 18 (11%) are devoted to using the findings. These pages describe 12 management challenges that outcome data can help agencies meet, but they offer little guidance on exactly what steps agencies can take to use the data for these purposes. For example, the UWA manual encourages agencies to use outcome data to improve the effectiveness of programs. Some agencies know specific steps for making this happen—for example, analyze the outcome data to pinpoint where the program is having more and less success, interpret the implications, brainstorm possible ways to improve services, implement trials, draw conclusions, and revise the program.

But for many other agencies, specific guidance about these sorts of steps would have been (and still would be) very helpful. As a result of the relative dearth of back-end guidance in written resource materials and training curricula, the effort is perceived by some as focused on research-oriented data gathering, not management-oriented program improvement (Kopczynski & Pritchard, 2004). This limitation may be harder for evaluators to help remedy because they too have traditionally been only minimally trained in the management uses of data, but perhaps the evaluation profession has more to offer than we realize.

- Fifth, the emphasis on quantitative measurement is difficult for local health and human service agencies that are used to telling their success stories with narrative vignettes. Guidance is needed to help agencies use their narratives to identify intended outcomes and see measurement as a way to learn whether the stories are unique or representative of other clients. Agencies could then use their vignettes to illustrate their outcome data rather than claiming the vignettes as evidence of outcomes. Perhaps evaluators could help by suggesting ways to build on qualitative approaches, such as Davies's most significant change method (Davies & Dart, 2005) or Brinkerhoff's success case method (Brinkerhoff, 2003).
- Sixth, although we applaud UWA for advocating a realistically long timeframe for agencies to implement program outcome measurement, early guidance underemphasized the importance of making regular progress within this time period. As a result, implementation was sometimes characterized by

significant activity immediately following agency training, then months of minimal thought or progress capped by frantic catch-up just before an annual report to the UW was due. An extreme example is one UW that gave agencies three years to implement outcome measurement, required very little of those agencies for more than two years, and then expected them to have reliable outcome data at the end of the third year. Although UWA later issued supplemental advice that UWs set up a schedule by which agencies would demonstrate progress in small but frequent steps, the message was not early or widespread enough to become standard procedure. Allowing a long time horizon is fine, but to develop an effective system the time must be filled with constant progress.

• A final topic, which may or may not represent a limitation of UWA's approach, is its decision not to create a comprehensive national repository of logic models, outcomes, indicators, and data collection tools. Nationwide, hundreds of UW-funded programs struggle with how best to measure outcomes for youth development, domestic violence, early learning, senior care, drug abuse, financial stability, housing, and a host of other health and human services. There is no doubt that these programs would save considerable time and effort, at least initially, if they could access materials already developed by similar programs. Many agencies have asked local UWs to make such resources available, and some UWs have created local repositories of materials to share among the agencies they fund. In turn, many UWs have asked UWA to gather, store, and share these materials on a national level, but UWA has consistently declined, guided by the recommendation of its national task force as described earlier. Is this a limitation of UWA's approach, or simply a fact?

An alternative approach to this challenge might be for UWA to identify and furnish links to compilations prepared by others, especially those with subject-matter expertise offering already vetted resources and tools. For example, the Girl Scouts of the USA has developed a series of outcome measurement materials for programs helping to develop young girls (see http://www.girlscouts.org/research/publications). UWA's second report on outcome measurement activities of national health and human service organizations (United Way of America, 2003b) included some of these program-specific references, and UWA's Outcome Measurement Resource Network originally supplied links to many such resources. Continuing these early efforts to create links to examples of program-specific outcomes and indicators and tested data collection tools would have been helpful to many UWs and programs.

Rolling Out the UWA Approach

Once the task force had developed UWA's approach to measuring program outcomes, UWA created a number of resources to roll out the approach throughout the entire UW system. Most well-known is its 170-page manual

for agencies, *Measuring Program Outcomes: A Practical Approach* (United Way of America, 1996). This manual, with lead authorship by Harry Hatry, a well-known performance measurement expert at the Urban Institute, is now in its 15th printing and has sold more than 160,000 copies. For the years 2003–2005, more than one-third of those purchasing the manual were government offices, foundations, students, consultants, or colleges and universities. Regarding the last group, nearly 100 colleges and universities purchased the manual during a recent 14-month period, and more than 50 purchased 10 or more copies, suggesting its use in a human services, social work, or public administration course.

The manual was one of five resources UWA released in 1996. The others were a practice-oriented training kit aligned with the content of the manual, a guide on implementing program outcome measurement for UW staff, a shortened version of this guide for UW board members and other volunteers, and a video introduction to basic outcome measurement terms and concepts.

Many additional resources have followed. Some help UWs deal with implementation issues, such as job descriptions for UW program outcome measurement staff; a report on the essential underpinnings that need to be in place for a UW to plan for, implement, sustain, use, and benefit from program outcome measurement (United Way of America, 2003a); examples of UWs measuring their own outcomes; and most recently a guide for building or buying software for a Web-based system UWs can use to manage agencies' outcome data (United Way of America, 2007a) and descriptions of existing software and systems that UWs can use for this purpose.

Other UWA resources help UWs build agency capacity for outcome measurement, including a 90-minute video with materials for delivering a three-hour introductory workshop; examples of how UWs have linked agencies with technical assistance and resources; examples of how UWs have collaborated with other funders to build agency capacity; information about commercially available software and Web-based systems agencies can use to manage outcome data (United Way of America, 2005b); and two snapshots (United Way of America, 1998; United Way of America, 2003b) of what other national organizations were doing to help local affiliates measure outcomes.

A second, integral component of the roll-out was a train-the-trainer strategy in which UW staff first received expert training on all aspects of the approach and then returned to their local areas to train a variety of persons—local agency staff, staff and volunteers from their UW, and sometimes other local funders, non-UW agencies, and local government bodies. To train the trainers, UWA developed a highly structured four-and-a-half-day training course, "Implementing a Focus on Program Outcomes." This training taught UW staff how to use a detailed training kit to build skills in applying the concepts and procedures in the UWA manual. As of March 2007, 641 persons from 281 UWs in 46 states and three foreign countries had been trained in this manner.

Third, to conduct these multi-UW training sessions, to have people available to train local UWs and agencies, and to offer in-depth technical assistance when requested, UWA recruited and trained a national consultant pool of six senior evaluators. These evaluators were experienced both in evaluation and in working with local agencies. In the first few years of the roll-out, members of the consultant pool led numerous training sessions around the country, each typically two days in length. They also furnished technical assistance to several hundred agencies under various arrangements with UWs.

A fourth important feature of the roll-out was two UWA Websites offering program outcome measurement resources. One of these is publicly available at www.unitedway.org/outcomes; it received 109,195 hits in 2006. The other is UWA's private UW-only Website. Each site is updated with new resources as they become available. There also is a private listsery for UW staff involved in program outcome measurement.

Fifth, UWA created a National Learning Project (NLP) to study efforts by seven leading-edge UWs to implement and use program outcome measurement and document the resulting impacts on UWs, local agencies, and communities. The UWs, UWA, and a consultant team formed a learning community to share experiences, tools, and lessons learned. The final NLP report (James Bell Associates, 2001) included cross-site findings and individual site reports with guidance and resource materials for other UWs.

Sixth, every other year during the first eight years of UWA's initiative, UWA held a popular national Forum on Outcomes. In plenary sessions and break-out groups, UWs were able to learn from and get practical advice from the early-adopting UWs, share lessons and solve problems with each other, and gain exposure to evaluation approaches from outside the UW system, presented by such evaluators as James Bell, Sidney Gardner, Harry Hatry, Astrid Merget, Michael Patton, James Sanders, John Seeley, Stacey Stockdill, and Joseph Wholey.

Finally, UWA's roll-out extended beyond the UW system. As its early materials were released, UWA conducted two trainings on UWA's approach and training methodology for staff of 32 national health and human service organizations. It has also made more than 40 presentations on program outcome measurement to regional, national, and executive-level workshops and conferences of national nonprofit organizations.

Our Assessment of the Roll-Out Process

As Table 2.3 shows, overall we are satisfied with several aspects of UWA's efforts to disseminate its approach to measuring outcomes:

• Several key products were developed quickly, an accomplishment helpful to those working at the local level.

- The manual's continuing popularity speaks for itself, as do the many visits to the UWA Website.
- Ongoing feedback on the training materials from sources within and beyond the UW system attest that they are methodologically sound, engage participants, and begin the process of developing products with real-world applications back home.
- The National Learning Project afforded invaluable insights into common experiences, lessons learned, and promising practices for the system. Shared tools and other resources helped accelerate implementation by other UWs and contributed to greater alignment of expectations among UWs.
- The trainings and multiple presentations for national health and human services organizations helped them know what was being asked of many of their local affiliates by local UWs. Many national organizations responded by disseminating UWA's materials and approach to their affiliates. Many also developed measurement resources for their affiliates that aligned with the UWA approach.

In sum, most people would probably agree that United Way's efforts have made it the most widely disseminated approach to outcome measurement in the nonprofit sector. At the same time, we can identify various elements of the roll-out that could have been improved.

• First, and perhaps most fundamental, was the heavy reliance on a train-the-trainer strategy to prepare UW staff to impart skills to agency staff. Because of this strategy, most local agency staff were trained not by experts or professional evaluators but by local UW staff, most of whom had themselves been trained only recently. This strategy overestimated the degree to which local UW staff could absorb all the necessary knowledge and skills in one week, remember it over time, and impart it effectively to local agencies. Being an effective trainer and technical assistance provider requires both training and consulting skills and a solid knowledge of the subject matter. Of UW staff who received UWA's training, there were many,

Table 2.3. Our Assessment of the Roll-Out Process

Strengths

Quick development of many products Very popular manual and Website Well-regarded training materials Structured learning from leading-edge UWs Dissemination via Websites, conferences UWA collaboration with other national organizations

Limitations

Overreliance on train-the-trainer strategy Too little guidance on how UWs can use outcome data No large-scale study of local agency progress since 1999 especially from larger UWs, who had solid prior evaluation training and experience. These already-knowledgeable individuals met the challenge well. But the majority of UW staff, for whom this was relatively new content, were not turned into evaluators in a week. Because staff of larger UWs had more agencies to prepare, they had more opportunity to deliver training and technical assistance and thus gain knowledge and experience. Staff of smaller UWs, with fewer agencies, had less opportunity for this learning and improvement.

This is not to suggest, however, that trained evaluators would necessarily have been more effective at training and providing technical assistance for outcome measurement. Some evaluators find it difficult to integrate UWA's approach to outcome measurement with their commitment to experimental design, or to understand the reality of local service delivery agencies. Some know evaluation but do not teach it effectively. Many UWs did engage evaluators to help with the task, and even though many were helpful each of us has heard horror stories from frustrated UWs and local agencies about professional evaluators or academics who gave guidance that ultimately led to overly complex processes or unrealistic expectations. Thus hiring an outside consultant is far from a panacea.

In retrospect, UWA's roll-out would have benefited from both increased focus on building UW staff skills in specific outcome measurement tasks and more practical guidance on identifying, engaging, and managing outside experts with the appropriate set of skills.

- Second, and as noted earlier, UWA's approach fell short in helping agencies use outcome data once collected. The same is true regarding guidance given to UWs on how they themselves could use program data. UWA conducted training on using program outcomes in funding decisions and conducted research on using outcome data to communicate UW impact. UWA did not, however, do enough to impart to local UWs the variety of other ways they could use the data and learning from programs to enhance their own work.
- A third limitation was that, even though a number of local UWs monitored the outcome measurement progress of the agencies they themselves were funding, UWA did not continuously and systematically track agency-level progress across the UW system. Early on, UWA did use grant funds for a 1999 survey of 391 agencies in six communities (discussed in the section "Outcome Measurement as Practiced by Local Agencies"), a survey that identified both the challenges these agencies encountered and the benefits they received. However, it would have been valuable to repeat this survey in later years in order to assess changes over time. One observer has pointed out the irony of an effort to encourage others to improve measurement and build knowledge not continuing to monitor agency-level progress when its work spanned more than 10 years and affected hundreds of UWs and thousands of agencies.

Implementation of the UWA Approach by Local United Ways

As mentioned earlier, the United Way system is a federation of 1,300 autonomous local UWs, each independently incorporated and governed by a local board of directors. In this respect, United Way is more akin to the real estate company RE/MAX, with its slogan "Each office independently owned and operated," than the restaurant giant McDonald's, which standardizes very precisely the exact materials and procedures each location must use. This autonomy means that each local UW decides for itself whether or not it will measure outcomes, and if so, how.

Some UWs were early pioneers of outcome measurement 15 or more years ago, and many of the lessons they learned became part of the UWA approach. One of us led the pioneering efforts in the United Way of Greater Milwaukee.² At the other extreme, many UWs have yet to measure outcomes and show no interest in beginning to do so.

Other UWs fall between these two extremes, in that they began measuring outcomes within the past 10 years—some proceeding in fits and starts—or are beginning now. In the past few months, one of us has worked with a UW just starting the process.³ In this case, the UW's long-time executive director was consistently opposed to measuring outcomes, and only her departure opened the door to moving in this direction.

Today, an estimated 450 UWs require agencies to measure program outcomes; this includes most of the 350 UWs having more than five full-time staff members. Together, these 450 UWs fund approximately 19,000 local health and human service agencies, making implementation fairly widespread.

However, simply because 450 UWs are requiring outcome measurement does not mean that each UW follows UWA's recommended approach. Exercising their local autonomy, some UWs use UWA materials and adhere closely to the suggested principles and procedures. Other UWs use UWA materials but do not follow UWA guidance regarding implementation. Other UWs use materials from other sources, materials that are not always consistent with the UWA approach. Still other UWs adopt a combination of approaches, or they sometimes develop an approach of their own.

The reality is that, even though UWA offers a uniform approach and materials at the *national* level, UWs are using a variety of approaches and materials at the *local* level. In some cases, in our view, the differences have had unfortunate results. In others, their excellent current practices reflect 10 years of learning, growth, and innovation since the time the basic UWA materials were released. In fact, one might argue that it would be disappointing if UWs had *not* moved beyond the original guidance by now.

Regardless of the approaches they use, most of the UWs involved in measuring outcomes seem committed to helping build agency capacity for the task. Many funders, in requiring their grantees to conduct measurement of various types, supply definitions, examples, or lists of variables or tools, but only some foundations have been as persistent as UWs in working to build grantees' outcome measurement capacity. In a 2002 survey of 215 UWs that had implemented or were implementing plans for program outcome measurement, 83% of 71 responding UWs reported offering both ongoing training and technical assistance to agencies, identifying it as "essential" for successful implementation (United Way of America, 2003a). Of course, the 33% response rate to this survey requires us to interpret these findings with caution.

The challenges of providing this training and technical assistance to agencies, however, have been significant. Regarding training, most UWs fund many programs; the largest UWs may fund 200 or more. This creates a need to train many people, both in agencies and in the UW. Local agencies and UWs also have staff turnover, which creates a continuing need for retraining. Although trained staff often move on to other human service agencies, thus building overall community capacity for outcome measurement, the disruptive implications of this turnover for individual agencies is significant. In addition to the immediate halt in activities when a key outcome measurement person leaves, it can be difficult to find an affordable replacement; and once found, it takes time and expense to bring the new person up to speed.

You will remember that, in the UWA approach, the primary reason to measure outcomes is to increase program effectiveness. Many UWs are encouraging agencies to use outcome data for exactly this purpose. For example, 75% of UWs responding to the 2002 survey consider a program's outcome-based learning and improvement in its funding decisions. However, very few UWs offer training, assistance, or practical guidance on the uses of outcome data for this purpose.

Although the guidance cautioned otherwise, early expectations were that UWs would move quickly to use outcome data to guide funding decisions. Perhaps they might give more money to programs with higher outcome scores and less or no money to programs not demonstrating such effectiveness. On the other hand, perhaps UWs might give low-performing programs *more* money, not *less*, in order to help them succeed. In any event, there were concerns that UWs might use the data in funding decisions before agencies had good measurement systems in place, might not take into account the very real variability among programs, and might focus on outcome data to the exclusion of other important criteria.

In actuality, as UW staff and volunteers came to understand the challenges of measurement the initial impulse for a rush to judgment subsided. In the 2002 survey referenced earlier, 95% of UWs said they were asking for outcome information from agencies, and 85% said that the information was an important factor in funding reviews. Nowhere, however, is it the only factor considered, and it appears that defunding programs solely because of outcome performance does not happen often. Several factors may be at play

here, among them reluctance to defund certain agencies (long-time partners, small grassroots organizations, large and influential agencies, agencies with influential board members, and so on) and UWs' understanding that factors other than program performance can affect results.

Despite these limitations, UWs implementing outcome measurement find that it offers several benefits. In the 2002 survey, more than two-thirds of responding UWs reported that measuring outcomes has helped them improve several facets of their operations: their accountability to donors and the community (70%); their marketing and fundraising messages (73%); their image and visibility in the community (70%); and their success in retaining, maintaining, and increasing dollars (68%). In these ways, measuring outcomes seems to produce tangible benefits for many UWs.

However, even while requiring outcome measurement of funded programs, most UWs do not use the tool for themselves. In the 2002 survey, although 66% said their UW sees the value of having an outcome focus in its own work, only 35% reported that their UW measures outcomes of internal divisions, and only 26% confirmed that outcomes are included in job descriptions and performance reviews.

Outside their funded agencies, some UWs have had an impact on outcome measurement among other public, private, and nonprofit groups. In the 2002 survey, 65% of UWs reported they were offering training to programs they did not fund. The Wisconsin state human service department asked the UW in Milwaukee to train human service agencies in that state. The UW in New Orleans conducted a two-day training program for the heads and key staff of Louisiana state departments. A volunteer at the UW in Cleveland, trained on the UWA approach, adapted UWA's manual for his small business loan program. Crime Victim Services, a local agency funded by the United Way of Greater Lima (Ohio), took to heart the UW training it received, and its outcome planning so impressed state officials that the Ohio attorney general mandated eight of this agency's local outcomes to be outcomes for the entire state. Summit Endowment, a community health foundation in rural Pennsylvania, has for several years used a variation of the UWA approach with the agencies it funds (Hendricks, 2006).

Our Assessment of Implementation by Local UWs

Table 2.4 shows our overall conclusion about implementation in local UWs: There is no one, single way outcome measurement has been implemented. Instead, there is a wide variation in efforts, ranging from those that are measured, thoughtful, and helpful to those that are rushed, ill-conceived, and nonproductive. We would like to believe there are more examples of the former than the latter, but firm data do not exist.

Such variability is inevitable in a federation such as the United Way, and to expect uniformity would be wishful thinking. One might as well

Table 2.4. Our Assessment of Implementation by Local UWs

Overall conclusion

Wide range of efforts, varying across UWs; some productive, some not

It might help UW systemwide performance to:

Update the 1996 manual

Encourage more UWs to:

Inform the UW allocations volunteers more fully

Coordinate with other local funders

Tap relevant local expertise

Help agencies obtain useful resources

Provide funds to support outcome measurement

expect each basketball coach in Indiana to teach her team the exact same offensive scheme. Given this variability, however, a few steps might improve the overall situation across the UW system:

- UWA might update its well-known manual to reemphasize important points from the original edition and incorporate important new topics, especially practical ways to analyze and use outcome data. The original writing occurred in 1996, and much has been learned in the past 12 years. Not every local UW would study the new edition in detail, but its publication would likely reaffirm UWA's commitment to program outcome measurement, reinvigorate local efforts, and bring a decade of learning and improvement to UWs, the nonprofit sector, and others. To those resisters who have believed (hoped?) that this is simply another passing phase, it would also serve as an important indicator that outcome measurement is here to stay.
- UWs might place more emphasis on informing their allocation volunteers of the challenges and limitations of outcome measurement, what it is reasonable to expect from agencies, and the criteria they should use to evaluate agency outcome results fairly. In each UW, the important funding decisions are essentially made by these volunteers, not by paid staff. Yet these volunteers have typically received less training from UW staff, largely because they are busy members of the community. As a result, they do not always know what is appropriate to expect from an agency.
- UWs might better coordinate their outcome measurement efforts with other local funders such as county governments or local foundations, and with contiguous UWs. Wise UWs do this already, and the benefits are obvious. When different local funders use their own approaches, terms, definitions, timeframes, and reporting requirements, local agencies are forced to change their procedures for each funder. This harms the agencies and undermines each funder's intentions; the agencies are obviously unable to

- truly absorb and integrate any one funder's approach. A coordinated effort, on the other hand, is more likely to represent a thoughtful implementation strategy that allows agencies to focus on substance without having to juggle competing idiosyncrasies.
- UWs might make a concerted effort to tap into relevant expertise available in the local community. Few UWs have all the necessary skills in-house, and effective implementation often requires drawing on talent from outside the UW. For example, staff and volunteers of agencies already pursuing outcome measurement, management assistance programs for nonprofit organizations, UW volunteers from businesses such as market research or public relations firms that collect and analyze descriptive data, or carefully selected and oriented university faculty or consultants could conduct training sessions and offer technical assistance to agencies. Retired business executives could help UW boards learn to use outcome data for funding decisions and in other ways.
- UWs might help agencies obtain useful resources by, for example, creating a
 directory of local resource people and organizations, developing a resource
 library of reference materials, organizing brown-bag lunches or technical
 assistance clinics on specific topics that agencies can attend when relevant,
 and encouraging agencies to work together and share experiences and
 resources.
- Finally, UWs might provide funds to help agencies develop their outcome measurement systems through direct grants, by requiring a certain percentage of each grant to be devoted to measuring outcomes, or by accepting part of the cost of developing and operating an outcome measurement system as an agency budget expense. Funds also might support the collaborative efforts of "affinity groups" of like-focused agencies to establish their measurement systems. As one expert notes, "Quality costs," and agencies cannot invest monies they do not have. Even if there are no direct costs for hiring outside experts, data collection, equipment, or services, there are certainly the indirect costs of time spent by agency staff.

Outcome Measurement as Practiced by Local Agencies

In this section, we reach the nub of the issue. UWA can develop its approach and materials, and a local UW can implement outcome measurement well or poorly. But neither of those groups is asked to identify desired outcomes, create logic models, develop measurable indicators, develop data collection instruments and procedures, gather data, analyze them, interpret the findings, and use those findings both to improve programs and for external accountability. Those tasks fall to local health and human service agencies, the place where, truly, the rubber meets the road.

Given the variability in UW implementation, it is not surprising that the experiences of local agencies also vary considerably. Some agencies readily grasp both the concepts and the recommended techniques, and for them

outcome measurement is a helpful tool. In the 1999 survey of 391 experienced local agencies mentioned earlier (United Way of America, 2000), more than three-quarters of the 298 respondents reported that measuring program outcomes helped them clarify program purpose (86%), focus staff on shared goals (88%), identify effective practices (84%), improve service delivery (76%), enhance recordkeeping (80%), communicate results to stakeholders (88%), and compete for resources (83%). Overall, 74% of agencies agreed that "on balance, implementing outcome measurement has had a positive impact on this program's ability to serve clients effectively," and 89% answered yes to the question, "Would you recommend to the director of a similar program that he/she consider implementing program outcome measurement?"

These 298 agencies also reported several difficulties in 1999. About 50% reported that measuring program outcomes overloaded their record-keeping capacity, diverted resources from existing activities, and led to a focus on measurable outcomes at the expense of other important results. More than 50% found it difficult to identify appropriate outcomes, indicators, methods, and data collection tools. About 60% reported insufficient staff time available and inadequate computer hardware and software capacity to store and manipulate outcome data. More than 60% were concerned about the cost of measuring outcomes. Some of this may have been the inevitable difficulties with any new effort, but many of these issues probably remain as concerns for local agencies.

These overall findings may be viewed as somewhat encouraging, but are they representative of all agencies today? The question is fair; these 298 agencies were surveyed explicitly because they were funded by one of six pioneering UWs involved in UWA's National Learning Project. These UWs were measuring outcomes before the UWA approach was developed—although all six UWs subsequently adopted the UWA approach—and they were atypically knowledgeable and experienced.

How, then, are those agencies faring that are funded and supported by more-typical UWs? We have no systematic data on agencies since the 1999 survey, but our collective experience from site visits, conversations, and various documents leads us to believe that the overall picture is probably mixed, with serious challenges emerging alongside productive benefits; some agencies might even report their overall experience to be negative. On the other hand, the additional years of experience with outcome measurement by individual agencies and among nonprofits in general may have allowed time for greater facility with, and more varied and effective uses of, the tool. We simply do not know in any systematic way; hence our earlier suggestion to gather local-level experiences more regularly.

Our Assessment of Agency Outcome Measurement

Just as there is no single experience shared by all UWs implementing outcome measurement, so is there no single experience shared by all agencies

Table 2.5. Our Assessment of Local Agency Outcome Measurement

Overall conclusion

Wide variability of experiences, varying across agencies; some positive, some not

UWs cannot influence

Inherent difficulty of outcome measurement Commitment of agency leadership Motivation of front-line staff

However, UWs can influence

Type and level of support given to agencies

striving to measure outcomes (see Table 2.5). Some agencies are "poster boys and girls" for the concept, and they inspire others with what can be done. Other agencies are only partially grasping the concepts and reaping only some of the benefits. Still other agencies may be frustrated at an effort they feel is time-consuming and unproductive. This variability is compounded by the fact that, for most agencies, UW funds make up only a small percentage of their overall funding; other funders typically supply more money and therefore carry more influence.

We are especially concerned that agencies are not using outcome data to derive as much benefit as they could. Too many agencies may feel they have "used" their outcome data if they submit an outcome report to their UW. As with the UWs that fund them, we suspect that too few agencies are using outcome data to increase the effectiveness of their programs. As one expert laments, the effort hasn't yet made the essential leap from outcome measurement to "outcome management," that is, from outcomes reporting to "outcomes usage." In our opinion, this may be the biggest agency challenge at the moment.

- We see four factors determining agency success in measuring outcomes, and unfortunately we believe a UW can influence only one of them. The first factor is the inherent difficulty of collecting, analyzing, interpreting, and using outcome data. Despite the motivating successes in some agencies, the task requires a specialized combination of analysis- and management-oriented skills, and few agencies have staff with these specialized skills. Guidance is increasingly available for steps such as analyzing outcome data (Hatry, Cowan, & Hendricks, 2004) and using it to benefit the program (Morley & Lampkin, 2004), and outside vendors are beginning to offer outcome measurement-related services to agencies, but agencies still need to guide these efforts.
- The second factor is the agency's leadership. All of us, and many of our colleagues, have seen the importance of top-level commitment to measuring and continuously improving outcomes. With this leadership, even an

underfunded, overworked agency can achieve surprising successes; without it, chances of success are much lower.

- The third factor is the agency's staff. Strong leadership is important, but front-line staff are the ones who actually fill out the forms and keep the records, and too often they don't recognize the value or relevance of doing so for themselves and their agency. In addition, though many agencies are quite stable, two UWs found that 28% and 33% of agency staff they had trained in outcome measurement the previous year had left their jobs.
- Our view is that UWs cannot influence to a significant degree any of these first three factors. This leaves the fourth factor: the type and level of support from the agency's UW, especially over time. We believe an agency is more likely to succeed if its UW offers them a vision of what is possible, a trusting partnership in which to discuss program weaknesses, and continuing training and technical assistance to do the task successfully. That is, it appears agencies are more successful with outcome measurement when their UWs help them measure well and make good use of the data. Agencies must engage fully and do their best, but UWs have an important role in creating the conditions for success.

Program Outcomes and Community Impact: An Evolution for the UW System

The focus of this chapter, and the priority of the early work of UWA in this area, has been on measuring program-level outcomes. But as previously noted, even in the early years of the effort UWA and its task force realized that UWs seek to affect more than the people served by UW-funded *programs*. They also aspire to affect and measure their impact on *community*-level change.

This broader aspiration has taken center stage as, starting in 2001, the UW system undertook a transformation to make community impact its central focus. In this transformation, *community impact* is about improving the lives of *community populations*—groups of people in the community—by changing policies, organizations, systems, neighborhoods, networks, and other factors that influence those populations. This contrasts with *program outcomes*, which are changes in *program clients* stimulated by program activities (United Way of America, 2007b).

UW is focusing current efforts on the challenge of achieving *community* impact, but this does not mean a lessening of the emphasis on *program* outcome measurement. UWA President and CEO Brian Gallagher has said, "In the move to community impact, program outcome measurement is table stakes: you don't get in the game without it." UWA has identified 17 examples of ways that program outcome measurement contributes to community impact in three strategic areas: demonstrating results of direct-service efforts, allowing a head start in implementing the community impact model, and helping to target community issues (United Way of America, 2006).

It has also issued guidance for and examples of using program learning to inform community impact objectives and strategies (United Way of America, 2005a).

As its early work suggested, UWA's approach continues to distinguish between measurement of *program* outcomes and *community* impact and to reiterate that simply aggregating program outcome data does not yield community impact data. This may be obvious to the evaluation community, but it is not readily apparent to some who are eager to use existing (program outcome) information to demonstrate results in the new community impact environment. UWA's recommendations and resources regarding measurement of community impact focus on tracking progress and results of both changes in community *conditions* being sought and the improvements for community *populations* that are the ultimate reason for the effort.

Measuring progress will be an integral part of the UW system for the foreseeable future. As UWs and agencies measure changes in program participants stimulated by program activities (program outcomes) and changes in a defined community population brought about by changing conditions within the community (community impact), they will continue to require access to essential evaluative skills and knowledge.

Notes

- 1. The UWA national consultant pool consists of Kenneth Fyfe, Michael Hendricks, Melanie Hwalek, Jane Reisman, John Seeley, and Dawn Hanson Smart.
- 2. Kathleen J. Pritchard led that effort.
- 3. Michael Hendricks worked with that UW.

References

Brinkerhoff, R. O. (2003). The success case method. San Francisco: Berrett-Koehler.

Davies, R., & Dart, J. (2005). The most significant change (MSC) technique: A guide to its use. Cambridge, UK: Rick Davies and Jess Dart. Full text available at http://www.mande.co.uk/docs/MSCGuide.pdf

Hatry, H. P., Cowan, J., & Hendricks, M. (2004). *Analyzing outcome information: Getting the most from data.* Washington, DC: Urban Institute Press. Full text available at http://www.urban.org/UploadedPDF/310973_OutcomeInformation.pdf

Hendricks, M. (2002). Outcome measurement in the nonprofit sector: Recent developments, incentives, and challenges. In K. Newcomer, E. T. Jennings, Jr., C. Broom, & A. Lomax (Eds.), *Meeting the challenges of performance-oriented government* (pp. 99–123). Washington, DC: American Society for Public Administration.

Hendricks, M. (2006, Nov.). *Mentoring local nonprofit agencies to manage for results*. Paper presented at the annual conference of the American Evaluation Association, Portland, OR.

James Bell Associates. (2001). National Learning Project on using program outcome data to create measurable change: Evaluation final report. Alexandria, VA: United Way of America.

Kopczynski, M. E., & Pritchard, K. (2004). The use of evaluation by nonprofit organizations. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds.), *Handbook of Practical Program Evaluation* (2nd ed., pp. 649–669). San Francisco: Jossey-Bass.

- Morley, E., & Lampkin, L. (2004). *Using outcome information: Making data pay off.* Washington, DC: Urban Institute Press. Full text available at http://www.urban.org/UploadedPDF/311040_OutcomeInformation.pdf
- Plantz, M. C., Greenway, M. T., & Hendricks, M. (1997). Outcome measurement: Showing results in the nonprofit sector. In K. E. Newcomer (Ed.), *Using performance measurement to improve public and nonprofit programs* (pp. 15–30). San Francisco: Jossey-Bass. Full text available at http://national.unitedway.org/outcomes/library/nd paper.cfm
- United Way of America. (1996). *Measuring program outcomes: A practical approach*. Alexandria, VA: United Way of America.
- United Way of America. (1998). Outcome measurement activities of national health and human service organizations. Alexandria, VA: United Way of America.
- United Way of America. (2000). *Agency experiences with outcome measurement*. Alexandria, VA: United Way of America. Full text available at http://national.unitedway.org/files/pdf/outcomes/agencyom.pdf
- United Way of America. (2003a). *Indicators that a United Way is prepared to plan for, implement, sustain, use and benefit from program outcome measurement.* Alexandria, VA: United Way of America.
- United Way of America. (2003b). *Outcome measurement in national health & human service and accrediting organizations*. Alexandria, VA: United Way of America. Full text available at http://national.unitedway.org/files/pdf/outcomes/natlorgsreportfinal.pdf
- United Way of America. (2005a). Connecting program outcome measurement to community impact. Alexandria, VA: United Way of America. Full text available at http://national.unitedway.org/files/pdf/outcomes/ConnectingPOM_toCI%20Final.pdf
- United Way of America. (2005b). Outcome measurement data management systems for agencies. Alexandria, VA: United Way of America. Full text available at http://national.unitedway.org/files/pdf/outcomes/dataMgt_0105_update.pdf
- United Way of America. (2006). How program outcome measurement contributes to community impact. Alexandria, VA: United Way of America. Full text available at http://national.unitedway.org/files/pdf/outcomes/POM-CI%203-page.pdf
- United Way of America. (2007a). *Managing agency and community-change initiative data: Guidelines for software selection*. Alexandria, VA: United Way of America. Full text available at http://national.unitedway.org/files/pdf/outcomes/Software_Guidelines.pdf
- United Way of America. (2007b). *Program outcomes and community outcomes:* What are the differences? Alexandria, VA: United Way of America. Full text available at http://national.unitedway.org/files/pdf/outcomes/ProgramOutcomesvsCommunity Outcomes.pdf
- U.S. Office of Management and Budget. (2007). *Guidance for completing 2007 PARTs*. Washington, DC: U.S. Office of Management and Budget.
- W. K. Kellogg Foundation. (1998). W. K. Kellogg Foundation evaluation handbook. Battle Creek, MI: W. K Kellogg Foundation. Full text available at http://www.wkkf.org/Pubs/Tools/Evaluation/Pub770.pdf

MICHAEL HENDRICKS is an independent evaluation consultant and trainer currently based in Oregon.

MARGARET C. PLANTZ is a director of Impact Design and Learning at the United Way of America in Alexandria, Virginia.

KATHLEEN J. PRITCHARD is the executive director of the Planning Council for Health and Human Services, Inc., in Milwaukee, Wisconsin.