Applying a Theory of Change Approach to the Evaluation of Comprehensive Community Initiatives: Progress, Prospects, and Problems

James P. Connell and Anne C. Kubisch

Since its birth in the late 1980s and early 1990s, the field of comprehensive community initiatives (CCIs) has been struggling to find evaluation strategies and methodologies that correspond well to the goals and designs of the initiatives themselves. Up to this point, CCIs have had three general options to follow: (1) retreat to process documentation of the initiatives and greatly reduce expectations about obtaining credible evidence of their impacts; (2) try to "force fit" the initiatives themselves into the procrustean bed of existing and accepted evaluation methods in order to estimate their impacts; and (3) put off evaluating CCIs until the initiatives are more "mature" and "ready" to be evaluated using existing strategies.

The field needs better options than these. Specifically, the field is seeking alternative approaches to evaluating CCIs that will meet both the need to estimate these initiatives' effects on interim and longer-term outcomes and the need for information on how the interventions produce those outcomes. This paper is a progress report on one such approach.

In this paper, we present what we are calling a "theory of change approach" to evaluating CCIs. We describe three stages in carrying out this approach:

- surfacing and articulating a theory of change
- measuring a CCI's activities and intended outcomes
- analyzing and interpreting the results of an evaluation, including their implications for adjusting the initiative's theory of change and its allocation of resources

In many ways, these stages-and the questions they raise-are similar to those of any evaluation process: What is the treatment or intervention? What are its intended and measurable outcomes? And, how are the data to be collected and analyzed such that the causal links between treatments and outcomes are described in the most compelling way? What is different about evaluating CCIs is that the answers to these three questions are often much more elusive.

For example, as we describe the steps that are necessary to surface and articulate a theory of change, the reader will see that, at their most general level, CCI theories are quite similar to many program theories: the initiative plans to do X in order to accomplish Y and Z. But, this similarity-and its implications for evaluating CCIs-vanishes quickly when one realizes that, unlike most programs, CCI theories have multiple strands (economic, political, and social), which operate at many levels (community, institutional, personal network, family, and individual), are co-constructed in a collaborative process by diverse stakeholders, and evolve over the course of the initiative. Each of these complicating factors can plague evaluation of more circumscribed programs, to be sure, but in CCIs these factors are defining traits. They are the rule, not the exception. Moreover, these complicating factors spill over into the other stages of evaluation; in later sections of the paper, we explore their implications for measurement, analysis, and interpretation.

We conclude the paper with reflections on the constraints and promise of the approach, including its capacity to reinforce the basic principles of a CCI and contribute to a knowledge base that can inform future neighborhood-based interventions.

Defining a Theory of Change Approach to Evaluation

What Is a Theory of Change Approach to Evaluation?

Weiss (1995) defines a theory of change quite simply and elegantly as a theory of how and why an initiative works.[1] Building on her work, we have defined a theory of change approach to CCI evaluation as a systematic and cumulative study of the links between activities, outcomes, and contexts of the initiative.

This definition suggests that the first step toward evaluating a CCI is to determine its intended outcomes, the activities it expects to implement to achieve those outcomes, and the contextual factors that may have an effect on implementation of activities and their potential to bring about desired outcomes. For example, the goal of many CCIs is to improve the well-being of children and families in the neighborhood. In this case, one of an initiative's primary activities might be to replace categorical and centralized services with integrated neighborhood-based family resource centers. An important contextual factor might be the policy environment, including the presence or absence of legislation allowing for pooled funding of state resources for innovative community-based initiatives. Another central activity might be to build social networks among families with young children, which in turn could be affected by local contextual factors such as the racial make-up of the neighborhood and its history of intergroup relations.

How do such theories of change assist CCI stakeholders? In 1972, Carol Weiss described the value of having any program evaluation rooted in good theory (1972). In a 1995 paper, Weiss described the potential contribution of this approach to the evaluation of CCIs. We now identify at least three reasons to begin the design and evaluation of a CCI with a good theory of change.

A theory of change approach can sharpen the planning and implementation of an initiative. Used during the design phase, it increases the likelihood that stakeholders will have clearly specified the initiative's intended outcomes, the activities that need to be implemented in order to achieve those outcomes, and the contextual factors that are likely to influence them. These are the building blocks of any good evaluation, but they are especially useful for mid-course feedback to managers and for developing a knowledge base about how and why CCIs work.[2]

With a theory of change in hand, the measurement and data collection elements of the evaluation process will be facilitated. For example, a theory of change asks that participants be as clear as possible about not only the ultimate outcomes and impacts they hope to achieve but also the avenues through which they expect to achieve them (Weiss, 1995). An evaluation based on a theory of change, therefore, identifies what to measure-ultimate and interim outcomes, and the implementation of activities intended to achieve these outcomes-and helps to guide choices about when and how to measure those elements. By providing guidelines for deciding among the various tools in the evaluation toolbox, the approach helps avoid the risk that evaluations will be driven by the tools themselves.

Articulating a theory of change at the outset and gaining agreement on it by all stakeholders reduces, but does not eliminate, problems associated with causal attribution of impact. A theory of change specifies, up front, how activities will lead to interim and longer-term outcomes and identifies the contextual conditions that may affect them. This helps strengthen the scientific case for attributing subsequent change in these outcomes (from initial levels) to the activities included in the initiative. A theory of change approach would seek agreement from all stakeholders that, for example, activities A1, A2, and A3, if properly

implemented (and with the ongoing presence of contextual factors X1, X2, and X3), should lead to outcomes O1, O2 and O3; and, if these activities, contextual supports, and outcomes all occur more or less as expected, the outcomes will be attributable to the intervention. Although this strategy cannot eliminate all alternative explanations for a particular outcome, it aligns the major actors in the initiative with a standard of evidence that will be convincing to them. Clearly, this will not be as powerful as evidence resulting from randomly assigned control and treatment groups, but, as has been noted elsewhere, random assignment of communities is not a feasible avenue of evaluation for CCIs (Hollister and Hill, 1995).

It should be noted at this point that we are advocating a theory of change *approach* to evaluation; it is not an evaluation method that stands on its own. Indeed, the approach relies upon and uses many methodologies that have been developed and refined over the years-quantitative and qualitative, impact and process oriented, traditional and non-traditional, and so on-for information collection, measurement, and analysis. But, if we are right about its promise to generate credible evidence of CCI impact, along with insight into the reasons for that impact, results from evaluations using a theory of change approach should respond to current needs in the field for information about whether CCIs are "working" and generate useful new scientific knowledge to enrich the design of future CCIs.

What Is a Good Theory of Change?

For the approach to achieve its potential, the theory of change guiding the CCI and its evaluation needs to be a good one. We have identified three attributes of a good theory of change that stakeholders should confirm are present before committing to an evaluation and, indeed, should revisit throughout the implementation and evaluation of the initiative:[3]

- **It should be *plausible*.** Do evidence and common sense suggest that the activities, if implemented, will lead to desired outcomes?
- **It should be *doable*.** Will the economic, technical, political, institutional, and human resources be available to carry out the initiative?
- **It should be *testable*.** Is the theory of change specific and complete enough for an evaluator to track its progress in credible and useful ways?

To develop plausible, doable, and testable theories of change, CCIs need to draw upon various sources of information-program experience, scientifically generated knowledge, and community residents' insights, to name some of the most important.[4] This is, in part, because neither social science nor experience-to-date nor participants' insights alone yet offers a complete picture of the processes of change that CCIs are seeking. Social science research and evaluation research, for example, are just beginning to investigate unplanned and planned community change (Sullivan, 1996) and how community change impinges on the lives of residents (Connell, Aber, and Walker, 1995; Aber, Berlin, Brooks-Gunn, and Love, 1997). This emerging research base must be combined with the insights and experience of past CCIs (for example, O'Connor, 1995; Halpern, 1995) and the new insights and experience of contemporary CCIs (for example, Kubisch et al., 1997; Stone, 1996) in the development of theories of change.

For example, suppose a CCI sets crime reduction in the target neighborhood as one of its longer-term desired outcomes. If the information guiding the selection of interim and early outcomes and initial activities for the CCI were drawn primarily from the field of law enforcement, the CCI designers might develop a theory of change that highlights increased police presence in the neighborhood. If the information were drawn from urban planning, the theory might focus on

improving the condition of open spaces in the neighborhood or tearing down abandoned housing. If the information were drawn from the field of family therapy, the theory might include activities designed to provide developmental supports for adolescents in households with track records of child or spousal abuse. Or, if the information were drawn from community residents' knowledge, well-known drug dealers might be identified for sting operations. By enriching the construction with information from all these sources, the planning process should yield a more plausible theory of change with respect to the longer-term outcome of crime reduction.

O'Connor (1995) has emphasized that one element that must also be incorporated into the theory of change is the external environment. Careful consideration of context helps the designer and evaluator gain clarity about factors that may have a significant bearing on a CCI's chances for achieving its intended outcomes but that the initiative itself is not initially able to influence. This should help ensure that activities are strategically implemented and that the evaluation yields sharp and compelling tests of its hypotheses. In the hypothetical theory of change for education reform presented later in this paper (grid 2 *325k*), for example, the local school task forces need to be aware of upcoming board elections at the local level, regulations for allocation of state education funds, and federal timelines for phasing out court-ordered bussing.

Carrying Out a Theory of Change Evaluation

Having attempted to describe a theory of change, we now turn to the task of describing how an evaluation based on a theory of change might proceed. On this front, we have made some progress and met some difficult challenges, and, in the next three sections, we share both. The first step is to map out a process that should produce a plausible, doable, and testable theory of change. This is where we have the most experience, and we discuss how to go about getting the theories surfaced, articulated, and aligned.

We then turn to the questions of how and when to measure activities and outcomes included in the theories. And finally, we examine whether the information being generated is credible enough to make judgments about how well the initiative is working.

Surfacing and Articulating a Theory of Change

In the introduction to New Approaches to Evaluating Community Initiatives (1995), Kubisch, Weiss, Schorr, and Connell conclude that CCIs are difficult to evaluate in part because their designs are underspecified at the outset of an initiative. As a consequence, one of the first things evaluators are commonly asked to do is help specify the theory underlying the intervention and thereby "unpack" the intervention itself.

Chen (1990) and Patton (1986) describe a process in which stakeholders and evaluators "co-construct" the initiative's theory so as to maximize its utility for all, as a planning and management tool, as a vehicle for participant empowerment, as a guide to resource allocation, and as a way of communicating with the field as a whole. This approach resonates with the CCI field's limited experience in this area (see, for example, the practitioners' reflections in this volume).

In our experience, surfacing and articulating a theory of change through a collective and collaborative process is as fraught with difficulties as it is full of promise. The case study by Milligan, Coulton, York, and Register in this volume confirms this and provides more detail about the nature of the real-life discussions that occur on the ground in order to produce such

theories of change. Although we are still in the early stages of learning about how to carry out a theory of change evaluation, some lessons and challenges that appear to be particular to the theory articulation process are beginning to emerge. They fall into two broad categories: generating a theory of change and reconciling multiple theories of change.

Generating a Theory of Change

From our perspective, the goal of the participatory planning process is to generate a theory of change that is viewed by its stakeholders as plausible, doable, and testable. With these three criteria in mind, we recommend that the following questions be considered as part of the planning process:

- What longer-term outcomes does the CCI seek to accomplish?
- What interim outcomes and contextual conditions are necessary and sufficient to produce those longer-term outcomes, beginning with penultimate outcomes and moving through intermediate to early outcomes?
- What activities should be initiated and what contextual supports are necessary to achieve the early and intermediate outcomes?
- What resources are required to implement the activities and maintain the contextual supports necessary for the activities to be effective, and how does the initiative gain the commitment of those resources?

If there is a gap between existing or projected resources available to the initiative and those deemed necessary to implement activities that will produce outcomes, the initiative will have to raise the resources necessary to close the gap. If the gap cannot be closed, first outcomes and then activities will have to be adjusted in order for the theory of change to remain "doable."

We believe that the sequence of steps shown in the diagram below is important to maintain: start with long-term outcomes, work backward toward initial activities, and then map required resources against existing resources. Beyond the general frame, however, much has to be determined locally. Who participates in the conversations? How, when, and where should these conversations take place? These questions must be resolved within the community setting.

Experience from a wide range of programs and CCIs shows that identifying and agreeing upon long-term outcomes is relatively easy, in part because long-term outcomes are generally so broad as to be uncontroversial: for example, improved high school graduation rates, greater "sense of community," or increased income levels. Likewise, identifying early activities is relatively straightforward. Intermediate and early outcomes are more difficult to specify because scientific and experiential knowledge about links between early, interim, and long-term outcomes is not well developed in many of the key areas in which CCIs operate. Defining interim activities and interim outcomes, and then linking those to longer-term outcomes, appears to be the hardest part of the theory articulation process.

For the purposes of this paper, we have generated two theories of change that could have resulted from this process. Each illustrates, at an admittedly general level, what an early version of a theory of change might look like for a CCI that is focusing on community building (grid 1 *227k*) or education reform (grid 2 *325k*).[5] Each grid includes longer-term outcomes, a set of interim outcomes that should lead to these longer-term outcomes, some early outcomes that logically precede the intermediate ones, and the initial activities that are meant to lead to the early outcomes. Each grid is also divided into four levels of outcomes that appear to be most relevant

for current CCIs: community, organizational/institutional, family/personal network, and individual.

The process of constructing these theories of change highlights the tenuous nature of the causal linkages guiding many CCIs. In the education reform theory of change ([grid 2](#) *325k*), evidence from school-reform research tells us that the long-term, individual outcome of improved student performance will probably occur when the specified changes at the school, or institutional, level are put in place. There is much weaker support, however, for the part of the theory that identifies the formation and activation of an education task force within the CCI as a critical step toward producing institutional change in the schools. In the case of the community building theory of change ([grid 1](#) *227k*), we have even weaker scientific and practical knowledge about how to produce some of the longer-term outcomes specified in the theory. Consider, for example, the outcome of residents feeling comfortable about taking action when a neighbor "does something wrong," such as using drugs publicly. Given that we have little evidence about how to produce that outcome, how can we have confidence that its hypothesized precursors-say, creating new block-level civic and social activities so that more residents know their neighbors-even if implemented, will take us down the right path?

In addition, specifying intermediate outcomes and how they may lead to long-term change can be a politically charged process, especially if those outcomes might imply major resource reallocation or power shifts. For example, in the educational reform theory of change ([grid 2](#) *325k*), we hypothesize that the strategic plan of the CCI education task force has the goal of making significant changes in school operations. These changes may require different uses of funds, new job descriptions, and perhaps even laying people off. Gaining consensus among all stakeholders, especially the education professionals, on those changes and how they will be made is more difficult than simply pronouncing that all children will show meaningful improvement in their educational performance.

With regard to a CCI's activities, early activities proposed by the CCI are often fairly well specified, while later activities and their links to later outcomes tend to remain underspecified. This lack of detail reflects the fact that few CCIs have gotten much beyond early outcomes in any of their operational areas. Moreover, CCIs are meant to be dynamic enterprises, and the initial theory of change serves as a map of mostly uncharted territory, a map that the CCI itself will have to revise as it makes its voyage. Thus, our experience suggests that a common answer to the question, "What do you expect to be doing in the fifth year of the initiative?" is "Ask us in the fourth year, and we'll tell you."

The inability of many stakeholders to make linkages between early activities and longer-term outcomes raises significant problems for evaluation design. Perhaps the greatest factor in determining the feasibility of the theory of change approach is the capacity of a CCI's stakeholders and evaluators to identify, prioritize, and then measure the key activities and contextual factors, not in retrospect but in advance. Evaluators and stakeholders alike are quite good at looking back on interventions and constructing compelling tales of why a particular result did or did not occur. The challenge posed by the theory of change approach is to theorize prospectively about these issues. This requires balancing the need for the theory of change to remain responsive to emerging opportunities and challenges with the need for investors in CCIs (including funders, implementers, and participants) to have some basis upon which to judge the likelihood of reaching the intended long-term outcomes.

Reconciling Multiple Theories of Change

Once the theory specification process begins, it quickly becomes apparent that various stakeholders in the initiative can, and often do, hold different views about what it will take to produce the long-term outcomes of the initiative. Indeed, one of the important contributions of this approach is that it points out that multiple theories of change may be operating simultaneously within a single CCI and that various CCI stakeholders may be working under different, and possibly even competing, theories of change.

For example, Weiss (1995) pointed out several assumptions CCI architects and funders are including in their theories of change: that $250,000-$500,000 per year is a significant enough amount of money to cause change at the neighborhood level; that a neighborhood is an appropriate unit on which to target an initiative's efforts; and that agency collaboration is required at the neighborhood level. One can imagine the director of a lead agency chosen to implement a new foundation-funded CCI operating under quite different assumptions. She might believe that $250,000-$500,000 is not nearly enough to implement the activities required to achieve the CCI's stated outcomes. She might think, therefore, that the initial grant should be seen as a vehicle to leverage additional funds. She might also believe that the neighborhood is not capable of achieving the change on its own and that significant efforts must be made to involve the neighborhood in citywide efforts. Finally, she might think that agency collaboration is nice but takes a lot of time and preempts the agencies' time to deliver specialized services that are badly needed in the neighborhood. In what ways are these different hypotheses in harmony, and possibly reinforcing, and in what ways do they imply different activities, timelines, and even outcomes? These views should be uncovered as stakeholders move through the steps presented earlier of surfacing and articulating the theories of change guiding the CCI.

It is not uncommon for CCIs to be launched without the various theories of change being articulated, much less reconciled. Some experts have noted that one of the great strengths of the CCI phenomenon might well be that it can accommodate multiple theories of change and move forward without their reconciliation. But CCIs that remain inclusive enough to accommodate these multiple theories cannot avoid integrating the theories at two points: the allocation of resources and the evaluation. It is, after all, in the decisions about which activities to invest in that priorities must be developed about which of the various hypotheses that link activities to outcomes are most promising. And it is in the process of designing an evaluation that specific decisions must be taken regarding what is meant by key terms (such as "collaboration"), the type and degree of change being sought, and the measures that would indicate whether change is occurring.[6]

In the example above, the difference between the funder's and director's views of the importance of agency collaboration would emerge as they decide upon early and interim outcomes that each believes will lead to long-term change. In the funder's eyes, credible evidence of collaboration might involve building structural institutional links, such as joint staffing of a new family service program and pooled resources. The lead agency director, on the other hand, might view collaboration so differently that her marker of progress might simply be increased evidence of referrals among agencies.

Resolving the challenges that these multiple theories pose is a political as well as scientific process. Patton (1996) and Usher (1996) warn that imposing strict standards of theory articulation too early in the process can undermine participation and stifle the dynamic nature of the CCI enterprise. At the same time, as suggested above, leaving the CCI's theory of change

ambiguous permits, and indeed encourages, various stakeholders to project their own preferences about activities and outcomes onto the initiative. This Rorschach-test model for CCI theories of change can set up false and unrealizable expectations among stakeholders that could become problematic during resource allocation and evaluation planning. The political question that confronts the CCI manager becomes: When is the CCI robust enough to have the true diversity of its stakeholders' theories surfaced and integrated into its overall theory? Or, when can it no longer operate without doing so? Will these different theories of change be included as parallel, integrated, or competing strands in the overall theory, or will some be selected for inclusion in the implementation and evaluation of the initiative and others not? The task of addressing these issues should not fall solely, or even primarily, to the evaluator, but the evaluation discussion may serve as the context within which they are played out.

The requirements that theories be articulated and that they be specific enough for stakeholders to make judgments about whether or not they are plausible, doable, and testable do not preclude those theories from incorporating multiple perspectives on what long-term outcomes are important, what the interim steps are to getting to those long-term outcomes, and what activities should be implemented. Our two hypothetical theories of change are complex and include multiple strategies for achieving long-term outcomes.[7] Plausible theories of change will no doubt be complex and pluralistic, but if they are to be implemented (doable) they cannot be contradictory and if they are to be evaluated (testable) they cannot be unarticulated.

Measuring Activities and Outcomes

In any evaluation, outcomes and activities must be translated into observable measures: How do we know that the treatment or program occurred, and how do we measure its results? That measurement process in CCIs is likely to be more complex and demanding than in typical program evaluations.

Even in evaluations of multifaceted programs-having different combinations of elements, multiple mediators and moderators of treatment effects, and multiple desired outcomes-the measurement process tends to reside at one unit of observation, typically the individual, and to occur in a fixed order, with treatment assessed first, mediators next, and outcomes last. Moreover, measurement in more circumscribed program evaluations tends to draw primarily on quantitative techniques, such as enumerating participation through administrative records, surveying clients to gauge exposure and experience, and actuarial activities to measure discrete client outcomes.

The measurement demands placed on evaluators of CCIs by a theory of change approach are quite different. Measures of outcomes and activities must be developed at multiple levels. In the case of outcomes, as in any other evaluation, some are more difficult to measure than others. CCI outcomes such as improvements in infant mortality, high school graduation, and employment rates are relatively easy to measure.[8] Measures of community attributes such as social capital, shared values, and strong networks are more elusive, as are institutional change indicators of service integration, responsiveness to community needs, and systems reform. CCI "treatments" are sets of activities that occur over time in inter-related clusters at different levels of observation, with later activities being shaped by the outcomes of earlier activities. Therefore, traditional single-point or fixed-interval longitudinal assessments of outcomes will not effectively capture the change process in CCIs, and staggered baselines may even be necessary. Clearly, the measurement burden in a theory of change evaluation is likely to be heavy.[9]

Although most CCIs are still in the early stages of measuring activities and outcomes, three measurement issues are emerging that are specific to a theory of change evaluation.[10]

Measurement of a CCI's activities is as important as measurement of its outcomes. To make a case for impact, the theory of change approach seeks to accumulate rigorous tests of links between an initiative's activities and their expected outcomes. Therefore, it must have compelling measures of both activities and outcomes and then link, through causal inference, change in one to change in the other, repeatedly and cumulatively over the early, intermediate, and later stages of the initiative. Thus, in this approach, process is recast into activities, and outcomes are expanded to include results occurring over the entire course of the initiative, not just the long term, and at multiple levels, not just the individual level. Some of those activities might relate more to the creation of the conditions or capacities for achieving outcomes, and their outcomes might be indicators of readiness to continue with the next set of activities.

The measurement discussion must resolve the issue, "How good is good enough?" For a theory of change to guide an initiative and its evaluation, performance standards must be set for the outcomes included in the theory as well as for the implementation of activities. In our education reform example, many of the critical features of school site reform included in the theory of change have specific targets in the intermediate and long terms: for example, reducing student/adult ratios to less than 15/1 during core instructional periods. For the evaluation to be useful during early stages, stakeholders will have to be precise about what the initial activities should look like and what early outcomes will be required to produce desired thresholds on intermediate outcomes. For example, what form of commitment from school officials to the reform plan will be sufficient to proceed to the first implementation step? Or, in our community building example, how will we know whether the CCI governing entity has good representation, good leadership, an adequate management structure, and appropriate staffing?

Certainly any evaluation would be well served by establishing expected threshold levels of change, but this issue is crucial in a CCI theory of change evaluation. Stakeholders must have a clear idea of how much change is "good enough" for them because there may not be an ex post analysis of whether CCI-induced change, particularly on early and intermediate outcomes, was "statistically" significant.

The process of surfacing and articulating the theory of change will provide important information about measuring activities and outcomes. For example, a good theory of change provides some rational structure for determining measurement points. Instead of setting arbitrary and perhaps inappropriate data-collection points (or expecting "constant" data gathering from the evaluator), the approach suggests that measurement points be based on when activities specified in the theory of change should occur and when their intended outcomes should occur. If a strong collaborative process is established in the early stages of a theory of change evaluation, it can help assure that stakeholders will work with the evaluator to explore creative measurement strategies.

Making the Case for Impact

The question of how to make a convincing case that a CCI has or has not worked remains problematic. The paper by Robert Granger in this volume takes up this topic in much greater detail, so we will raise only a few points at this time.

At the most general level, the theory of change approach contends that the more the events predicted by theory actually occur over the course of the CCI, the more confidence evaluators and others should have that the initiative's theory is right. We suggest, then, that the major audiences for an evaluation of a CCI-including community residents, initiative managers and funders, and policy makers-should be convinced that the initiative "worked" if four points can be demonstrated:

- up front and along the way, a well-specified and plausible theory of change described steps toward an anticipated change (from historical baselines) in important outcomes for the community, its institutions, and its residents
- the activities of the CCI that were part of these steps were implemented at expected thresholds
- the magnitude of changes in the early, intermediate, and long-term outcomes that followed these activities met predicted thresholds
- no obvious and pervasive contextual shift occurred that could otherwise account for all these predicted sequences of activities and outcomes

Would these criteria be sufficient for the most skeptical researchers to agree that the CCI worked, that it had impact? Probably not. But ruling out all alternative explanations through randomized experimental methods is not feasible as the *primary* inferential tool for attributing impact to a CCI. The theory of change approach for establishing impact draws on tried and true scientific traditions of testing hypotheses about cause and effect relationships, including methods used in physical, biological, and other social sciences.

What sets the social policy evaluation field apart from most of the rest of the scientific world has been its conclusion that all forms of hypothesis-testing are inadequate relative to experimental approaches that include random assignment of units of observation (individuals, institutions, communities) to treatment and control conditions (Hollister and Hill, 1995). The argument made by those who seek to retain this single standard of evidence for impact is that it is the *only* way to establish a convincing counterfactual. But, part of the reason for the dominance of the experimentalists is that the magnitude of the change expected to occur as a function of many social interventions has been very small. The smaller the change expected, the more solid the counterfactual must be in order to attribute cause. However, in most current CCIs, the magnitude of change expected in the long-term outcomes is not small. They seek neighborhood transformation and meaningful improvements in individual and family well-being. Whether or not stakeholders agree that these expectations are realistic-that is, whether the theory of change is plausible and doable-should be determined before they are tested in an evaluation. The more significant the change that occurs, we and others (such as Gueron, 1996) would argue, the less the need for airtight counterfactuals to attribute impact to CCIs.

Theory of Change Evaluation Reports

The theory of change approach recasts traditional distinctions made in the program evaluation literature. For example, *a systematic and cumulative study of links between activities and outcomes* replaces process documentation, implementation, and outcome studies. This recasting has implications for what the products of such an evaluation might look like.

Most evaluations of complex initiatives focus their measurement activities on long-term outcomes. As a secondary component, "process" or "documentation" studies collect descriptive

data (usually qualitative) on planning and implementation activities. Implementation studies discuss whether or not and how completely the program or initiative was implemented. These studies are then written up separately from the outcomes studies or loosely coupled in a summary report. The message delivered by this form of reporting is clear: what is really important is whether the long-term outcomes changed; how the change occurred and why the intervention did or didn't work are secondary.

Alternatively, and more frequently in evaluation reports on CCIs, process documentation and implementation studies are all that is presented because long-term outcomes are not expected to have shown any change until years after the initial activities are implemented and because there is no strong, a priori theory of change linking early activities to early outcomes. Process documentation has not been unwelcome, since it often provides both funders and directors of initiatives with formative feedback on operational issues useful for mid-course corrections. Even so, those who have been involved with CCIs for some time are well aware that it is important for evaluators to avoid becoming so enmeshed in information about process that they lose sight of the importance of whether and how activities are leading to desired outcomes. Assessing quality of implementation is a necessary but not a sufficient condition for good evaluation, even in the short term. What early outcomes are these activities supposed to affect? How are these outcomes being measured? How are the links between initial activities and early outcomes established? How is the movement toward intermediate and longer-term outcomes being tracked?

Thus, the theory of change approach replaces the two principal, and often independent, types of evaluation reports with one that explicitly and deliberately covers both activities (formerly process and implementation) and outcomes (formerly long-term outcomes only). As in most approaches to evaluation, the very first reports are likely to provide baseline readings on desired early, intermediate, and long-term effects. But soon thereafter, we should expect evaluation reports to focus on links between initial activities and early outcomes. We believe that this form of reporting helps the theory of change approach supply all stakeholders with timely, useful, and rigorous information about the progress of their initiative and can provide early guidance if the theory of change needs to be revised.

Strengths and Challenges of the Theory of Change Approach

The major strength of the theory of change approach is its inherent common sense. Its major *competitive* advantage is the inability of other currently available approaches to do the job.

We have described the potential benefits of the theory of change approach from the initial planning of a CCI, through the measurement of its outcomes and activities, to the analysis and interpretation of the data. It should generate useful learning over the life span of the initiative and could spawn cross-initiative learning as well. But perhaps its most powerful contribution to the evaluation endeavor is its emphasis on understanding not only *whether* activities produce effects but *how* and *why*, throughout the course of the initiative. This is useful at the level of the CCI itself, where feedback from the evaluation can help stakeholders make informed decisions about whether to stay the course or modify goals and strategies. But the audience for information about the "how" and "why" is also much wider. It includes other program directors, designers, or funders who need guidance as they develop new efforts or attempt to learn from successful initiatives. They need specific information about actual mechanisms that are related to good outcomes, and an evaluation guided by a theory of change can provide that level of detail (Weiss, 1997). And, given that CCIs are explicitly designed as experiments in how to transform distressed communities, a theory of change evaluation contributes to the "social learning"

objectives of evaluations and gives them a national and perhaps even international audience in policy and research. For example, the approach exposes gaps in existing knowledge and helps lay out a research agenda in many fields of inquiry.

The theory of change planning and evaluation process can also reinforce the broader goals of the initiative. In particular, the process can be a powerful tool for promoting collaboration and engagement at the community level focused on products and outcomes. For example, building capacity for ongoing problem solving at the community level is a goal of virtually all CCIs, and, because a theory of change evaluation explicitly values community knowledge, vehicles for resident involvement are easily built into its structure. Brown (1995) suggests that an evaluation, if so designed, can contribute to that goal by ensuring that data about the community are collected, analyzed, and then fed back in a way that can educate and mobilize residents to participate in an ongoing community planning process. Also, although all participants may agree that evaluation is important, it can be difficult to get agency staff or residents to invest the requisite time and energy in the evaluation enterprise, which often seems secondary to activities that lead more directly to improved outcomes. We suggest that a theory of change approach to evaluation, when developed with or shared with staff and residents, can help to overcome that reluctance. It helps to make explicit the capacity-building agenda of the initiative and the role evaluation can play in that agenda.

At the same time, creating new partnerships, fostering collaboration, and developing community capacity can lead to a diffusion of responsibility that is detrimental to the impact and sustainability of an initiative. Again, a theory of change is helpful, in this case to clarify accountability pathways in the initiative. The more explicit the theory of change, the more explicit and consensually validated the accountability structure can be.

Thus, the approach breaks down the line between formative and summative evaluation. It diminishes the perceived trade-off between rigor and programmatic utility of information being collected. It aspires to both simultaneously, while adding the goals of social learning and capacity building. The approach, as we see it, recognizes that there is neither the need nor the time to collect separate kinds or different qualities of information in order to serve all of these important goals.

Before closing, however, we feel obligated to be clear about what we foresee as some of the burdens of a theory of change approach to evaluation. In order for it to be most useful, all stakeholders will need to invest time and political capital in developing plausible, doable, and testable theories of change. This is true for all participants but especially for evaluators, who will need to broaden their view of their role to include eliciting complex theories of change and translating them into evaluation designs; using or developing multiple, and often sophisticated, measures of activities and outcomes; ascertaining linkages between activities and outcomes on a continuous basis; detecting differences between espoused theories and program implementation; and supporting strategic efforts to convince skeptics that evaluation results are compelling.

Of course, many of these investments would be required for any good evaluation. What is not yet known is whether a theory of change approach suggests still greater investments on these fronts or whether a well-articulated theory of change will create offsetting efficiencies as well. In all likelihood, the answer will vary greatly by initiative, by evaluator, and by the availability of funding.

Finally, it is important to reiterate that the theory of change approach to CCI evaluation is *only an approach*. It provides a framework for embarking on the evaluation of a complex initiative that promises to be useful to the evaluator and other stakeholders. It helps to identify what should be evaluated and clarifies the research questions. But it will employ familiar methodological tools to measure outcomes and activities and to strengthen the credibility of its conclusions. It does not solve all the evaluator's challenges in working with these complex initiatives, but it might make the job more tractable, useful, and rewarding.

The next and most important question is how this approach holds up in the crucible of practice. Fortunately, versions of the approach are already being tested in a number of initiatives, some of which are described elsewhere in this volume. Some insights on the utility of the approach should be forthcoming from these initiatives and others in the near future. The true test of the approach will come as stakeholders in these and other CCIs take stock and decide whether the evaluation results are clear, compelling, and useful for their purposes.

Notes

1. Carol Weiss introduced the notion of theory-guided program evaluation in 1972 (Weiss, 1972). Huey-tsyh Chen has been working on theory-driven evaluation approaches since 1981, particularly as applied to program evaluation, and many others have contributed to its elaboration and evolution along the way. See, for example, Bickman, 1987; Chen, 1990; Chen and Rossi, 1992; Fetterman et al., 1996; Gaventa, 1995; Patton, 1986; Yin, 1989.
2. See Connell (1996) for a fuller elaboration of these arguments.
3. Recent field-based work by Connell suggests that a fourth criterion be added: that outcomes included in the theory of change be meaningful to all stakeholders. One can imagine a theory of change that meets the first three criteria but is focused on goals and outcomes that are considered unimportant or even trivial by one or more stakeholder groups.
4. Existing literature on program evaluation suggests that there are different types of theories and that, for both design and evaluation purposes, there are tradeoffs among

them. For example, Chen (1990) makes the distinction between normative (experience based) and causative (scientifically based) program theories, and Rossi (1996) argues that scientifically based theories are more useful for program evaluation. Our view is that CCIs are such new and complex phenomena that we must mine multiple sources of information in order to try to develop the strongest possible theories of change.

5. See Connell (1996) for a fuller examination of an education reform theory of change.
6. If, on the other hand, the initiative and the evaluation are extremely well funded, multiple theories of change might be implemented and tested.
7. We have encountered other cases where multiple strategies are being pursued in the service of a common goal: for example, building developmental supports for youth by establishing service-based referrals to church youth groups, community service programs for youth, and support networks for their grandparents. In other cases, multiple goals are served by a single set of activities: for example, putting resources into an existing "community kitchen" to improve physical health by providing nutritious meals, to offer jobs to otherwise unemployed residents, and to build "individual responsibility" by insisting that those who receive free meals contribute time to the kitchen or other neighborhood service activity.
8. The "face validity" of any measure could be questioned by some stakeholders in CCIs. For example, change in the proportion of students completing high school is relatively easy to measure, but some stakeholders may challenge this as an indicator of "progress" if they perceive standards being lowered to increase proportions of students achieving this outcome.
9. These challenges of measurement and others-when and how often to collect measures of these activities and outcomes, or who or what should be measured to capture neighborhood, institutional, family and individual change-are discussed in more detail in the chapter by Michelle Gambone in this volume. The Roundtable's steering committee on evaluation is also developing an annotated catalogue of measures that will give potential users important information about available instruments and strategies.
10. These issues are addressed in greater depth by Michelle Gambone, in this volume.

References

Aber, J. Lawrence, Lisa Berlin, Jeanne Brooks-Gunn, and John Love. 1997. "Comprehensive Community Initiatives: Their Role in Promoting the Development of Young Children and Their Families." Draft report. Washington, DC: Aspen Institute.

Argyis, Chris, and Donald Schon. 1992. *Theory in Practice: Increasing Professional Effectiveness*. San Francisco: Jossey-Bass.

Brown, Prudence. 1995. "The Role of the Evaluator in Comprehensive Community Initiatives." In *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*, ed. James Connell et al. Washington, DC: Aspen Institute.

Chen, Huey-tsyh. 1990. *Theory Driven Evaluations*. Thousand Oaks, CA: Sage Publications.

Chen, Huey-tsyh, and Peter H. Rossi., eds. 1992. *Using Theory to Improve Program and Policy Evaluations*. New York: Greenwood Press.

Connell, James P. 1996. *First Things First: A Framework for Successful School-site Reform*. White paper prepared for the E. M. Kauffman Foundation. Kansas City: E. M. Kauffman Foundation.

Connell, James P., Anne C. Kubisch, Lisbeth B. Schorr, and Carol H. Weiss, eds. 1995. *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*. Washington, DC: Aspen Institute.

Connell, James P., J. Lawrence Aber, and Gary Walker. 1995. "How Do Urban Communities Affect Youth? Using Social Science Research to Inform the Design and Evaluation of Comprehensive Community Initiatives." In *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*, ed. James Connell et al. Washington, DC: Aspen Institute.

Fetterman, David M., Shakeh J. Kaftarian, and Abraham Wandersman, eds. 1996. *Empowerment Evaluation: Knowledge and Tools for Self-Assessment and Accountability*. Thousand Oaks, CA: Sage Publications.

Gaventa, John, Janice Morrissey, and Wanda Edwards. 1995. *The Evaluation and Learning Initiative of the National Empowerment Zone and Enterprise Community Program: Review and Recommendations for Phase II Support. Volumes I and II*. Knoxville: University of Tennessee.

Granger, Robert. 1995. Personal communication.

Gueron, Judith. 1996. Personal communication.

Halpern, Robert. 1995. *Rebuilding the Inner City: A History of Neighborhood Initiatives to Address Poverty in the United States*. New York: Columbia University Press.

Hirota, Janice. 1995. *Interim Report: Fieldwork and a Theory-based Approach to Evaluation. Draft report to the evaluation steering committee of the Roundtable on Comprehensive Community Initiatives for Children and Families*. Washington, DC:Aspen Institute.

Hollister, Robinson G., and Jennifer Hill. 1995. "Problems in the Evaluation of Community-Wide Initiatives." In *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*, ed. James Connell et al. Washington, DC: Aspen Institute.

Kubisch, Anne C., Carol H. Weiss, Lisbeth B. Schorr, and James P. Connell. 1995. "Introduction." In *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*, ed. James Connell et al. Washington, DC: Aspen Institute.

Kubisch, Anne C., Prudence Brown, Robert Chaskin, Janice Hirota, Mark Joseph, Harold Richman, and Michelle Roberts. 1997. *Voices from the Field: Learning from Comprehensive Community Initiatives*. Washington, DC: Aspen Institute.

Milligan, Sharon, Claudia Coulton, Peter York, and Ronald Register. 1996. *Implementing a Theories of Change Evaluation in the Cleveland Community Building Initiative*. Cleveland: Center on Urban Poverty and Social Change, Case Western Reserve University.

Mueller, Elizabeth, Xavier de Souza Briggs, and Mercer Sullivan. 1996. *From Neighborhood to Community: Evidence on the Social Effects of Community Development*. New York: Community Development Research Center, New School for Social Research.

O'Connor, Alice. 1995. "Evaluating Comprehensive Community Initiatives: A View from History." In *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*, ed. James Connell et al. Washington, DC: Aspen Institute.

Patton, Michael. 1986. *Utilization-Focused Evaluation*. Newbury Park, CA: Sage Publications.

Patton, Michael. 1996. Personal communication.

Rossi, Peter. Forthcoming. "Evaluating Community Development Programs: Problems and Prospects." In *Urban Problems and Community Development*, ed. Ron Ferguson and William Dickens. Washington, DC: Brookings Institute.

Stone, Rebecca, ed. 1996. *Core Issues in Comprehensive Community-Building Initiatives*. Chicago: Chapin Hall Center for Children.

Sullivan, Mercer. 1996. "Where Is the Community in Comprehensive Community Initiatives? Lessons from Community Studies for Design and Evaluation." Draft Report. Washington, DC: Aspen Institute.

Trochim, William M. K., and Judith Cook. 1992. "Pattern Matching in Theory-Driven Evaluation: A Field Example from Psychiatric Rehabilitation." In *Using Theory to Improve Program and Policy Evaluations*, ed. Huey-tsyh Chen and Peter H. Rossi. New York: Greenwood Press.

Usher, C. Lynn. 1996. Personal communication.

Weiss, Carol Hirschon. 1997. "How Can Theory-Based Evaluation Make Greater Headway?" *Evaluation Review* 21, No. 4.

Weiss, Carol Hirschon. 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 Contexts*, ed. James Connell et al. Washington, DC: Aspen Institute.

Weiss, Carol Hirschon. 1972. *Evaluation Research: Methods of Assessing Program Effectiveness*. Englewood Cliffs, NJ: Prentice-Hall.

Yin, Robert K. 1992. "The Role of Theory in Doing Case Study Research and Evaluations." In *Using Theory to Improve Program and Policy Evaluations*, ed. Huey-tsyh Chen and Peter H. Rossi. New York: Greenwood Press