The Development and Use of a Theory of Change to Align Programs and Evaluation in a Complex, National Initiative

Written by

Wendy M. DuBow and Elizabeth Litzler

ABSTRACT

In a complex, multisite initiative, a well-developed theory of change can help an organization's leadership make strategic decisions about program elements, organizational direction and priorities, as well as provide structure and accountability for evaluation. This article describes the multi-step process used by the evaluation team of a large nonprofit to iteratively develop a theory of change for a complex national initiative, including challenges encountered during the creation process and the initial implementation phase. We describe the benefits derived from the development process and the uses of the theory of change by the program staff and by the evaluation team. Although the theory of change highlighted weaknesses in the evaluation, it also provided us with an opportunity to improve the evaluation. Recommendations are included for those evaluators seeking to develop a theory of change for their own organizations.

INTRODUCTION

Developing a logic model, figuring out measurable objectives and outcomes of interest, then collecting

and analyzing data to assess a program's effectiveness is arguably the essence of our profession. These

steps in themselves can be complex and require tremendous thought and careful navigation with

program stakeholders. But many of us are faced with evaluation projects that are even more

complicated than a single program evaluation. Some projects challenge us to take a bird's eye view of

our object of evaluation, to step back and try to understand how many moving parts fit together. The

authors of this paper are in just such a situation, and as one of our techniques for addressing this

complexity, we developed and used an organizational theory of change. We are the senior members of

the internal and external evaluation team who are charged with developing, conducting, and reporting

on a complex, highly distributed organization that is aiming at social change on a national scale, the

National Center for Women & Information Technology (NCWIT). DuBow is NCWIT's internal director of

evaluation and Litzler is the director of NCWIT's external evaluation firm. In this paper, we will share the

specific steps we used to work with the program staff and other relevant stakeholders to develop a

theory of change , and how we were able to achieve buy in and, ultimately, utilization. The paper will

also share how the theory of change informed not only the organization's direction but also the external

evaluation plan, ultimately strengthening both through heightened focus.

WHY USE A THEORY OF CHANGE?

A common starting point for many evaluators is to construct a logic model based on the program to be

evaluated. This is in fact chapter 1 of the *Handbook of Practical Program Evaluation*. Often this is done

in conjunction with program staff to ensure buy-in and accuracy (Fawcett et al., 2003). However, faced

with a complex, social change initiative with multiple programs and efforts, a logic model can constrain

rather than clarify the initiative (Breuer et al., 2016). As others have noted, it is not just logic models

that don't fit complex social change initiatives, but many other traditional evaluation designs also don't apply (Weitzman, Silver, & Dillman, 2002).

Initiatives with ambitious goals, such as social or cultural change, are inherently complicated and can, therefore, be quite difficult to evaluate (Weitzman et al., 2002). They often include many different types of programs, occurring on multiple fronts at once, sometimes each with their own program evaluations. As the literature has noted, developing a theory of change in these situations brings many advantages both to program staff and to the evaluation (Rogers, 2007; Weiss, 1997), including program planning, program development, and clarifying assumptions about and reasons for activities (Weiss, 1997). Two decades ago, Kubisch, Fulbright-Anderson, and Connell (1998) identified some of the specific difficulties inherent in the evaluation of complex social initiatives and proposed a theory of change as a method to enhance the quality and accountability of these initiatives. From these early foundations, program theory evaluation, also called theory-based evaluation, evolved in an attempt to address some of the issues inherent in complex initiatives (Rogers, Petrosino, Huebner, & Hacsi, 2000; Christie & Azzam, 2005).

Carol Weiss, an early proponent, believes that one of the reasons theory-based evaluation caught on in the mid-1990s was because it seemed to promise a resolution to evaluators' inability to show causal connections between program activities and program outcomes. In the absence of random assignment, a theory of change could "show what processes lead to the outcomes observed" (p.70, 1997). Thus, if some of those processes were not implemented, then the intended outcomes could not be expected to occur. Any issues in program theory or implementation are made evident by a solid theory of change, and the theoretical connections between activities and outcomes are made explicit.

Because the terminology used for evaluations based on theories of change is still in flux (see, for example, Dahler-Larsen, 2001; Rogers, 2007; Christie & Alkin, 2003; Coryn, Noakes, Westine, & Schroter, 2011), it is important to establish a shared definition before going further. We refer our readers to Weitzman et al.'s (2002) definition of a "theory of change," as it accurately describes our approach:

> Theory, in this context, refers to the specific guiding model of those responsible for developing, directing, and implementing the program. In this approach, evaluators work with program staff to make explicit the theory and assumptions implicit in the programs they are undertaking (p.373).

In the formal evaluation literature, the term "theory of change" is used less often than program theory, though there seems to be some vocabulary slippage. "Program theory" is the nomenclature often used to differentiate the *theory* being referenced in a theory of change and that of a *theory* drawn from social or behavioral sciences. Donaldson and Lipsey (2006) coined the term "program theory" to call attention to the distinctive type of theory being discussed in the evaluation literature. Also seeking to differentiate between theories drawn from social science or behavioral sciences, Chen and Turner (2012) offered the terms "stakeholder intervention theories" (aka program theories) and "formal theory-based interventions" (aka social science theories) (p. 396) to clarify what kind of theory is being used. While "stakeholder intervention theory" emphasizes the source of the theory and Donaldson and Lipsey (2006) emphasize the focus of the theory, both attempt to clarify that a social or behavioral science theory is not what is being used. Still, in our experience, neither of these terms is immediately understandable to a lay audience, and no single term is used exclusively in the evaluation literature.

Definitions of "theory of change" that have not been published in the evaluation literature, but which nevertheless are influential--at least within the nonprofit world we work within--come from capacity-

4

building nonprofit organizations. For example, Karlie Silver with the Skoll Foundation (2014), in extolling the virtues of using a theory of change, defined it as "a blueprint for achieving large-scale, long-term goals. It identifies the preconditions, pathways and interventions necessary for an initiative's success." We found this "blueprint" language helpful in discussing a theory of change with program stakeholders.

Theoryofchange.org, an organization devoted to promoting the use of theories of change and teaching people how to use them, defines a theory of change as "a comprehensive description and illustration of how and why a desired change is expected to happen in a particular context… by first identifying the desired long-term goals and then works back from these to identify all the conditions (outcomes) that must be in place (and how these related to one another causally) for the goals to occur." Theoryofchange.org's definition begins to lay out a process by which evaluators can work with program stakeholders to "make explicit the theory and assumptions implicit in the programs they are undertaking" (Weitzman et al., 2002, p.373).

Not everyone in the evaluation field has been open to using a theory of change. Stufflebeam (2001) and Scriven (1998) have argued against using a theory specific to the program being evaluated in part because, they argue, such an approach lacks sufficient validation and proven methodology.   Others in the evaluation field have noted that a theory of change in itself is insufficient to evaluate an initiative because it does not include the counterfactual (Hollister, 1998; Cook, 2000). While advocates argue that these shortcomings can be addressed through the use of empirical data (e.g., Weitzman et al., 2002), there are only a few examples in the literature that actually implement this combined theoretical-empirical approach (see, for example, Crew & Anderson, 2003). Despite not all being believers, in the field evaluators are using theories of change and finding them to be useful. In the next section, we

explore some of this theory-based work to provide a framework for our experience of how a theory of change can be used.

HOW THEORIES OF CHANGE HAVE BEEN USED

Among those writing on these topics, there is near universal agreement that a theory-based approach can be useful for program staff. As Michael Quinn Patton (1997) said two decades ago, program staff often have no idea what their implicit theory is. Evaluators can help program staff engage in a constructive exercise to develop a customized theory of change. Christie and Alkin (2003) argue that a theory-based evaluation can be responsive to program needs, so that relevance ensures utilization. Done thoughtfully, the process of developing a theory of change can in itself function as a capacity building activity for the organization and its staff (Weiss, 1997), as well as, we would argue, a method of participatory evaluation that will create heightened staff buy-in (Patton, 2012). When evaluators develop a theory of change with program staff, it enables staff to both articulate and understand the linkages between activities and intended outcomes. Program theory has been described as both explanatory and prescriptive for the organization (Chen & Rossi, 1987; Chen, 1990). Using a theory can improve programming, clarify goals, assess effectiveness of the program's overall implementation as well as its individual components (Chen, 1990).

The clarification and assessment aspects of theory usage are, of course, helpful not only to program staff but to evaluators as well. Christie and Alkin (2003) described how a program theory actually can define the evaluation through highlighting the important connections between program activities and desired impact. Crew and Anderson (2003) have used program theory to develop hypotheses that the evaluation could test. Coryn et al., in their 2011 review, describe the many different terms evaluators use to apply theory to their work in order to explicate a program model and/or investigate a program's

causal relationship to its outcomes. Breuer and colleagues (2016), in a review of 45 public health

evaluations with theories of change, found that evaluators used the theory of change in a number of

ways. These ranged from helping to design the evaluation plan to contributing to instrument

development, although most uses were not well articulated in their published articles. Many were done

in collaboration with stakeholders, though not all appeared to be. Some evaluators advocate using a

theory of change as an evaluation tool to measure the overlap or disparities between how an initiative

should unfold--per the initiative's theory of change--and how it actually does unfold (Weitzman et al.,

2002). But on the whole, there is little relatively little information on how evaluators used a theory of

change to evaluate program implementation (Breuer et al., 2014; Coryn et al., 2011)

There is little published about how evaluators develop a theory of change and even less about how the

theory of change was utilized by program staff. Our process of developing a theory of change was highly

iterative and time consuming, but yielded tremendous payoff in terms of organizational clarifying and

evaluation refinement. Thus, it felt important to share the process and outcomes with fellow evaluators

so that others can realize the same benefits, and hopefully, avoid some of the early pitfalls.

In the pages that follow, we describe the real world process of creating a theory of change for a complex

social change initiative, including the challenges we encountered during the creation process and the

implementation phase. We also describe the benefits we derived from the process of creation and the

final, completed theory of change, including examples of how it was used by program staff and by the

evaluation team.

A COMPLEX SOCIAL CHANGE INITIATIVE: THE NATIONAL CENTER FOR WOMEN & IT

NCWIT is a non-profit organization with a mission to significantly increase the meaningful participation of women in technology. To do so, it brings together corporations and startups, academic institutions, government agencies, and non-profit organizations to make the organizational and social changes in their various sectors that will enable movement forward on this entrenched, historic problem. It is an organization comprised of other organizations, providing them with evidence-based practices for recruiting and retaining girls and women in computing, from K-12 through workforce. Led by a former Bell Labs executive trained in computer science, NCWIT operates more like an agile start-up than the large national organization it is, as it is constantly creating new programs and trying new methods for engaging its stakeholders in largescale social change.

As of this writing, NCWIT employs about 45 part-and full-time staff and contractors, has more than 1,000 member organizations, and has a multi-million dollar annual budget through funding from the National Science Foundation and corporate sponsors. While each of the 40-plus programs and campaigns NCWIT runs could have a logic model of its own, many of its efforts can't even be described as programs, efforts such as social media communications, academic presentations and papers, primary research, policy discussions with government representatives, awareness-raising meetings with industry leaders, and so on. The complexity of the organization and its large reach left us as the evaluators with an evaluation challenge. Our challenge is common to evaluators of large, complex, multi-site initiatives that evolve quickly. One of the ways we chose to address the complexity was by developing a theory of change for the organization as a whole.

Our approach addresses a central shortcoming described in the literature that theory-driven approaches often do not go beyond program staff conceptions of how their program works, and that program activities and/or evaluation are not always aligned with the theory (Rogers, 2007; Weiss, 1997). Our

multi-year iterative process, including both internal and external evaluation, overcame these shortcomings.

PROCESS OF DEVELOPING A THEORY OF CHANGE

The organization was founded in 2004 without a theory of change, but with many good intentions, as most nonprofit organizations are. The approach NCWIT took was largely based on the assumptions that if they made sound social science research on the issues of diversity in technology accessible and user-friendly, and if they were successful in getting people to use these research-based practices, then they could transform the field of computing and make it more inclusive for women and other historically underrepresented minorities. In 2006 when the external evaluator was hired (Litzler), a logic model was created. Very soon after, this logic model felt inadequate to keep up with the growing, entrepreneurially organized nonprofit. In 2009, with an internal evaluator on board (DuBow), a theory of change began to be developed. There was precious little information on how to effectively develop a theory of change, let alone map one out succinctly. Back in 2009, when our process began, DuBow reviewed Weiss's (1995) book *New Approaches to Evaluating Comprehensive Community Initiatives* and spent time on theoryofchange.org, the Skoll Foundation, and the Aspen Roundtable discussions of theories of change.

DuBow began by interviewing each of the three founders about what prompted them to start the organization, what they had hoped to accomplish, and how they had imagined the organization would reach its goal of gender parity in technology. The interviews also included questions about how each one would describe the evolution in their thinking as they put thought to action. Following these interviews, DuBow interviewed the founding social scientists associated with NCWIT, asking about the state of knowledge in the field about how to tackle the problem of persistent underrepresentation of certain populations in the computing/information technology.

From the perspective of all of these key informants, when NCWIT was founded, the little research that

did exist about this issue did not get used by practitioners interested in reform. Instead, high school

computing teachers, post-secondary computing faculty around the nation tried approaches largely

unfounded in the education and social science literature. Or they developed interesting approaches in

parallel, without realizing what had already been done. Diversity was not on the tech industry radar yet

in the way it has since become as a result of high profile legal cases (e.g., venture capitalist Ellen Pao,

former Uber CEO Travis Kalanick, etc.) and a proliferation of articles in the popular press (e.g., *New York

Times*, *Fast Company*, *Atlantic Monthly*, etc.).

While the key informant interviews were a critical building block of creating this theory of change, the

founders spoke more in terms of strategies than theoretical constructs because they were coming from

a decidedly personal and practical perspective. Indeed, the early theory of change discussions at NCWIT

tended to devolve into discussions of action, or strategies, rather than tactics. They said things like,

"find/create facts for others to use," "develop and convene a network," "evaluate reform activities,"

and "create awareness." Although they could not step back enough to name them explicitly, their

underlying assumptions were articulated by DuBow (Table 1).

[Table 1. Early Assumptions of the NCWIT Founders]

Based on these interviews and discussions with external evaluator Litzler, DuBow began to develop the

first visual models of NCWIT's theory of change (Figures 1, 2), based on a model seen in Community

Change: Theories, Practice, and Evidence (Aspen Institute, 2006.)

[Figure 1. NCWIT Theory of Change (circa 2009)]

[Figure 2. Pop-out of the graphic in the Theory of Change describing the NCWIT three-pronged strategy]

The explanation of the graphic (Figure 1) was as follows:

To understand the NCWIT Theory of Change (a.k.a., the "pathway of change"), we start with the goal in green at the top. We hold these assumptions ("what we believe to be true") about achieving our goal. So, NCWIT assumes that: (1) Our mission will not be achieved without intervention of some sort, (2) No single organization can do it alone, (3) Shared information about how to increase women's meaningful participation is currently lacking, (4) An infrastructure is necessary to link those organizations with similar goals and allow sharing of practices and cross-fertilization of promising ideas, (5) Knowledge and awareness of the issue can, and should, be raised for critical parties to make changes, (6) Gender parity must be addressed at all stages of pipeline.

Achieving this goal requires certain preconditions ("what will need to occur"). We name five here, there are more. (1) Girls and women must have positive exposure to IT in order to pursue IT in meaningful ways, (2) Organizations throughout the pipeline must be aware of current gender biases and their implications, (3) Organizations throughout the pipeline must have knowledge of how to recruit, retain, and promote females in IT, (4) Evidence-based promising practices must exist and must be shared for orgs and individuals to move forward, (5) Alliances between key players are necessary to enhance the infrastructure.

This is where NCWIT's three-pronged strategy comes in. The three approaches are: (1) Build Capacity for NCWIT Member Organizations, (2) Create and Distribute Multimedia Resources, (3) Propagate Awareness and Outreach. Through increased knowledge and awareness, enhanced networking among pipeline organizations, increased capacity of frontline organizations (i.e., K-

12 schools, higher education institutions, workplaces, and entrepreneurs), and through reaching

enough organizations to achieve system-wide impact, NCWIT believes that it can affect a shift in

societal perspective on issues related to IT, and thereby, achieve its mission.

In sum, NCWIT provides the research, the networking linkages, the information and the

resources to help frontline organizations make the differences in the educational and workplace

environments to enable both incremental and radical changes until a cultural shift has been

achieved.

The graphic was shown to stakeholders of the organization, and explained as above. This theory of

change did not make sufficient sense to staff to be used other than as a single slide in a site-visit

presentation for a funder. At the time, DuBow's conclusion was that the organization was not ready for

a theory of change, that they were not able to step back from the activities they were doing and their

passion for their cause, to see it from a more abstract perspective. In retrospect, this may have been

true; however, it was also true that this template did not convey the theory adequately. In addition,

DuBow's status as a relative newcomer to the organization likely inhibited its adoption as well.

Working under the assumption that the theory was under-developed and under-explained, DuBow next

brought it to NCWIT's external Social Science Advisory Board (SSAB) in 2010. At the time, the SSAB

consisted of more than a dozen researchers from various social science post-secondary departments

around the country who were all studying some aspect of gender and technology. All of them had a

much longer tenure with the organization than had DuBow. DuBow presented them with some

background on theory of change, and her three-fold reason for pursuing a theory of change. These were,

alliteratively: "Exposing the assumptions we have as an organization; Ensuring that we are all on the

same page; Encouraging accountability."

The SSAB welcomed the development of a theory of change, and spent more than an hour discussing it at their meeting. Discussion included examining the weaknesses of some of the assumptions, the notion of whether or not a static theory of change was even possible for an organization as dynamic and entrepreneurial as NCWIT, and in the end, they provided contradictory feedback which was, by its nature, difficult to incorporate. Thus, the NCWIT theory of change went into hibernation, and the internal and external evaluations went on as before. The external evaluation focused on evaluating the effectiveness of the three-pronged strategy, and the internal evaluation went on evaluating the effectiveness of the individual programs and their own individual program goals.

In 2013, increasingly challenged by the quickly growing complexity of the organization, and emboldened by more sessions at the American Evaluation Association conference about theory of change, the internal and external evaluator once again turned to a theory of change as a method of explaining the organization's approach and underlying assumptions. We told fellow AEA members that NCWIT's theory of change explained why the organization decided to tackle the problem of low gender diversity in technology in the manner they had, describing how NCWIT believed these approaches would result in the ultimate outcome they hoped to achieve. In other words, the theory of change was intended to make explicit the organization's underlying assumptions and explain how short-term and medium-term outcomes move NCWIT toward its ultimate goal: Increased meaningful participation of women and other under-represented groups in technology as well as the increased innovation that will result from this participation.

We were motivated to develop a robust theory of change because we felt it would unify our work and that of the organization, but since no one else was yet convinced, and we had our regular evaluation duties to perform, this development could not be a front-burner activity. Thus, the authors spent about

two years developing the current theory of change. We drew many versions of the theory of change on scratch paper in the airport after AEA meetings, and in other stolen moments. Eventually, we began drafting graphics in PowerPoint (Figure 3), and began explicating each stage and its assumptions, writing a long explanation of each part of the theory of change, which, as of this writing, consists of a dense 14-page white paper. We felt comfortable initiating this without additional feedback from program staff and leadership because by that point we each had been working with NCWIT for several years (DuBow, 4 years and Litzler, 7 years).

[Figure 3. Early Version of NCWIT Theory of Change (circa 2013)]

When we felt it was comprehensible to others and had been sufficiently thought through, we brought it to the organization's leadership and staff for face validation. First, we discussed the entire prose explanation and a revised graphic (Figure 4) with the internal staff social science team (a 10-person team within NCWIT, consisting of social scientists from various disciplines, including education, communications, evaluation, sociology, etc.). From this conversation, we made many changes, most of them minor, many of them related to syntax. Surprised to have received so little criticism from this otherwise very analytical group, we incorporated all the changes that were easy to make, leaving other conceptual connections for a later date revision. Next, we brought it to the lead social scientist, the founding CEO, and the current assistant director (the only one who had not been privy to the 2009 theory of change).

[Figure 4. NCWIT Theory of Change (circa August 2016)]

This leadership group received it with enthusiasm, but also had comments on wording and questions

about connections between sections. They had read the two-page summary we had written, not the 14-

page white paper. Reminiscent of the early strategy vs. theory orientation of the founders, the directors

wanted to see all of the NCWIT programs mapped onto the theory of change [see Section VI. Challenges

Encountered]. They were so excited about having an overarching theory at this point that the founder

seized upon it as the organizing principle of the next National Science Foundation (NSF) grant proposal

the organization was preparing. In order to use a graphic as shorthand for the entire theory, DuBow

edited it according to their word preferences and other suggestions they had, then took it to the NCWIT

graphic design team to make it look more professional. Once the graphic had been redesigned, DuBow

was to take it to stakeholders as well as outside parties for talk-alouds.

It turned out that working with the graphic design team was in itself an excellent step. The designers not

only understand the organization from a completely different perspective than do the evaluators or

social scientists, but also are used to translating complex ideas into appealing and self-explanatory

graphics. Their questions and suggestions and our being forced to once again talk it through step by step

served to greatly enhance the illustration (Figure 5) as well as the two-page summary. Thus, working

with the graphic designers helped develop the theory of change to be more understandable to a broad

audience, as they brought not only their expertise communicating technical concepts to lay audiences

but also their perspective on visual communication. Working with them helped to ensure the theory of

change's core ideas and assumptions would be apparent from the graphic alone, and it helped us to be

even clearer in our two-page summary as we clarified sections that had been sticking points in their

understanding.

[Figure 5. NCWIT Theory of Change (circa October 2016)]

At AEA 2016, DuBow asked five evaluators of computing initiatives, and one first-time AEA attendee who did not work on STEM evaluation, to talk through what they thought the theory of change graphic (circa October 2016) was depicting. They mostly understood it the way we intended it to be understood, but they had a few common misconceptions, which DuBow brought back to the design team and Litzler to revise the graphic to its current version (Figure 6).

[Figure 6. NCWIT Theory of Change (circa December 2016)]

The evolution of the theory of change reflected the incorporation of diverse stakeholder perceptions, vocabulary changes in messaging over time, and graphic improvements to make it simpler and more easily understandable. NCWIT's essential philosophy and approach remained consistent throughout this time.

HOW THE NCWIT THEORY OF CHANGE HAS BEEN APPLIED

There are three main ways in which the NCWIT theory of change has been used: (1) to prioritize NCWIT activities, (2) to provide an organizing principle for funder reports and proposals, and (3) to shift the external evaluation framework to make a stronger connection between activities and outcomes (Core Principle 3 in Coryn et al, 2011). These uses are described below.

Theory of Change Helps to Prioritize Activities

It became clear when we examined our entire organization's activities that most were taking place in the "Necessary Conditions" column rather than in the later stages of the theory of change. A number of

shifts occurred in response to this observation. The authors had conversations with NCWIT leadership to discuss gaps identified by the theory of change and how NCWIT might undertake new activities, or re-develop current activities, to fill those gaps. Also, the NCWIT social science team, a team responsible for deciding upon and developing new resources (i.e., mission-related outreach and educational materials), refocused which types of resources should be made, and which should not, after observing that key areas of the theory of change graphic were being relatively overlooked, such as moving stakeholders from awareness to action.

Also, NCWIT program managers started looking to the theory of change as a reference to help guide their activities. For example, the program staff working with NCWIT's industry member organizations used it to prompt "self-assessments" about where each organization is on the continuum pictured in the theory of change—i.e., learning, awareness, individual change, organizational change. Another program focused on K-12 initiatives, reported that the theory of change allows them to "say no to some new ideas" whereas before they felt compelled to take on everything suggested to them. Broadly, the evaluation team has found that the theory of change provides a shared vocabulary and point of reference, which has helped make staff discussions more productive. The shared vocabulary has also helped staff with how to present the NCWIT approach as part of a *systemic* approach to change.

Theory of Change Serves as Organizing Principle

NCWIT has used this theory of change in reporting to funders and in proposals for new funding. The NCWIT leadership decided to use the theory of change as an organizing mechanism for a poster session held for one of its major funders. A simplified version of the theory of change became an icon on each of the posters to indicate how a given program area related to the theory of change (Figure 7). Visitors of any program poster could then see which part(s) of the theory of change that program related to, and

staff at the posters were instructed to talk about the programs in the context of the overarching theory

of change.

[Figure 7. Wayfinding Icon Example for Poster Session with Funder]

In addition, after conversations with organization leadership about gaps that the theory of change made

visible in [NONPROFIT ORGANIZATION'S] menu of programs and approaches to that point, the

leadership team tailored its next large funding proposal to include new strategies and activities, and to

remove other activities. In other words, the theory of change brought into relief areas the organization

was not addressing sufficiently (e.g., prompting stakeholders to action) as well as areas the organization

had been over-addressing (e.g., raising awareness). The educational and informational materials NCWIT

has been known for have historically focused on raising awareness about diversity issues and increasing

knowledge of how to make change. Materials produced post-theory of change, however, have been

focused much more explicitly on social change by sharing specific strategies and tools with new and

more actionable content.

Theory of Change Shifts External Evaluation

Lastly, the external evaluation team shifted the evaluation to be more in line with the theory of change.

Litzler and her team re-organized the metrics around the theory of change so that the connection

between NCWIT activities and expected outcomes would be stronger and easier to understand. The

theory of change forced the evaluators to focus more on outcomes rather than outputs , as it clarified

which outcomes were critical to the different stages of the theory of change. For example, Litzler's team

stopped tracking the number of press releases NCWIT produced each year, focusing instead on the types

and sources of press mentions in order to gauge influence and largescale social change. The theory of

change precipitated a shift in outcome focus as well. While the early years of the external evaluation had focused on stakeholder learning and awareness, now the team began collecting more data around how organizations and individuals were enacting change. New survey items were created to focus on how individuals were changing related behaviors and the extent to which organizations were making changes to improve gender equity.

This shift in the evaluation is enabling the external evaluator to tell a more compelling story about the organization and its outcomes. Data from this new individual and organizational change metric has already proven to be a much more direct measure of the organization's impact on social change. Previously the member survey had asked whether or not respondents collaborated with someone new as a result of membership (one of several indirect measures of increased capacity). Although the theory of change highlighted weaknesses in the evaluation, it also provided us with an opportunity to improve the evaluation. This tension highlights how the use of a program theory can be both a benefit and a challenge to evaluators.

CHALLENGES ENCOUNTERED

We encountered both expected and unexpected challenges, during the theory's development as well as during its application.

Development Challenges

The development of the theory of change over such a long time period required us to make decisions about how to incorporate our retrospective interviews with founders and the evaluators' experience over the years into a theory of change that would fit the current organization and its current goals. It didn't make sense to stick with a theory of change that fit the organization in 2009, since the

organization had made significant changes since then. While the mission had stayed relatively constant, the organization structure and strategies had evolved quite a bit, as had the language used to describe them. For example, the three-pronged strategy language changed over the years, from Build Capacity for NCWIT Member Organizations, Create and Distribute Multimedia Resources, and Propagate Awareness and Outreach in 2009, to Convene, Equip, Unite in 2013. These shifts were not only linguistic, but also strategic, and had to be incorporated into our framing of the model. We were able to connect the early and later theory of change work because of our embeddedness in the organization; that is, we knew what "convene," "equip" and "unite" meant both conceptually and tactically.

We also faced logistical challenges that affected our progress. We both had very full to-do lists each day, and practical urgencies abounded, so it was difficult finding time to be theoretical and do the deep thought-work that was needed. Designating two mornings a month for meetings, before other work could take over, proved to useful for our collaborative thinking and writing.

Some of our thought-work was focused on breaking the model down into small individual pieces and explaining the assumptions of each piece, and then revising to try to express these complex ideas more simply. There were countless iterations of the model that occurred every time we brought in a new perspective. Sometimes the new perspectives were our own after we had taken a break from working on the model and found we wanted to revise both the prose explanation as well as our early versions of the graphic. Working on both the graphic and the prose really helped to strengthen both of them, as one would highlight issues with the other. Still, it was challenging to come back to this project after a multi-month break and have to completely re-familiarize ourselves with our conceptualizing to that point; on the other hand, we found that time away sometimes was essential to enabling us to see gaps and inconsistencies.

Challenges in Application of the Theory

It exceeded our hopes that the theory of change has been applied beyond the evaluation team and used by the organization's staff and leadership, thus improving both program planning and the impact evaluation. While we relish this as a success, there were nonetheless challenges in the application of the theory of change, which are important to note.

One concern we had about application of the theory was in relation to the organization's staff. Because we worked for more than two years to fully develop the theory of change, we were concerned that staff would not buy in to the theory when it was re-introduced. Thankfully, this was not the case (as demonstrated by the many ways described above that organizational leadership and staff embraced the theory of change). We also had concerns about whether the organizational leadership might reject the model because it could highlight organizational weaknesses, such as linkages where the organization was not spending enough programmatic time. Instead, leaders were interested in exploring the theory of change model to understand how it could help them develop funding proposals and make programmatic decisions.

As Carol Weiss (1997) has pointed out, one challenge theory-based evaluations face is that the evaluation sometimes does not actually follow the program theory. We found this to be somewhat true in our case because the original evaluation framework was created in 2006. Once the theory of change had been confirmed with the relevant stakeholders, the external evaluator used it to reorganize the evaluation framework. This was a significant and challenging undertaking which took many hours of examining and revising the current framework and metrics. On the part of both the evaluators and the program staff, there is perceived value in longstanding metrics, so retiring them was not an action taken

lightly. Some of the metrics had been collected for 10 years, making it especially difficult to decide to retire them, given the sunk costs invested in them. In this process, we identified areas of the theory of change that lacked compelling metrics, thus highlighting gaps in the evaluation. The starkest example of this is that the theory of change "Actions" column did not have corresponding metrics or measures; we added new survey items to assess individual and organizational change, and conducted interviews with a sample of NCWIT's members. This was helpful to the evaluation in the long term, but forced a recognition that weaknesses had existed prior. It also prompted us to examine the linkages and assumptions in the theory of change. For example, we are now using longitudinal survey data to understand the extent to which learning something new and using NCWIT's materials are impacting our member academic institutions' female enrollment or graduation numbers.

This application of the theory of change to the evaluation challenged us from a time and energy standpoint (there were 40+ metrics before the reassessment) as well as from an emotional standpoint. It was difficult to relinquish established ways of doing this evaluation and to recognize its weaknesses. We asked ourselves, "How had we not measured that before?" We felt vulnerable, but it gave us the opportunity to grow as evaluators and improve the work we do for our clients. Fortunately, NCWIT's leadership was receptive to changes because she recognized this shift would yield more and better impact data. The decade-long relationship and trust the external evaluator had built with the organization contributed to the organizational receptiveness to shift gears.

As both a challenge and a sign of positive progress, we found that NCWIT staff and leadership really wanted to map each of NCWIT's many programs onto the theory of change. As mentioned earlier, staff chose to use the theory of change as an organizing mechanism for a poster session for a major funder. We found this heartening as it reflected their buy-in to the theory, and their desire to use it to

understand the organization better. However, their desire to map programs reflected a lack of understanding of how the theory was supposed to work. Few programs lie in only one sector of the theory of change, so programs could not simply be overlaid onto the theory of change model. For example, a program may spark learning and awareness *as well as* lead to organizational change. Still, this desire on their part prodded us to explain the theory of change in more practical terms, helped them be more comfortable with it, and made it more likely they would use it appropriately. It also prompted us to point out which areas of the theory of change were well addressed by NCWIT's various programs, and which areas were less well addressed.

Moving forward, the challenge most critical for the authors to consider is twofold: First, many of the people to whom we introduced the theory of change did not understand, without explanation, the distinction the evaluation community would make between "program theory" and "social science theory." There was an expectation that our "theory of change" would be based in some sort of social science theory derived from research in a relevant field—whether it be a theory about organizational change, or a behavior modification theory, for instance. We need to do a better job of explaining this distinction. Second, even after a thorough explanation of the theory of change and how the evaluation is examining the linkages to assess impact, some audiences still felt that a causal connection was needed, and missing from the model.

CONCLUSIONS

To develop a well-thought-out theory of change requires intensive contemplation and collaboration time, as well as input from many parties, all of which is difficult to find in our busy schedules. Fortunately in our case, our timing turned out to be fortuitous, as program staff and leadership were

receptive to having an overarching framework at the point when we had completed it--and perhaps would not have been prior. Leadership had been looking for a way to bring together all NCWIT had done and planned to do in the future to tell a coherent story to funders and stakeholders, and the theory of change served that function. Indeed, staff and the leadership were eager for a way to tie together the disparate NCWIT programs and strategies under a single umbrella that didn't oversimplify and could absorb the entrepreneurial evolution of the organization over time.

The theory of change was a priority to leadership during proposal writing time, but it has not yet translated to a priority to sharing broadly via the NCWIT website, in part due to a backlog of website edits to be made. However, NCWIT leadership is sharing the model with their Board of Directors, and is happy to have us share it as well. In addition, the theory continues to have internal influence in programmatic and strategic discussions.

While we valued engagement with the organization's staff and leadership in developing this theory of change, we didn't involve them at every stage, but rather at the beginning and in the final stages. The hard thinking came from the evaluators in response to observing the organization, acting on feedback, and listening to staff and leadership discuss the organization's activities, strategies, and future plans. Realistically, staff and leadership could not have given us sufficient time and energy at critical thinking points. In addition, the various stakeholders would have introduced so many disparate perspectives that the project simply could not have moved forward with democratic buy-in. Thus, we do not advocate a completely participatory approach. At the same time, we also do not advocate developing a theory of change in isolation from the organization. In our long-time embedded roles as evaluators, we had an informed vantage point from which to develop a theory that could be digestible to all stakeholders. We contend that the evaluators' simultaneous insider/outsider perspective is powerful and should not be

dismissed as a tool for developing a theory of change. Also, while the design process was not intentionally drawn out over time, the lengthy time period ended up being fruitful for our development.

While we would like to believe that the successful utilization of this theory of change is due to our hard work, we recognize that the fortuitous timing and context of its re-introduction was helpful. Thus, evaluators may want to consider the timing and context when sharing a theory of change.

This theory of change development process was strengthened immensely by the collaboration of the internal and external evaluators. We each have a different view into the organization, with deeper knowledge in certain areas. This collaboration was key to a final model that resonated with the organization.

As a result of our experience, we offer these recommendations for evaluators interested in developing a theory of change:

- Collaborate with someone else who can think and talk about the organization from a high-level perspective. It is critical to understand program specifics without getting bogged down in them.

- Put the theory of change in front of people with different perspectives when you are at development junctures where you can incorporate feedback.

- Be prepared to write up a long white paper that elaborates on the precise details of the theory of change, and at the same time, recognize that no one else will read it. Our experience suggests this long, detailed write-up is a necessary step for deep thought and collaboration.

- Be open to the theory of change highlighting weaknesses in the evaluation. Be prepared to let go of long-standing metrics and develop new ones.

- Develop a standalone graphic to convey the theory of change. Unlike a logic model, this graphic cannot be built from a template because it needs to grow out of, and reflect, the organization's specific theory of change. Working with a graphic designer is helpful, but may not be necessary, as long as feedback on the graphic is collected from individuals who have graphic design expertise.

- Write a one- to two-page summary that explains both the theory of change concepts and linkages and explains the graphic. This is not only a useful exercise for communicating the theory of change in a digestible format, but also for testing how well your graphic works to convey the essence of the theory of change.

- Don't be concerned if the development process is taking a long time. Just make sure that the final product is relevant to the organization's objectives and terminology at the time you introduce it. You can use a long development period to your advantage to incorporate changing perspectives and test the robustness of the theory of change over time.

- When preparing for utilization, look for points of organizational receptivity and ways the organization can leverage the work immediately.

Ultimately, we learned that even after an organization has been underway for many years, it can still be a clarifying and useful exercise to develop a theory of change. Internally, it can provide a complex initiative with an organizing principle by rendering explicit what had heretofore been conceptual and, perhaps, not understood in the same way by all staff and leadership. Because most nonprofits' programmatic interventions are not started by evaluators or theorists but rather by well-meaning activists, developing a theory of change can prove to be a fruitful way of clarifying the organization's future strategies. Most importantly for this audience, we found that the painstaking development of a

thoughtful theory of change greatly improved our ability to measure the metrics most appropriate to

determining a complex organization's success.

In light of some stakeholder's comments that causation between elements was missing, we have

realized that we can improve our communication of the evaluation that accompanies the theory of

change. We can clarify how the outcome metrics will demonstrate impact. We are also working on

applying theory of change-based evaluation questions for all of NCWIT's programs, so that program staff

can self-assess the extent to which their programs bring the organization closer to its intended impact.

By making use of outcome data in relation to the organization's theory of change, we can implement a

combined theoretical-empirical approach that will enable a fuller understanding of this complex

organization's overarching impact.

REFERENCES

Aspen Institute Roundtable on Community Change. Community change: Theories, practice, and

      evidence. Fulbright-Anderson, K. & Auspos, P (Eds.) (2006). Washington, D.C.: Aspen Institute,

      2006.

Breuer, E., De Silva, M. J., Fekadu, A., Luitel, N. P., Murhar, V., Nakku, J., & Lund, C. (2014). Using

      workshops to develop theories of change in five low and middle income countries: lessons from

      the programme for improving mental health care (PRIME). *International Journal of Mental*

      Health Systems, *8*(1), 5.

Breuer, E., Lee, L. De Silva, M. & Lund, C. (2016). Using theory of change to design and evaluate public

      health interventions: a systematic review. *Implementation Science*, 11, 63.

Chen, H.T. (1990). Issues in constructing program theory. *New Directions for Evaluation, 47,* 7-18.

Chen, H. T., & Rossi, P. H. (1987). The theory-driven approach to validity. *Evaluation and Program*

      Planning, 10, 95-103.

Chen, H.T. & Turner, N.C. (2012). Formal theory versus stakeholder theory: New insights from a tobacco-

      focused prevention program evaluation. *American Journal of Evaluation*, *33*(3), 395-413

Christie, C.A. & Alkin, M.C. (2003). The user-oriented evaluator's role in formulating a program theory:

      using a theory-driven approach. *American Journal of Evaluation*, *24*(3), 373-385.

Christie, C.A. & Azzam, T. (2005). What theorists say they do: A brief description of theorist's

      approaches. *New Directions for Evaluation, 106,* 15-26*.*

Cook, T.D. (2000). The false choice between theory-based evaluation and experimentation. *New*

      Directions for Evaluation, 87, 27-34.

Coryn, C. L., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2011). A systematic review of theory-driven

      evaluation practice from 1990 to 2009. *American Journal of Evaluation*, 32(2), 199-226.

Crew Jr, R. E., & Anderson, M. R. (2003). Accountability and performance in charter schools in Florida: A

    theory-based evaluation. *American Journal of Evaluation*, *24*(2), 189-212.

Dahler-Larsen, P. (2001). From programme theory to constructivism: on tragic, magic, and competing

    programmes. *Evaluation, 7*(3), 331-349.

Donaldson, S.I. & & Lipsey, M.W. (2006). Roles for theory in contemporary evaluation practice:

    Developing practical knowledge. In I. Shaw, J.C. Greene, & M. M. Mark (Eds.) (2006). *The*

    handbook of evaluation: Policies, programs and practices (pp. 56-75). London, England: Sage.

Fawcett, S.B., Boothroyd, R., Schultz, J.A., Francisco, V.T., Carson, V., & Bremby, R. (2003). Building

    capacity for participatory evaluation within community initiatives. *Journal of Prevention &*

    Intervention in the Community, 26 (2), 21-36.

Fetterman, D., & Wandersman, A. (2005). *Empowerment evaluation principles in practice*. New York:

    Guilford.

Hollister, R. (2003). The growth in after-school programs and their impact. *Brookings Institute*. Retrieved

    from: https://www.brookings.edu/wp-content/uploads/2016/06/20030225.pdf

Kubisch, A. C., Fulbright-Anderson, K., & Connell, J. P. (1998). Evaluating community initiatives: A

    progress report. In *Fullbright-Anderson, K., Kubisch, A.C., & Connell, J.P. New approaches to*

    evaluating community initiatives: Theory, measurement, and analysis (Vol. 2).

Patton, M.Q. (2012). Utilization-focused evaluation, 5th edition. Thousand Oaks, CA: Sage.

Rogers, P. (2007). Theory-based evaluation: Reflections ten years on. *New Directions for Evaluation 114*:

    63-67.

Rogers, P. J., Petrosino, A., Huebner, T. A., & Hacsi, T. A. (2000). Program theory evaluation: Practice,

    promise, and problems. *New Directions for Evaluation*, 87, 5–13.

Scriven, M. (1998). Minimalist theory: The least theory that practice requires. *The American Journal of Evaluation*, *19*(1), 57-70.Silver, K. (2014). Theories of change and common metrics drive search for big impact in global health.

Skoll Foundation Archives. Retrieved from: http://archive.skoll.org/2014/08/18/theories-change-common-metrics-drive-search-big-impact-global-health/

Stufflebeam, D. (2001). Evaluation models. *New Directions for Evaluation*, *2001*(89), 7-98.

Weiss, C.H. (1997). Theory based evaluation: Past, present, and future. *New Directions for Evaluation*, *76*: 68-81.

Weiss, C. H. (1998). *Evaluation: Methods for studying programs and* policies (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Weitzman, B.C., Silver, D. & Dillman, K. (2002). Integrating a Comparison Group Design into a Theory of Change Evaluation: The Case of the Urban Health Initiative. *American Journal of Evaluation*, *23*(4): 371-385.

Wholey, J.S., Hatry, H.P., & Newcomer, K.E., (2004). *Handbook of practical program evaluation*. Jossey-Bass, 2nd ed