CEHC
Cultural Ecology of Health and Change

The Cultural Ecology of Health and Change (CEHC) Working Papers Series
Working Paper

Traditional Approaches to the Evaluation of Community Based Interventions: Strengths and Limitations

By

Tony L. Whitehead, Ph.D., MS.Hyg.
Professor of Anthropology and Director, CuSAG

Department of Anthropology
University of Maryland
College Park , Maryland 20742

September 15, 2002.

BACKGROUND: Outcome Evaluation and the Dominance and Strengths of Quantitative Methods in the Evaluation of Community Intervention

In this document, community based interventions (CBIs) are defined as *programs designed to address health or social problems within a community through the significant participation of members of that community in the planning, implementation, and evaluation of such efforts.* CBIs are a type of community intervention, but not all community interventions are community based. Every year hundreds of millions of dollars are awarded in sponsorship of community interventions to achieve the particular social or health issue that they propose to address. With the growth in the funding of such efforts came an increasing demand for the assessment of whether such programs actually achieved what they proposed. The increasing trend in demands for CBI evaluation can be traced back to the 1960s and 1970s when there evolved a "national interest in [the] evaluation of public health, medical care, education, and social services program." (Windsor et al 1994:6).

During the early years of this interest in program evaluation, so call "quantitative" methods (e.g., *survey and measurement methods*, *hypotheses testing*, large *randomized samples*, *reliable and valid instrumentation*, *experimental designs*, etc.) would dominate the field of evaluation as it did in most social science research. We continue to strongly rely on such methods because, as argued by some, they are the most reliable tools that we have for achieving maximum possible reliability and validity in exploring hypothesized relationships between factors associated with human behavior and health status, and programs designed to change such behaviors and status. As such, these methods were thought to be the most reliable and valid way to demonstrate whether the large public and private investments in such programs were justified.

Quantitative methods were thought to be our strongest weapons for assuring *external validity* (the degree to which project outcomes can be generalized to a larger population than that exposed to the interventions) and *internal validity* (the degree to which the outcomes attributed to the interventions, were in fact due to the interventions. External validity is irrelevant without first overcoming the threats to internal validity. According to Windsor and colleagues (1994:149-150), and others[1], the following are the primary threats to internal validity posed in community interventions: :

1. **History,** or unanticipated events that may occur in the intervention or the community during the life of the intervention that may have some impact on intervention strategies or outcomes;

2. **Maturation,** changes in the program or the participant through such maturation processes as gaining unexpected skills or knowledge, unanticipated influences by others, natural maturation through the passage of time, unanticipated changes that occur through developmental phases;

3. **Testing or Observation,** or the effect of taking a test, being interviewed, or being observed on outcomes. the significant, unplanned

---

[1] Campbell and Stanley 1966; Cook and Campbell 1983. .

4. **Instrumentation,** or bias resulting from changes in the "characteristics of measuring instruments, observation methods, or data collection processes, that is factors affecting the reliability and validity of the data collecting instruments;

5. **Statistical Regression and Artifacts**, or the selection of an experimental, control, or comparison group on the basis of an unusually high or low level of characteristics that may yield changes in subsequent measurements;

6. **Selection,** or the identification of a control or comparison group not equivalent to the treatment group because of demographic, psychosocial, or behavioral characteristics.

7. **Participant Attrition,** or bias due to non-random assignment and/or excessive attrition (10% or more) from the experimental or control group.

8. **Interactive Effects**, or any combination of the preceding seven factors.

Experimental and Quasi-Experimental Project Designs

In designing community interventions to overcome such threats to internal validity, *experimental designs* are considered the most powerful because they are structured for this particular purpose. Experimental designs are research plans that *randomly* assign study participants to both an *experimental group* (those exposed to the intervention) and to a *control group* (those not exposed to the intervention), and the observation of both groups, usually *prior to,* and always *after the application* of the intervention. (Windsor et al, 1994:148). These procedures (random assignment of study participants, experimental and control groups, and pre-and post-test observations) are considered to be the most reliable research protocols for demonstrating with the greatest level of *measurable precision* that program strategies have been the principal contributors to proposed program outcomes, rather than to some other factor within the research environment. Experimental designs were thought to enhance the internal validity.

Conducting experimental designs in community settings, however, are frequently not possible because of the difficulties of achieving truly comparative experimental and control communities, and of randomly assigning study participants in both types of communities. As such, *quasi-experimental* designs are attempted more frequently than true experimental designs in the evaluation of CBIs. Quasi-experimental designs are considered the second most powerful technology that we have in controlling for rival factors, or *internal validity*, because they have the same attributes as experimental designs (experimental and control groups, and pre- and post-test observations), except for the randomness of participant assignments.

Early Approaches to *Formative*, *Process*, and *Impact* Evaluations to Overcome Problems in Traditional Approaches to Evaluating Community Interventions

In utilizing quasi-experimental or true experimental designs in early (since the 1960s) approaches to the evaluation of CBIs, emphases were placed on simply assessing whether the proposed or

hypothesized outcomes were achieved through the program strategies or interventions used. Little attention was given to understanding *why* certain intervention worked or did not work. In such orientations, most CBIs were evaluated as *failures*, or at least not having achieved the level of success proposed or desired by those implementing and/or sponsoring such programs. There were those however who begin to question whether such "failures" were due to the weakness of the interventions, or in their implementations, or were they due to inadequacies in the evaluation designs. This led to further development of designs appropriate to the evaluation of CBIs. There emerged the demand for evaluation designs that not only focused on the proposed *outcomes* of a project, but attention should be given to the evaluation of various stages of project implementation. That is attention should not only be given to *outcome (or summative) evaluation*, but also to *formative evaluation* carried out during the developmental phase of an intervention to assess the feasibility for project success, and *process evaluation,* carried out during the mid-term of a project. Summative evaluation included both outcome evaluation, or the evaluation of *final* project outcomes, and *impact evaluation,* or the evaluation of *short term* outcomes

Even while formative, process, and impact evaluation techniques were added to the evaluation of community intervention (along with outcome evaluation techniques), there were still problems in terms of addressing the various factors that could affect the implementation of community interventions, and thus the desired or proposed outcomes. The central problem was that in early approaches to the use of formative, process, and impact evaluation strategies, quantitative approaches (surveys or structured interviews as primary forms of data collection) continued to dominate. Formative and impact evaluation strategies *still tended to measure outcomes*, only now measuring them in project phases, as well as at the end of the project (outcome evaluation). That is, outcomes were developed and measured for the development phase (formative evaluation), at midterm (impact evaluation), and final (outcome evaluation). Process evaluation was associated with formative and impact evaluation in that it consisted of assessing whether proposed program procedures were followed, and using the findings from the formative and impact evaluations as demonstrations as to whether procedures were being followed. The problem at this point was that little attention was being given to the *complexity of the human dynamics* involved in the planning, implementation, and evaluation of community interventions that could affect both implementation and project outcomes.

The planning and implementation of community interventions were becoming increasingly complex as increasing emphasis was placed on community participation and the development of partnerships or coalitions. Not only were those involved who were responsible for conceptualizing, designing, sponsoring, and implementing community interventions, but there were also representatives from numerous community groups, many of whom might have different agendas or perceptions about what the project was about, and what it should try to achieve. As Goodman (1998:38) states:

> "Community initiatives often have broad and multiple goals that require complex interactions, multiple levels of intervention that occur in diverse settings, and multiple change strategies that require broad and repeated exposure….The responsibility for program implementation is likely to be shared by many community organizations and individuals, ;…Community health promotion may center on elusive concepts like

community empowerment, ownership, leadership, and capacity building that are difficult to measure….[Moreover]….Community programs often take many years to produce results….[and may have] competing interests of different program stakeholders…"

Further Developments in the Evaluation of Community Interventions: Qualitative Methods Become Part of the Evaluator's Tool Kit

It would become apparent to a number of those involved in the sponsoring, planning, implementation and evaluation of community interventions that the dependence on quantitative methods alone could not address the possible impacts on project planning and implementation that come from such complexities. As a consequence *qualitative approaches* to evaluation of community interventions were adopted by an increasing number of evaluators to help understand and address these complexities, if not address them. Such qualitative methods as open-ended interviews (as compared to closed coded responses as practiced in quantitative surveys), *group interviews* (rather than the focus on individual interviews emphasized in quantitative approaches so as to prevent environmental bias), *indepth case studies* (as compared to the emphasis on large samples of aggregated individuals found in quantitative studies), *"snowball" and other less rigid sampling* techniques, (as compared to the randomized sampling emphasized in quantitative approaches), *observations (*of stakeholder interactions, implementation organization activities, intervention sessions, and activities in the community targeted by the intervention),and *content analysis* (which focuses on the analysis of narrative themes as compared to the focus on numerical analysis found in quantitative studies) became part of the evaluation scene.

While data collection methods using quantitative approaches are quite rigid in their design and implementation so as to prevent the threat to internal validity of *instrumentation* (the same instruments, observation methods, and other data collection processes are administered to each study participant without change during project duration), qualitative methods emphasize *flexibility* (to address the complexity of community interventions discussed above), and *responsiveness to changing local needs and conditions….and the ability to adapt in the middle of the assessment process*. (Goodman 1998:38). The more open ended and flexible approach of qualitative methods also allowed program evaluators to better assess the meanings or interpretations that different project stakeholders and community residents had about the project and its possibilities for the community, and about the various socio-cultural contexts (political, economic, religious, etc.) that could have an impact on the implementation of the project and its success in achieving its outcomes.

Thus in the midst of the continued dominance of quantitative research designs in health behavior and health services research, in recent years there has been a small but growing trend toward greater multi-method and multi-disciplinary approaches to the evaluation of community interventions. Included among the new multi-method approaches are attempts in the "triangulation" or combining of traditional quantitative methods with the more "discovery" oriented qualitative approaches listed above[2].

---

[2] Unfortunately, many times these efforts at true triangulation is compromised as those who strongly adhere to quantitative paradigms continue to believe that qualitative methods are not

Collaborative and Empowerment Evaluation

The growing acceptance of qualitative methods in program evaluation and other social science research coincided with the emergence of a number of theoretical and research orientations in the social sciences and humanities that broke sharply with the more positivist orientations among which quantitative research methods were a primary component. By the 1960s the positivist idea of total investigator control of the research environment so as to be able to get to some objective truth by controlling rival factors to hypothesized relations had dominated the social sciences, including evaluation. However, the rise of theoretical and research orientations such as *cultural theory, critical theory*, and related orientations such as *poststructuralism, postmodernism, feminism, ethnic models, neo-Marxism* and *materialism, constructivism,* and *participatory, action* and *advocacy* research strongly contested the old ideas of *positivism,* that have been outlined by Guba and Lincoln (1998:204) as;

> (1) there is an apprehendable reality that is assumed to exist, driven by immutable natural laws and mechanism;

> (2) that research can in principle, converge on the "true" state of affairs;

> (3) the investigator and the investigated "object" are assumed to be independent entities, and the investigator is capable of studying the object without influencing or being influenced;

> (4) when influenced in either direction (threats to validity) is recognized, or even suspected, various strategies are followed to reduce or eliminate it;

> (5) values and biases are prevented from influencing outcomes, so long as the prescribed procedures are rigorously followed; and

> (6) replicable findings are in fact "true."

For cultural theorists, in social relations, there is no apprehendable reality or truth driven by immutable natural laws or mechanisms, but that truth is a socially or constructed phenomenon, varying based on who is creating it. For critical theorists, cutting across numerous disciplines, efforts should be made to deconstruct those truths created by powerful groups for their own benefit, often times to the detriment of those outside of these groups, but who are also socialized to believe such truths. Similar theoretical orientations are also found in poststructuralism and postmodernism, as well as in neo-Marxist and materialist orientations that focus on the role of

---

science, take too long to do, are too expensive, and if you have to do them, make a quick stab at it, through for example focus groups, so that we can quickly get on with the "real" scientific research. This attitude towards qualitative methods undermines the complementary strength such approaches could bring to our understanding, and the weaknesses of quantitative methods in providing an understanding to context, process, and meaning continue.

history and political economy in the production and reproduction of hegemony. Because there are gender, ethnic, and national attributes to hegemonic structures, feministic and ethnic models call for greater attempts to understand the realities of females, and non-powerful ethnic and national groups. People with these theorietical orientations often view truth or reality as having the following characteristics, again according to Guba and Lincoln (1998:206:207):

1) Realities are apprehendable in the form of multiple intangible mental constructions, socially and experientally based, local and specific in nature (although elements are often shared among many individuals and even across cultures, and dependent for their form and content on the individual persons or groups holding the constructions.

2) Constructions *are not* more or less "true" in any absolute sense, but simply more or less informed and /or sophistications.

3) Constructions are alterable, as are their associated "realities."

4) The investigator and the objective of investigation are assumed to be interactively linked so that the findings are literally created as the investigation proceeds.

5) The variable and personal (intramental) nature of social constructions suggests that individual constructions can be elicited and refined *only* through interaction between and among investigator and respondents.

6) These varying constructions …are compared and contrasted through dialectical interchange.

7) The final aim is to distill a consensus construction that is more informed and sophisticate than any of the predecessor constructions (including…the *etic*[3] construction of the investigator).

Such orientations underlie Buchanan's suggestion that we need to move from a positivist orientation towards social change to a more humanistic theorizing approach. . He points out that while positivism is not only not the only method for gaining valid knowledge, but it is a powerful ideology that thwarts the field's [social change] interests in alleviating suffering and promoting social justice. Buchanan (1998): cites Sullivan (1995:165-166), who writes:

"The dimensions of the challenge [of resolving social problems] become clearer as we ask what stands in the way of an effective response. The answer is disturbing, though manifest. Beside, or rather behind the, the inertia of entrenched interests there is also an ideal and a regime of knowledge which is strongly resistant to the stance of social engagement and moral inquiry demanded of today's challenges. This is the epistemic regime of postivism…The continuing hold of positivism dogma over the thinking and practice of

---

[3] Simply stated, the emic, is an *internal* perspective, from those being studied, about their world, while the etic is the interpretation of that world based on the (social science) methods, knowledge, and skills brought from outside of that world by the researcher.

higher education is a key problem which must be confronted by anyone who concludes that the needs of our time demand a reshaping of professional knowledge."

In his humanistic theorizing approach to planned change, Buchanan (1998:443-447) points out the need for the following:

1) Make assumptions explicit

2) Provide Understanding, rather than explanation

3) Make Sense, which involves the constructivist notions discussed above.

4) Sensitization, wherein theory should heighten people's awareness of a broader range of the particular contingencies bearing on a specific situation.

5) Critique, wherein theory was historically conceived in terms of social criticism in the service of social change.

These anit-positivist theoretical orientations underlie participatory, action, or advocacy research. These research orientations call for a break from the positivist notion of the researcher as the unbiased investigator searching for a truth that should not be contaminated by action on the part of the investigator on behalf of the object of study. In this orientation, not only should the researcher participate in the research with the object of study, but the goal of the research is to assist the object of study in *empowering* themselves to overcome certain health or social problems that they are experiencing. It is within the context of such an intellectual environment that "empowerment evaluation" became part of the evaluation landscape.

Fetterman (2001:3) describes empowerment evaluation as:

" the use of evaluation concepts, techniques, and findings to foster improvement and self-determination…..It is designed to help people help themselves and improve their programs using a form of self-evaluation and reflection."

Fetterman, an anthropologist by training, further describes empowerment evaluation in three simple steps:

1) developing a mission, vision, or unifying purpose;

2) taking stock or determining where the program stands, including strengths and weaknesses; and

3) planning for the future by establishing goals and helping participants determine their own strategies to accomplish program goals and objectives.

In addition, says Fetterman, empowerment evaluators help program staff members and participants *determine the type of evidence* required to document and monitor credibly the progress toward their goals. (Fetterman 2001:23-24). Since its inception, empowerment evaluation has grown in acceptance and popularity, having been "adopted in higher education, government, inner city public education, nonprofit corporations, and foundations throughout the United States and abroad." (Fetterman 2001:1).

While Fetterman states that empowerment evaluation uses qualitative and quantitative methods, qualitative methods predominant, grounded in the constructivist and participatory theoretical models discussed above. It seems a very valuable tool for formative and process evaluation, but there seems to be something missing regarding summative (impact and outcome) evaluation. This could be because while attempting to bring together the *emic* and the *etic*[4] used by anthropologists in their approach to ethnography. While a primary goal in anthropological ethnography is come to an understanding of the world of those with whom they are studying or working *from their perspective*, (the emic) the ethnographer should not lose sight of the perspectives offered by the scientific methods, knowledge and skills that they have gained from their training and professional experiences (the etic). It seems to me that even after one goes through the valuable exercises of empowerment evaluation, one should be careful about leaving it completely to program staff and participants to *determine the type of evidence* required to document and monitor. It is here that some of my colleagues have called me a positivist because I am concerned that an over reliance on the etic side could introduce bias if there is little etic attention given to achieving outcomes and impacts that might be used or generalized by other programs. While I include some of the principles from empowerment evaluation in my own approach to evaluation, which I called *Ethnographic Assessment and Evaluation Systems* (EAES—Whitehead 2002d), in particular in the formative and process components of the EAES, I rely more on approaches like those of Robert Goodman for a more holistic evaluation model that uses qualitative and quantitative research methods. Goodman (1998) suggests five principles that one should follow to address both formative and summative aspects of evaluating community based initiatives:

1) **Evaluation of community programs should include an assessment of program theory**. Here Goodman (1998:40) points out these theories are sometimes referred to as theories of action or *logic models*, or *succinct logical series* of statements link the problems that the program is attempting to address, with *how* it will address them, and what the expected result is. Goodman cites Weiss's (1995) suggestion of program models as a series of micro-steps that contain important assumptions of if/ then statements (If staff are trained well for their respective tasks, then those tasks are more likely to be achieved).

2) **Evaluation instruments that are used to measure community programs must be contoured to each individual community.**

---

[4] Simply stated, the emic, is an *external* perspective, from those being studied, about their world, while the etic is the interpretation of that world based on the (social science) methods, knowledge, and skills brought from outside of that world by the researcher.

3) **Evaluation approaches used should be guided by the questions asked and often require both a quantitative and qualitative orientation**. Here Goodman cites Steckler and colleagues (1992:4).

4) "…health education and health promotion programs are complex phenomena which require the application of multiple methodologies in order to properly understand or evaluate them….Today, the issue no longer is whether to use quantitative or qualitative methods, but rather how they can be combined to produce more effective results"

4) **Evaluation should be informed by social ecology and social system concepts.** Here Goodman discusses theoretical orientations similar to the anthropological concept of holism, and which I discuss in some detail in describing the EAES (Whitehead 2002d)

5) **Community evaluation should involve local stakeholders in meaningful ways.** Here Goodman's point is similar to those found in empowerment evaluation.

References Cited

Buchanan, David R. (1998) "Beyond Positivism: Humanistic Perspectives on Theory and Research in Health Education" Health Education Research. 13(3):439-450

Campbell, D. and J. Stanley (1966). Experimental and Quasi-experimental Designs for Research. Chicago: Rand McNally.;

Cook, T.D. and D.T. Campbell (1983). "The Design and Conduct of Quasi-Experiments and True Experiments in Field Settings. In Handbook of Industial and Organizational Psychology. M.D. Dunnette (ed). New York: Wiley.

Fetterman David M. (2001), Foundations of Empowerment Evaluation, Thousand Oaks, CA: Sage Publications.

Goodman, R.M., "Principles and Tools for Evaluating Community-Based Preventions and Health" Journal of Public Health Management Practice 4(2):37-47

Sullivan, W. (1982). Reconstructing Public Philosophy. Berkeley: University of California Press

Weiss, C.H. (1995)/ "Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families." In New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Context. J.P. Connel et al; (eds). Washington, D.C.: The Aspen Institute,. Pp.65-92.

Whitehead, T.L., (2002d). Ethnographic Assessment & Evaluation Systems, The Cultural Systems Analysis Group, CEHC Working Paper #4, College Park, MD: The University of Maryland

Windsor, R. T. Baronowski, N. Clark, and G. Cutter (1994, 2nd Edition). Evaluation of Health Promotion, Health Education, and Disease Prevention Programs. Mountain View, CA: Mayfield Publishing Co.