

Evaluating Housing Revitalization Projects: Critical Lessons for all Evaluators

RALPH RENGER, OMAR PASSONS,
AND ADRIANA CIMETTA

ABSTRACT

The article describes the challenges faced by the authors in evaluating a neighborhood revitalization project. The challenges are placed in the context of three of the Program Evaluation Standards published by the Joint Committee on Standards for Educational Evaluation: Values Identification, Fiscal Responsibility, and Analysis of Quantitative Information. For each problem presented, the authors provide solutions that should assist all evaluators working on other similar types of broad-based community initiatives to conduct their evaluations in a more efficient and timely manner.

EVALUATING HOUSING REVITALIZATION PROJECTS: CRITICAL LESSONS FOR ALL EVALUATORS

Following World War II, a substantial need for new housing existed in the U.S. The market for single-family homes mushroomed and new suburbs blossomed in areas geographically adjacent to major metropolitan areas. Despite this growth, the need for low-income housing, especially in urban areas, continued to increase throughout the 1950s. In the 1960s, as one component of President Lyndon Johnson's "Great Society" initiative, the federal government responded to this need by building large public housing complexes under the direction of the Department of Housing and Urban Development (HUD). Though these complexes addressed the need for low-income housing, they were and continue today to be plagued with problems. They are densely populated, often built with substandard materials, and poorly maintained, often resulting in significant health problems for the residents (Krieger & Higgins, 2002). Health problems are compounded by the fact that there is a lack of access to essential health care services in and surrounding public housing structures (McAllister & Boyle, 1998). Public housing structures are also characterized by high crime (McAllister & Boyle, 1998; Popkin,

Ralph Renger • College of Public Health at University of Arizona, 1435 N. Fremont Ave., Tucson, AZ 85712, USA;
Tel: (1) 520-882-5852, ext. 18; Fax: (1) 310-206-6293; E-mail: renger@u.arizona.edu.

American Journal of Evaluation, Vol. 24, No. 1, 2003, pp. 51–64. All rights of reproduction in any form reserved.
ISSN: 1098-2140 © 2002 by American Evaluation Association. Published by Elsevier Science Inc. All rights reserved.

Olson, Lurigio, Gwiasda, & Carter, 1995), a high incidence of aggravated assault (Holzman, Hyatt, & Dempster, 2001) and high violence (Durant, Getts, Cadenhead, & Woods, 1995). Some of these public housing structures are now infamous and include such names as Cabrini Green in Chicago and Queensbridge and Marcy in New York City.

Housing Opportunities for People Everywhere (HOPE) is a HUD initiative designed to address the current public housing crises in the United States. In addition to addressing substandard housing, HOPE also provides funding for support services to address the human element, aimed at moving public housing residents toward self-sufficiency. Within HUD, this broader focus has been coined "beyond bricks and mortar." HUD has received increasing pressure from Congress to be accountable for the enormous financial investment made by taxpayers in HOPE. This pressure created an urgency to demonstrate the effectiveness of the HOPE initiative in meeting its objectives, particularly those related to the impact of support services. While many anecdotal reports are available to suggest the initiative is successful, HUD recognizes the need for empirical data demonstrating the effectiveness of HOPE (U.S. Department of Housing and Urban Development, 2000). In response to this need and to the desire for a neutral party to conduct evaluations, the most recent cycle of HOPE applications, HOPE VI, mandate that public housing authorities partner with local academic institutions to implement comprehensive plans of evaluation.

The purpose of this article is to report on the evaluation efforts for two HOPE VI projects in the City of Tucson in the Santa Rosa and South Park neighborhoods. The Santa Rosa Project is a \$14.7 million, 5-year project that began in 1996. The South Park project is a \$12.7 million, 4-year project that began in 2000. The Santa Rosa and South Park projects are addressing the needs of 200 and 87 families, respectively, who are living in public housing.

This article will highlight several challenges we faced in the evaluation of Santa Rosa and the strategies we used to surmount them. We will also describe how we applied the lessons that we learned in the Santa Rosa Project to our evaluation of South Park.

The challenges have been placed within the context of the Program Evaluation Standards published by the Joint Committee on Standards for Educational Evaluation (Joint Committee, 1994). In all there are 30 standards that "... provide a working philosophy for evaluation. They define the Joint Committee's conception of the principles that should guide and govern program evaluation efforts, and they offer practical suggestions for observing these principles" (Joint Committee, 1994, p. xviii).

Ideally, the standards are best used to guide the program evaluation from the inception of a program. However, as evaluators our attention is also naturally drawn to the standards where conflict or concern becomes evident in the evaluation process. In retrospect, as the authors reviewed the problems faced in evaluating these HOPE VI projects, it became evident that particular aspects of the Program Evaluation Standards (Joint Committee, 1994) were salient and helped frame these challenges. The Santa Rosa and South Park projects provide a convenient context through which we may illustrate some specific issues and strategies to improving the evaluation of the projects. It is the authors' hope that by sharing our experiences other evaluators currently working on similar types of broad-based community initiatives will be able to conduct their evaluations in a more efficient and timely manner.

CHALLENGE NUMBER 1: VALUES IDENTIFICATION UNCLEAR

The perspectives, procedures, and rationale used to interpret the findings should be carefully described, so that the bases for value judgments are clear. (Joint Committee, 1994, p. 31)

As the evaluation team became integrated into the operations of the Santa Rosa Project, it became evident that standards by which the success of their work could be judged were lacking. Clear standards of performance are foundational to a fair, just, and valid evaluation. Clear standards also aid evaluator credibility.

To ensure that the basis for value judgments is clear, well-written objectives are critical. "Objectives [must] state *who* is expected to experience *how much* of *what change* by *when*" (Green & Kreuter, 1999, p. 221). In the Santa Rosa Project there were many problems with the written objectives, including an absence of a standard of acceptability, failure to operationally define concepts, objectives for which there are no targeted programs, no defined end date, and no identified proximal/impact outcomes. These factors contributed significantly to a lack of clarity. The problem facing the evaluation team was how to improve clarity given that our evaluation did not begin until 2 years *after* the Santa Rosa HOPE VI project had begun and all of the objectives had already been written and were approved by HUD.

Absence of a Standard of Acceptability

To be evaluated, objectives must have a clearly stated standard of acceptability or target. This standard can be: (a) arbitrary, (b) based on scientific evidence on expected levels of improvement, (c) based on successes of similar programs elsewhere, or (d) based on state or county normative data (Fink, 1993; Green & Kreuter, 1999). While some objectives did state a standard of acceptability, most did not.

One explanation for a lack of stated targets in the Santa Rosa Project is that there was no point of reference, such as accomplishments from previous projects or perhaps from other cities, on which to base realistic expectations for change. In the absence of such information, the City was reluctant to set a target for fear of being held accountable for an unrealistically high standard (Nottingham, 1999). This is reflected in the manner in which many objectives were worded. Objectives were written with vague terms like "increase/decrease significantly," "create viable," "improve" and so forth.

For objectives using the term significantly, the possibility of applying a statistical standard of acceptability (e.g., $p < .05$) was examined. However, in most cases the sample size at baseline (e.g., number of residents, number of families, number of children) was too small to conduct statistical tests. This was further complicated by attrition, which meant even fewer cases were available for follow-up assessments of change.

Failure to Operationally Define Concepts

Many of the Santa Rosa objectives related to multifaceted and difficult to define constructs. In the absence of clear operational definitions it was impossible to determine how to measure these objectives. Significant time was spent trying to operationally define these constructs. In most cases, the operational definition did not accurately capture the complexity of the construct. For example, traffic flow was defined in terms of traffic volume, residential character as the proportion of vacant land, and cultural heritage as the ethnic mix. These definitions do not adequately capture the richness and diversity of these concepts.

Objectives for Which There are No Targeted Programs

It is not uncommon, and indeed perhaps even noble, for agencies to want to effect change in all that ails the communities in which they work. The result of this enthusiasm is that a list of

objectives is established that is too optimistic (Posavac & Carey, 1997), rather than one based upon those aspects of the community for which the agency can reasonably expect to make a genuine impact.

A good evaluation plan includes clear objectives that relate specific strategies (i.e., programs) to identified areas of need (McKillip, 1987; Posavac & Carey, 1997). In the Santa Rosa neighborhood, there were many needs. The City wrote objectives to address *all* these needs. However, many of the needs were beyond the scope of what HOPE VI funding could address. As a consequence, there were instances when there were no funded programs to target the written objectives. For example, objectives were written to reduce teen pregnancy rates and neighborhood crime, however there were no programs to target these important problems.

No End Date

None of the 67 objectives of the Santa Rosa Project included a date by which the standard of acceptability needed to be met. The City's position was that all objectives needed to be met by the end of the funding cycle. The failure to establish intermediate targets, however, made any meaningful interim assessments of progress challenging. That is, the evaluation report could not comment on the *rate* at which progress was occurring toward meeting an objective. It was limited to simply reporting on whether the final target had been met.

To illustrate this point, consider an objective which specifies that 20 families should be self-sufficient by the end of a 4-year program. Assume that, half way through the program, only seven families have achieved self-sufficiency. Is this cause for alarm? One might argue by extrapolation that a logical intermediate target would be that 10 families are self-sufficient half way through the program. If so, there is cause for alarm, and the evaluation report would highlight this as an area requiring attention. However, what if the first 2 years of the self-sufficiency program are spent recruiting participants? In this case, a more realistic intermediate target might be that five families achieve self-sufficiency by the mid point of the program. Under this scenario, the evaluation report would note that there is good progress toward meeting the objective.

No Identified Impact/Proximal Outcomes

It is relatively easy to establish and document outcomes for objectives related to tracking simple outputs, like new houses built, number of vacant lots changed, and so forth. In fact, one might argue that for many of these types of objectives that there truly are no proximal outcomes, only distal outcomes or endpoints and that they are discrete. A house is either built or not. A lot is vacant or not. A neighborhood center is built or not, and so forth.

Establishing objectives related to assessing change in more complex constructs (i.e., those involving human beings), such as perceptions, attitudes, and behaviors, are more difficult to define and to set standards of acceptability. Objectives related to the assessment of antecedent conditions were completely lacking from the Santa Rosa Project. This kind of evaluation, referred to by some as a black box evaluation (Cronbach, 1982; Posavac & Carey, 1997), has limited utility. Cronbach is rather pointed in his perception of such evaluation. "Some evaluators mistakenly adopt a black box, input-output analysis that tests a program over its longest possible reach and relegates evidence on intermediates to an appendix" (Cronbach, p. 223). He goes on to suggest that a "series of short-reach evaluations" can help to identify influential variables or contributing factors that may provide information to help identify both successful

and unsuccessful aspects of the program. In other words, intermediate data may prove to be more informative to the process of the project than data presented in final compilations.

There are numerous planning and evaluation frameworks that facilitate the understanding of antecedent conditions and the identification of intermediate objectives/data, including Precede-Proceed (Green & Kreuter, 1999), Model for Health Education Planning (Ross & Mico, 1980), Comprehensive Health Education Model (Sullivan, 1973), and the ATM Approach (Renger & Titcomb, 2002).

The Santa Rosa objectives only included reference to changing distal outcomes. This is not unexpected, as these are the outcomes from which concern and call to action often originate. High crime, high drop out rates, high unemployment, and so forth are examples of issues that incite a call for action. Thus, it seems logical to want to see change in these outcomes. The problem is that distal outcomes, or outputs, are often *symptoms* of a problem. There are a host of other antecedent conditions that must be impacted before change in the distal outcome can be expected or observed.

For example, one of the Santa Rosa objectives targeted reducing high school drop out rates. The assessment of changes in drop out rates is relatively straightforward through an analysis of school records. However, effecting change in these high school rates requires a much more extensive understanding of the factors that contribute to a student dropping out. One theory is that students who drop out lack self-esteem. If this is true, then one needs to understand the process by which self-esteem (itself a symptom) is built. For example, Renger, Kalbfleisch, Smolak, and Crago (1999) presented a theoretical framework that self-esteem is built when a youth is provided with opportunities to experience mastery in various life skill areas that they value. Thus, an evaluation plan for youth programs targeting the high school drop out and self-esteem issue might include an assessment of the impact on proximal factors, like improvements in mastery. Not only are such assessments more meaningful, for they delve into the black box, but changes in these outcomes are more likely to occur during the course of the HOPE VI program than those that are more distal.

Immediate Strategies Applied to the Santa Rosa Project

As the objectives had already been established and approved by HUD there was little that could be done in terms of writing better objectives. Further, to point out that the objectives were poor or to begin rewriting objectives that City personnel spent significant time writing would only undermine our credibility and destroy trust. Therefore, the focus of the evaluation team became to *demonstrate* the problem with the manner in which objectives were written. During regular scheduled meetings we would update the group with the status of our efforts to evaluate the list of objectives. Each meeting we would discuss progress we were making toward evaluating many of the objectives. Although there were multiple problem objectives, we purposively chose not to discuss more than one at each meeting. For example, during one meeting we noted to the group that collecting data on teen pregnancy rates would be problematic and suggested that if the objective were altered to assess teen birth rates then data would be forthcoming from the county and state. During subsequent meetings we systematically addressed each problem objective such as those related to defining traffic flow and residential character (discussed above). By spreading the discussion of problem objectives over time and by demonstrating the barriers to evaluating each objective as they were originally written, we were able to negotiate changes to some objectives in a professional and non-threatening manner.

Strategies Applied to South Park Based on Lessons Learned in Santa Rosa

Several steps for improving South Park objectives were implemented. First, a standard of acceptability was included for all objectives. Terms like “significantly” have been dropped in favor of quantifiable, measurable objectives such as the percentage improvement or a set number of clients to be served. Previous work conducted in Santa Rosa and a review of literature of similar programs helped set realistic targets. Because the requirement for evaluation is relatively new to the HOPE VI initiative, a frame of reference for setting targets is just beginning to emerge. For example, in the Santa Rosa Project, one objective was, “To increase the number of public housing residents successfully completing ESL programs.” Based on the Santa Rosa experience, this was improved for the South Park project to read, “50% of those who enroll in ESL certification will progress to the next level of ESL certification by the end of the 5-year period.”

Second, significant time was spent on the South Park application operationally defining complex constructs. If affordable indices could not be found to measure the construct or if the indices did not accurately capture the complexity of the construct, then the objective was dropped. It is important to note, however, that this did not mean that the program or initiative itself was dropped, but that efforts to evaluate this impact of the program were not pursued.

Third, a concerted effort was made in the South Park application to be realistic. This meant ensuring that strategies were clearly linked to identified needs and that objectives were only written for areas targeted by these strategies.

Fourth, the basic principles of writing objectives were applied. Each objective was reviewed to ensure it contained reference to who, what, how much, and when (Green & Kreuter, 1999). Equally important, we have identified intermediate standards of acceptability. Graduated standards of acceptability related to expected progression have been set for the midpoint of the project and the end of the project, at 3 and 5 years, respectively. The fact that the intermediate targets in South Park fall in the middle of the funding cycle is coincidental. The chronological midpoint of a funding cycle is not necessarily the best time to collect data on intermediate targets. The timeline for assessing intermediate targets should be set such that there is enough time to act on the results of the evaluation. This is the essence of the Utility Standard of Report Timeliness and Dissemination (Joint Committee, 1994). The experts implementing the programs, not the evaluator, are best able to establish this timeline. Once the intermediate targets were set, we were able to meaningfully link the evaluation reports to these time frames collecting data at baseline, 3 and 5 years.

It is important to note, however, that establishing intermediate targets does not preclude collecting data annually. With respect to working with community agencies, out of sight *does* equal out of mind. Therefore, it is important for evaluators to develop a routine with those agencies and organizations that provide data. It is our experience that making contact on at least an annual basis serves to strengthen the routine and allows evaluators to track shifts in organizational structure (e.g., new hires, new databases) to ensure that needed data will still be collected and made available.

Finally, we examined the feasibility of available methodologies (e.g., Renger & Titcomb, 2002) to develop a more in-depth evaluation plan, one that identified and assessed proximal outcomes (i.e., within the black box) as opposed to simply documenting changes in distal outcomes (i.e., outside the black box). The immediate barrier we faced was the lack of resources to implement the evaluation plan. The evaluation budget depended largely on staff collecting data. However, collecting data for the purpose of an in-depth evaluation meant added responsibilities to existing workloads for staff.

Two approaches are being pursued concurrently. First, the evaluators are articulating the dilemma to the funding agency (HUD). In correspondence with the funding agency the evaluators are pointing out that they agree with the need for more meaningful data on impact outcomes, that this data is only forthcoming from partner agencies in the community, and that additional funding is needed to provide these agencies with the technical assistance to produce the needed data. Second, more in-depth plans of evaluation are only being explored for a few flagship projects within the South Park HOPE VI project. The authors concur with the sentiments of Posavac and Carey (1997) when they state "It is better to carry out an evaluation with modest aspirations that one can trust than to plan an ambitious project that can only be done poorly given the limitations of resources available" (p. 63).

CHALLENGE NUMBER 2: ACHIEVING FISCAL RESPONSIBILITY

The evaluator's allocation expenditure of resources should reflect sound accountability procedures and otherwise be prudent and ethically responsible, so that expenditures are accounted for and appropriate. (Joint Committee, 1994, p. 121)

In striving for fiscal responsibility it is recommended that evaluators "be frugal in expending resources for evaluation" (Joint Committee, 1994, p. 122). The problem the evaluators faced in these projects was that the budget was already established and, as demonstrated below, significantly under-funded to assess all of the objectives of interest.

At the National Conference for Evaluation it was announced that evaluation budget for new awardees ranged from \$25,000 to over \$200,000 per year (U.S. Department of Housing and Urban Development, 2000). Certainly there is a positive correlation between the size of the budget and the size and length of the project. Typically larger projects require a proportionately smaller allocation for evaluation than smaller projects (Posavac & Carey, 1997). However, only 0.27 and 0.22% of the budget was allocated for evaluation in the Santa Rosa and South Park projects, respectively. Cummings (1992) suggests that $7 \pm 3\%$ be budgeted for evaluation. Given these guidelines, it is clear that evaluation is significantly under funded for both projects. As a result, the problem for the evaluators was how to complete the evaluation of the stated objectives with limited funds.

Immediate Strategies Applied to the Santa Rosa Project

One strategy being used to cut costs is to limit primary data collection. This begins by working with community agencies to determine what types of data are currently being collected and determining the extent to which it might be used to assess the objectives. As a result of an analysis of available secondary data sets, gaps in data needed to assess objectives will be identified. These gaps can then be met by subsequent primary data collection directed specifically toward program objectives for which no secondary data sources have been identified.

One limitation of this strategy is that the data being collected by agencies is for purposes other than the evaluation of the Santa Rosa Project and therefore may not be useful in assessing objectives. For example, one objective pertains to increasing the number of residents receiving English as Second Language (ESL) certification. From the Santa Rosa Project, we learned that the local adult education department tracked this data, but that individuals can only successfully move up levels and that there is no fixed endpoint such as certification.

Thus, the objective was rewritten in terms that are consistent with the type of data available. The second limitation in working with data provided by community agencies is that often the instruments are only administered one time, usually at intake. This makes the assessment of change difficult. Therefore, we are working with agencies to identify relevant questions on instruments they are using and request that they reassess clients at regular intervals. A third limitation is that the HOPE VI project is not a priority for many of the community agencies. The evaluators' experience in Santa Rosa was that shifts in inter-organizational priorities resulted in data no longer be collected at regular intervals or perhaps being dropped all together.

The second strategy to control costs is to stretch other budget line items, in this case HOPE VI labor dollars. In Santa Rosa, this was accomplished by hiring and training public housing residents to collect and enter data. Public housing residents were paired with public health graduate students to complete door-to-door surveys of public housing and neighborhood residents. These pairs also worked together to double enter the survey data, another critical component of the Standard for Systematic Information (Joint Committee, 1994). Employing resident-student teams is not only cost-effective, but it consistent with the spirit of the HOPE VI (i.e., assist interested public housing residents in developing new skills). Another advantage of employing residents is that funding for their salaries is available in other HOPE VI budget line items, thus limiting the strain on the small evaluation budget. Because of the success of this approach in Santa Rosa, we will be employing a similar strategy in South Park.

Strategies Applied to South Park Based on Lessons Learned in Santa Rosa

When the evaluation process is not included as an initial component of a program, the fiscal needs for a sound evaluation can present a genuine challenge. Likewise, the needs for particular data can also present a challenge if that data has not been consistently recorded for the duration of a program. One strategy being used in South Park is to develop and sign Memorandums of Understandings (MOUs) with agencies delivering services that assess more proximal outcomes that are precursors to improved quality of life. MOU's are normally limited to an agreement of what services will be provided in exchange for direct funds. The concept of MOUs is the same as that of the Formal Agreements Standard (Joint Committee, 1994) and includes many of the elements described in the guidelines. In addition to including specific reference to covering costs associated with evaluation, MOU's help to ensure accountability and delivery of data even if the agency experiences change in personnel in critical decision making positions or shifts in agency priorities.

Another strategy being examined to control costs in South Park is the use of unobtrusive measures (Webb, Campbell, Schwartz, & Sechrest, 1966). Using unobtrusive measures allows evaluators to be true to the Cost Effectiveness Standard by minimizing disruptions and, because they are typically less expensive to collect, conducting evaluations as economically as possible (Joint Committee, 1994). For example, during recent interviews South Park residents were asked about neighborhood safety and how they would know whether there neighborhood was a safer place. One of the residents replied, "If my neighborhood were safer, I would let my children play outside after dark." Thus, in the evaluation of safety in South Park we will examine the feasibility of collecting observations of children playing after dark as opposed to the more lengthy door-to-door survey used to assess perceptions of safety in Santa Rosa.

CHALLENGE NUMBER 3: PROBLEMS IN ANALYSIS OF QUANTITATIVE INFORMATION

Quantitative information in an evaluation should be appropriately and systematically analyzed so that evaluation questions are effectively answered. (Joint Committee, 1994, p. 165)

Many of the objectives of interest to the City focused on assessing change using quantitative data. Of course, to obtain an accurate assessment of *change*, baseline data is needed *prior* to the start of a project. The implementation of the Santa Rosa Project began on or about June 30, 1996. The contract for the evaluation was not finalized until February of 1999. Thus, the challenge was to determine whether baseline data could be obtained retrospectively for several of the objectives.

A second consideration in the analysis of quantitative information was selecting the appropriate unit of measurement (Joint Committee, 1994). Neighborhood level change is particularly important to the City because there is a desire to understand the impact that changes to public housing and supporting infrastructure has on the broader neighborhood. Thus, many objectives required measurement at both an individual and neighborhood level. For example, objectives were written assessing the impact of HOPE VI on public housing residents (i.e., individual level) and on residents for the Santa Rosa neighborhood (i.e., neighborhood level) regarding improving employment, educational levels, and reducing teen birth and crime rates. Most databases proved problematic for assessing neighborhood level objectives. Although data is stored at an individual level (i.e., by name, date of birth), there was no way to match available data to the geographical area of interest. Therefore, data for residents living in the neighborhood had to be extracted using other levels of data organization such as census tracts and zip codes. The problem with this approach is that the manner in which the data is organized does not align with the geographic boundaries of the neighborhood. The result of this mismatch is that estimates of impact are either too liberal or too conservative, depending on the extent to which data organized in the secondary databases overlaps with the geographic boundary of the neighborhood.

Selecting the appropriate unit of measurement was also salient with respect to measuring change at an individual or family level. Many objectives established a standard of acceptability in terms of a percentage of individuals or families. The challenge posed to evaluators is similar to that experienced by census takers (Brownrigg & de la Puente, 1992) in that there was uncertainty as to the exact number of individuals or families living in a household. Records maintained by the City as to the number of residents living in public housing were at odds with data collected by door-to-door survey. Some of the discrepancy can be explained by the failure to regularly update the database with alias names, new births, extended family members or evictions. Also, we learned that the database maintained by the City was not historical. That is, new information being entered into the system replaced existing information. This made tracking very difficult.

Immediate Strategies Applied to the Santa Rosa Project

Given that baseline measures had not been incorporated into the Santa Rosa Project when it was initiated, a decision had to be made regarding what constituted optimal, if not best practice. Given these circumstances, the methods we employed were adaptive and sought to deliver the most substantive and relevant data possible for evaluation of Santa Rosa (Patton, 1997).

The first tactic was to examine whether secondary data sources were available to establish baseline measures. Secondary data sources are archival (historical) and, if available, might date back to the start of the project. However, it was unrealistic to expect that secondary databases would be available to assess all of the objectives. Thus, the strategy was to *first* exhaust potential secondary data sources for as many objectives as possible and *then* use primary data collection techniques for objectives for which no secondary data source could be identified.

There are several advantages to this strategy. First, secondary data sources are less costly because the data is already being collected and entered by a third party. Second, the practical implications of identifying secondary data sources meant the amount of information that needed to be collected using primary data sources would be significantly reduced. Requiring less primary data collection also meets the evaluators ethical and financial responsibility to respect participants' time by only collecting data needed to complete the evaluation (Joint Committee, 1994; Kidder, 1981).

The search for secondary data sources led to the doorsteps of numerous agencies, each with differing levels of sophistication in data management. When requested, limited technical assistance (i.e., within budget constraints) was provided to agencies to improve current methods of data collection. This had a dual effect. First, it helped establish a good rapport with these agencies. Second, it provided an opportunity to modify the current data collection practices in favor of a method that enabled data specific to the HOPE VI objectives to be more easily extracted in the future.

It was the evaluators' experience that the most useful databases are those maintained by the tax assessor's office and the Police Department. Data are maintained at the parcel level in each of these instances. This is important as it allows the evaluator to define and collect data on the exact geographic boundaries of the neighborhood. One limitation in using this information is that the tax assessor is often backlogged, resulting in a lag in the time between when a change has been made and when it is reported. Another limitation is that the evaluator must become competent using the Geographical Information System (GIS) to extract and plot this data (Renger, Cimetta, Pettygrove, & Rogan, 2002).

Having exhausted the available secondary databases, the evaluators then attempted to fill in gaps for neighborhood level assessments by collecting data from neighborhood residents using a door-to-door survey. Significant resources were invested in staff training and salaries. A 1-month intensive effort only produced 63 surveys, representing 12.5% of the estimated households in the area. One possible explanation for these low response rates is the fear of outsiders, which is similar to that documented by Brownrigg and Martin (1989) as a reason for undercounts by the census bureau in low-income neighborhoods.

As a result of the low return rate, the external validity of the findings to the broader neighborhood is limited (e.g., only included respondents who were home during standard working hour). Safety issues did not permit conducting interviews after dark. Further, labor restrictions did not permit the hired personnel to collect data on the weekends.

Strategies Applied to South Park Based on Lessons Learned in Santa Rosa

In an attempt to gather more representative neighborhood level data we are exploring the feasibility of employing a central location intercept technique (Green & Kreuter, 1999) by identifying community events where large numbers of residents are likely to gather. Based on the experience from Santa Rosa this approach is likely to be more successful early in the project, when interest and enthusiasm for the new HOPE VI initiative is high. However,

records from the meetings at Santa Rosa show that attendance at community meetings dropped a few years after the project began, and attendance at neighborhood association meetings has always been low. This poses a few problems from an evaluation standpoint. First, the types of people attending community gathering may vary significantly over time. Thus, it will not be possible to assess whether any observed change is due to shifting perceptions or the shifting characteristics of the sample of neighborhood residents. Second, dwindling attendance might mean that the sample size available later in the project timeline may be too small to draw comparisons. Because of these potential problems, the option of assessing neighborhood impact by employing a qualitative method is being explored. Key stakeholders including neighborhood association presidents, business owners, church representatives and a few volunteer residents will be interviewed to determine their perceptions of neighborhood level changes. This approach achieved a better balance as recommended by the Standard for Analysis of Quantitative Information (Joint Committee, 1994).

Another strategy being considered to address the need for neighborhood level data in South Park is to limit objectives to those where it is reasonable to expect that the HOPE VI initiative could have an impact. As noted earlier, it is important to temper optimism and focus on changes that can realistically be expected, given the focus of the program.

In assessing individual versus family level change it is critical to maintain an accurate database of the number of individuals and residents in the target population. The database for South Park will be updated monthly. Every effort will be made from the onset to obtain the full names and aliases of the public housing residents being targeted by HOPE VI. We will work with staff and residents to help ensure that the initial registry is as complete as possible. On a monthly basis, the registry will be updated with information about residents who may have been evicted, relocated, moved, or who have had a recent addition to the family. This will be done in concert with a few key informants of the public housing community. The database will also contain the date of birth of each member so that data pertaining to youth (e.g., high school drop-out, teen births, participation in youth programs) can be easily tracked. As noted above, however, the database is not historical. That is, as information is updated each quarter it overwrites data from the previous quarter. This necessitates performing "data dumps" each quarter and developing strict rules detailing cut-off dates for data entry.

SUMMARY

During the course of an evaluation situations arise that may require ongoing changes to the methods being used (Patton, 1997). Though we may attempt to extract static quantitative measures for evaluation, the evaluation process is itself dynamic. More than 25 years ago, McTavish, Brent, Cleary, and Knudsen (1975) studied the problem of research implementation in a variety of projects funded through Health, Education and Welfare (HEW) agencies. Their report noted:

Our primary conclusion from the Predictability Study is that the quality of final report methodology is essentially not predictable from proposal or interim report documentation. This appears to be due to a number of factors. First, research is characterized by significant change as it develops over time. Second, unanticipated events force shifts in direction. Third, the character and quality of information available early in a piece of research makes assessment of some features of methodology difficult or impossible. Finally there appear

to be important and meaningful differences between raters in their professional judgements about the project's methodology. (McTavish et al., 1975, p. 62-63).

In essence, McTavish and his colleagues concluded that evaluation of most programs must be responsive to changes and fluctuations in those programs. This often proves to be a primary challenge to effective evaluation. Even though these observations were made almost 3 decades ago, they still clearly applied to the Santa Rosa program.

Santa Rosa HOPE VI is a multifaceted, longitudinal project that presented several challenges for evaluators. The solution for many of these problems lies in involving the evaluators early in the project, during the planning process (Weiss, 1995). The value of doing so is being realized in the new South Park initiative. Clear objectives are being written; ones with stated standards of acceptability, which are linked to strategies that target areas of need, and that can be measured. By involving the evaluation team before the groundbreaking ceremonies, the South Park evaluation will have more complete baseline data necessary for documenting change. Data collection will be an integrated component of the program rather than an addendum to the program. A comprehensive tracking system is being implemented at the onset of the project to track what is a moving target (i.e., residents living on-site, relocated, moved out). The early involvement of the evaluators will provide a more clear and organized plan of evaluation, which will result in significant cost-savings.

The evaluation process for the Santa Rosa and South Park programs of HOPE VI illustrates the usefulness of the Program Evaluation Standards both in terms of designing evaluations and in diagnosing problems that may develop later in the evaluation process. Using the Program Evaluation Standards as a diagnostic tool holds great potential for those evaluations that are undertaken *after* a program is already underway. The Standards can be engaged to focus on specific challenges that might be hindering effective evaluation of a program.

The authors have illustrated some specific challenges that constrained the evaluation process for the Santa Rosa program. Through innovative use of resources and data, the authors were able to optimize the evaluation process and satisfy the evaluation needs for the program. Evaluation of the Santa Rosa program provided some lessons for the evaluation design for the South Park program.

The innovations developed for evaluation of the Santa Rosa program are reflective of Patton's (1997) Utilization-Focused Evaluation. Patton points out that utilization-focused evaluation is not a formal model or template for evaluation, rather, it is an approach to the evaluation process. Key to this approach is utility, and the requirements are identification of the end users of the evaluation information and the ability of the evaluators to respond "actively, reactively and adaptively" to the information that is produced through the evaluation process. The ability to respond "adaptively" to various aspects of the evaluation process proved to be crucial to evaluation of the Santa Rosa program. This adaptive response also proved valuable in laying the foundation for the evaluation process eventually developed for the South Park program.

It is possible and, under certain constraints, even advisable to take an adaptive approach to the evaluation process while employing the Program Evaluation Standards as a guide. Adaptability allows the evaluator to respond to the unique challenges and situations presented by a specific program. Use of the Program Evaluation Standards helps to ensure that the evaluation process has integrity and is of the highest quality possible.

Our experience serves to remind all evaluators about the importance of educating those with whom we interact at local, state, and federal levels about the role of evaluators. Evaluators

must educate these stakeholders about the kind of resources needed to gather different kinds of evaluation data, especially data related to assessing impact. Evaluators must also act as innovators by providing cost-effective alternatives to implementing an evaluation plan, such as training residents to collect and enter data. Above all, evaluators must educate these stakeholders about the necessity of engaging an evaluator early in the project, in the planning phase.

ACKNOWLEDGMENTS

The authors would like to thank Tom Kingsley from the Urban Institute and the Community Services Department of the City of Tucson for their contributions to the article.

REFERENCES

- Brownrigg, L. A., & de la Puente, M. (1992). *Sociocultural behaviors correlated with census undercount*. Presented at the American Sociological Conference.
- Brownrigg, L. A., & Martin, E. A. (1989). *Proposed study plan for ethnographic evaluation of the behavioral causes of undercount*. Paper prepared for the Census Advisory Committee of the American Statistical Association and the Census Advisory Committee on Population Statistics at the Joint Advisory Committee Meeting, April 13–14, Alexandria, VA.
- Cronbach, L. (1982). *Designing evaluations of educational and social programs*. San Francisco: Jossey-Bass.
- Cummings, O. W. (1992). Evaluation: How much is enough? *Measurement and Evaluation in Counseling and Development*, 24, 150–154.
- Durant, R. H., Getts, A. G., Cadenhead, C., & Woods, E. R. (1995). The association between weapon carrying and the use of violence among adolescents living in and around public housing. *Journal of Adolescent Health*, 17, 376–380.
- Fink, A. (1993). *Evaluation fundamentals: Guiding health programs, research, and policy*. Newbury Park: Sage.
- Green, L. W., & Kreuter, M. W. (1999). *Health promotion planning: An educational and ecological approach* (3rd ed.). Mountain View: Mayfield.
- Holzman, H. R., Hyatt, R. A., & Dempster, J. M. (2001). Patterns of aggravated assault in public housing: Mapping the nexus of offense, place, gender, and race. *Violence Against Women Special Issue: Violence Against Women in Public Housing*, 7, 662–684.
- Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards* (2nd ed.). Thousand Oaks, CA: Sage.
- Kidder, L. (1981). *Research methods in social relations* (4th ed.). New York: Holt, Rinehart and Winston.
- Krieger, J., & Higgins, D. L. (2002). Housing and health: Time again for public health action. *American Journal of Public Health*, 92, 758–768.
- McAllister, L. E., & Boyle, J. S. (1998). Without money, means, or men: African American women receiving prenatal care in a housing project. *Family and Community Health*, 21(3), 67–79.
- McKillip, J. (1987). *Need analysis: Tools for human services and education*. Beverly Hills, CA: Sage.
- McTavish, D., Brent, E., Cleary, J., & Knudsen, K. R. (1975). *The systematic assessment and prediction of research methodology: Advisory report* (Vol. 1). Final Report on Grant OEO 005-P-20-2-74, Minnesota Continuing Program for the Assessment and Improvement of Research. Minneapolis: University of Minnesota.
- Nottingham, E. (1999, Spring). Personal communication.
- Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Thousand Oaks, CA: Sage.