Changes in Library Evaluation: Responding to External Pressures in the Institution of Museum and Library Services’ *Measuring Success Imitative* for the Grants to States Program

This chapter examines the challenges of developing and implementing a new national evaluation approach in a complex library funding program.  The approach shifts a prior outcome-based evaluation legacy using logic models to one relying on non-linear logic mappin*g.*  The new approach is explored by studying the *Measuring Success* initiative, launched in March 2011 for the largest funded library services program in the US, the Institute for Museum and Library Services formula based Grants to States program. The chapter explores the relative benefits of non-linear logic maps and emphasizes the importance of scaling evaluation from individual projects towards clusters of similar library services and activities.  The introduction of this new evaluation approach required a new conceptual frame, drawing on diffusion, strategic planning, and other current evaluation theories. The new approach can be widely generalized to many library services, although its focus is on a uniform inter-organizational social network embedded in service delivery.  The chapter offers a new evaluation perspective for library service professionals by moving from narrow methodological concerns involving measurement to broader administrative issues including diffusion of library use, effective integration of systematic data into program planning and administration, and strengthening multi-stakeholder communication.

Key words:  Institute of Museum and Library Services, Grants to States, Library and Services Technology Act, evaluation, results based management, outcome based evaluation, logic models, logic mapping, systems theory, strategic planning, diffusion, social networks.

Matthew Birnbaum, PhD, Institute for Museum and Library Services

Kim Okahara, Institute for Museum and Library Services

Mallory Warner, Institute for Museum and Library Services

**Corresponding Author**

Matthew Birnbaum, Institute of Museum and Library Services, 1800 M Street NW, 9th Floor, NW, Washington, DC 20036.

Email: mbirnbaum@imls.gov**Introduction**

The US political environment creates intense competition for social services and informal learning at all levels of government. This is not expected to change in the foreseeable future given economic circumstances. Evaluation research may be the most effective tool to help libraries respond to ever-changing technology demands, expanding diversity, and rapidly changing social needs in this political-economic environment.

The Institute for Museum and Library Services (IMLS) grant programs have incorporated program evaluation for about fifteen years, but the practice remains nascent when compared to that of related sectors such as education, social services, and public health. Improving the quality and usefulness of library program evaluation requires a cultural shift in how the field approaches planning for the programs it plans and administers, and the role of assessment in both.

Library educators, executives, and professionals must alter the customs embedded by a prior generation of evaluation models. Lacking any straight-forward blueprint for change, this shift requires continuous adaptation to enable and sustain clearer articulation and understanding of what library programs intend and what actually happens. Constrained library budgets, increased demand for library services, and heightened requirements for persuasive evidence of tangible results exacerbate the need.

This paper describes an initiative by IMLS and its state grant partners to revise program evaluation protocols for the largest US-funded library services program, formula grants to State Library Administrative Agencies (SLAAs) the Grants to States program under the Library and Services Technology Act (LSTA). The authors and collaborators have led this initiative, *Measuring Success*, since its inception in March 2011.

The next section summarizes theory shaping the initiative. The subsequent section describes progress of *Measuring Success* to date. The final section offers lessons learned to guide future evaluation research and practice in library services.

**Theoretical Background**

Program evaluation includes more than monitoring and measurement. It involves systematic thinking about a program, raising meaningful questions, gathering and assessing evidence to provide answers, and applying all to strengthen a program (Russ-Eft and Preskill, 2009). There are as many approaches to program evaluation as to library program administration.

Library programs are dynamic and continuously shaped by external influences; useful evaluation also relies on the capacity of library program administrators to effectively adapt. When external influences shift substantially enough, the evaluation approach also must change to retain value. Two primary reasons may make an evaluation approach stagnant and ineffective, consequently requiring adaptation. First, people act surrounded by uncertainties, and as in other areas of society, the knowledge base changes continuously. A ten-year-old “best practice” may no longer reflect the strongest current possibilities. Second, people resist change. Innovation takes effort. Repeated and significant change can cause burn out, compromising the effectiveness of an evaluation approach selected in the past.

### Federal emphasis on evaluation started in the late 1960s with the growth of Great Society Programs (see, e.g., Weiss, 1972). Evaluation practice matured as part of a larger public administrative shift towards “results-based management,” known more commonly in library services as “outcome-based evaluation” (OBE) (see, e.g., Rossi, Freeman, and Lipsey, 2004). OBE followed a similar trajectory in other public and non-profit settings around the world, introducing new public management approaches that sought to better integrate business practice to governmental settings (see, e.g., Boston et al., 1996).

### A major emphasis of OBE was to turn attention away from strict monitoring of whether funds were spent appropriately and protocols followed properly. This “input” approach to management was replaced by a private sector concern with bottom line results. The private sector challenge in prioritizing results ultimately boils down to maximizing profits, but the core philosophy driving the public and non-profit sectors’ bottom line focuses on beneficial changes to the public, and particularly to segments of the public that are targeted because they share a specific need.

### As the new approach to public administration matured, OBE, tightly interwoven with logic models, became a dominant feature of the evaluation landscape in the US government and elsewhere. Multiple federal agencies, and philanthropic organizations such as the Centers for Disease Control and the United Way, have been in the forefront of this approach for several decades. It has remained the dominant federal approach since the 1990s in response to directives by the federal Office of Management and Budget (OMB) and the White House, associated with the Government Performance and Results Act of 1993 (GPRA). Consequently IMLS began to introduce OBE and logic models to their grantee communities as key tools for planning and evaluation by the latter part of that decade.

The features of the basic logic model are easy to understand. A set of program “inputs” like labor, funding, and capital are allocated. Inputs shape program “activities” (or services). The activities lead to products or “outputs,” which in turn lead to “outcomes” (see Figure 1.1).

***Insert Figure 1.1***

Generally, outcomes constitute immediate changes in program participants. Over time, longer-term outcomes, if they are sustained and shared with larger populations, are expected to create more sweeping changes in participants, communities, and/or other segments of society—these are often called impacts. Once a program logic model is complete, metrics are developed to assess whether projected changes in outputs, outcomes, and, ideally, impacts, happen and to what extent.

Consider an early child literacy program. A logic model articulates program intentions. Funds and staff are allotted as inputs. Staff designs and administers activities like curriculum design, instructor training, and outreach. These activities lead to an output of direct participation for pre-school or very early readers (and usually caregivers). As a result of participation, the children are expected, for instance, to increase their cognitive capacities and/or to become better prepared to participate in schooling. It is anticipated that the participating children will experience longer-term improvements in learning and higher academic success (impacts). Once a logic model and its embedded theory is complete, metrics can assess the validity for any included propositions. For instance, output metrics might count the number of participating children. Outcome metrics might estimate changes in participating children’s cognitive abilities during the program’s life.

The logic model approach is useful for program planning and assessment. It helps practitioners clarify program intentions by articulating an underlying logic. It helps frame data for collection and analysis. It moves administrator and policy maker focus to program performance by shifting attention from inputs and towards outputs, outcomes, and impacts.

Unfortunately, the model has four major weaknesses. First, the linear design can oversimplify the projected sequence of change in a frequently non-linear world. Second, the model does not test alternative explanations or contradictions. At best, it recognizes external influences. Third, in practice the model generally lists *arrays* of potentially causal items under the various headings (inputs, activities, outputs, outcomes, and impacts), complicating both theoretical program logic and its subsequent assessment. Fourth, the model presumes objectivity, but programs are often created and implemented through processes requiring political compromise among individuals with different ideological perspectives, interests, and influence, situated across different institutional settings (Weiss, 1995).

Despite its weaknesses, the logic model remains highly recommended and applied in the US government and elsewhere. Its terms have become part of the vernacular of global public administration. While many evaluators grumble openly about the model’s limits, no cohesive alternative replaced it. A group of evaluation theorists and practitioners, many affiliated with the Systems Thinking Topical Interest Group in the American Evaluation Association, has begun to seed an alternative (see, e.g., Rogers, 2008 and 2010). Current social science discussions surrounding complexity theory are moving away from the mechanistic features of the logic model and towards more subtle interweaving of program dynamics within a larger environment and embracing complexity theories. For instance Patton (2011) has introduced a framework for logic mapping using non-linear interactions at individual, organizational, and societal levels.

Two limits of the logic model are significant for this paper. First, the societal impact of most public programs is due to non-linear “tipping points” (Schelling, 1971; Gladwell, 2000). These arise when individuals act based on their perceptions of others’ experiences. As ideas gain traction among a segment, they reach a tipping point where others become increasingly willing to adopt the notion. Then the notion “sticks” and becomes part of common lore (Heath and Heath, 2007); diffusion happens quickly and at a large scale. Second, external factors frequently influence both a program and the responses of participating individuals. Programs do not operate in a vacuum; they must be considered within their larger environment.

Both limits affect an important theoretical issue underlying IMLS’s *Measuring Success* initiative*.* Library services and programs operate on an implicit “theory of change”[[1]](#footnote-1) associated with informal learning. A prevailing interpretation of informal learning is based on variations of this theory of change: acquisition of some *knowledge* leads to *attitude* change that in turn leads to a change in *behavior*. The “KAB” theory of change (Figure 1.2.a) remains dominant in the social sciences and associated professional fields, including library services.

***Insert Fig. 1.2.***

While elegant, KAB theory is overly simplistic. Social scientists and other learning theorists have invested substantial energies over decades to devise and test substantial refinements of KAB theory and others to understand how learning can lead to desired behavioral (or other types of) change. While nothing has yet removed the basic KAB theory from its pedestal, this is a central part of the arena. [[2]](#footnote-2)

One leading alternative with supporting empirical evidence involves diffusion theory (see groundbreaking work of Rogers, 1962). This theory is illustrated in Figure 1.2.b.

Diffusion theory works at both individual and societal levels. The top row of Figure 1.2.b. focuses on individuals. Diffusion theorists divide individual learning change into five steps. First, an individual becomes aware of desirable *knowledge*. Then this person must be *persuaded* of its value. Once persuaded, the person must *decide* to access it (e.g., through program participation). Afterwards, the person must *implement* the knowledge (i.e., apply it) to validate usefulness to his/her circumstance. If the net benefit is positive, the individual ultimately *confirms* its personal value and adopts the desired behavioral change. In short, this model suggests that any program seeking to benefit individuals through *knowledge* must first persuade them that acquiring the knowledge is worthwhile, and then hope the individual decides it is valuable enough to sustain it through a behavioral change.

There are other applicable theories beyond KAB and diffusion. Using insights from strategic planning theory (see, e.g., Bryson, 2004), program staff contend with influences outside direct organizational control (external influences) in seeking to affect target populations. These involve “opportunities” and “threats” (or “barriers”). Consider again a library early childhood literacy program. Even if program staff appropriately target a group of families who could greatly benefit from such service, that group may not participate even if they agree gains will result. Their reasons may not be irrational but reflect perceived barriers like practical transportation access or child care constraints. Conversely, innovations in information technology may offer new opportunities for participation in childhood literacy programs, even without changes in target group awareness of its benefits.

As Figure 1.1.b also shows, diffusion theory provides insight into how individual change can aggregate into societal change.[[3]](#footnote-3) Using economic terminology, “spill-overs” happen when something affects one individual and then spreads to another. Diffusion theorists articulate such spread by assuming that some individuals are more prone to lead and others to follow. While Rogers’ (1962) diffusion theory described five types of individuals, people more realistically fall on a continuum. At one extreme, there are risk-takers willing to immediately adopt innovation, like participating in a new library program. Others, more cautious in varying degrees, will wait and assess others’ experiences before opting to participate and adopt their learning to change their circumstances. At the far extreme are risk-adverse individuals who may never choose participation and potential change.[[4]](#footnote-4) This aspect of diffusion theory is key in understanding how library programs can gradually increase popularity and impact; it is based on the continued intermingling of former, current, and potential participants and not just on the direct interactions between library staff and some clientele at one point in time.

We now turn our attention to the IMLS’s *Measuring Success* initiative Grants to States.

**Political Context for Evaluation in Grants to States**

There are three important policy features of IMLS Grants to States that have influenced choices for building a refined evaluation system. First, this has always been a formula-based federal program, resulting in complex intergovernmental relationships. The federal authorizing legislation places responsibility for the program’s oversight in IMLS, but provides SLAAs with substantial discretion in the services and activities they carry out. Further, many SLAAs have redistributed a large share of funds to local public libraries through project grants to tailor services and programs to various communities. Consequently, three governmental actors interface--IMLS, the SLAAs, and local libraries. Much multi-level cooperation is needed. Perhaps reflecting the absence of regulatory laws, the level of intergovernmental cooperation between IMLS and the SLAAs always has been high.

Since the formal parameters of Grants to States have been shaped through the legislative process, like most such products, its intentions are purposely broad. The LSTA has always stated over-arching purposes, such as rural library services, technology infrastructure, and more. Recently, library-to-library collaboration and library partnerships with non-library organizations have been highlighted. The LSTA however has never articulated explicit outcomes through this program that are intended to benefit participant circumstances.

Third, LSTA legislation for Grants to States is tied to Congressional authorization for IMLS. Since the first Library Services bill was signed on June 19, 1956, each federal authorization has stipulated that each state receiving funding must provide a formal plan for its use. The federal government began to institute evaluation requirements in 1990. IMLS was created in 1996 with funding both the library and museum communities amalgamated in the Congressional legislation. That legislation required that each SLAA conduct an independent evaluation of the results of its five-year plan. Processes were left to each SLAA. The federal legislation however complicates fulfilling evaluation requirements due to a discontinuity in the timing of the required SLAA five-year plans and five-year evaluations. As a result, the five year SLAA evaluations can only assess three to four years of their five-year plans. The evaluations and plans nonetheless are perceived as important contributors to transparency and to strengthening agency level performance at a federal level even if research strongly suggests that elected and appointed senior governmental policy makers rarely read the required evaluation reports (McDavid and Huse, 2012).

OBE based on logic models was first introduced to the Grants to States Program in 1997 and applied through 2011 It is the focus of the next section.

**Introduction of Logic Models to Grants to States: 1997-2011**

Grants to States, like nearly every other public program, historically has focused its administrative oversight on monitoring inputs to ensure that funds have been spent properly. While the 1990 and 1996 federal legislation highlighted increased federal concern for monitoring outcomes, fiscal accountability has remained part of the dominant program culture.

As IMLS began to offer guidance and leadership in OBE to Grants to States in response to the federal evaluation directives, it had to balance several concerns. First, it had to balance the monitoring of inputs and their outcomes. Second, as a state formula grant program, it had to balance federal government interests versus those of state governments. Third, since many SLAA investments were made through project awards to local public libraries, any new type of evaluation approach introduced to Grants to States had to be integrated with project grant administration.

Against this backdrop, IMLS began to introduce OBE with logic models to the Grants to States program in 1997. It hired an evaluation officer and contracted with a third-party expert in program evaluation. Training workshops were introduced to participating SLAA partners in using logic models for planning and assessment of outcomes. Every state had to select at least one project for OBE assessment. IMLS staff, in turn, availed themselves to support SLAA capacity to undertake new and increased programmatic evaluative responsibilities.

While federal legislation required each SLAA to conduct a five-year evaluation, IMLS gave each the option to decide whether to include OBE and in what form. The SLAAs increasingly placed emphasis on their own project grants to public libraries. Some adopted a requirement for a simplified logic model and OBE in project grants to libraries, and some built tools and training for that purpose.

This move to OBE required changes in the IMLS Grants to States web-based annual reporting system. The focus remained at a project level, with much of the architecture continuing to monitor expended funds. This reporting system allowed OBE reporting for both the SLAA as a whole and for its project grants. Reporting progress made in building SLAA OBE capacity was optional, with more detailed information requested at the project level. Fields in the online reporting tool were created for projects to enter information on activities, outputs, and outcomes, paralleling logic model architecture.

Substantial effort was devoted to creating a taxonomy to classify projects related to legislated priorities, but no standardized measures at the project level were incorporated for outputs or outcomes. The reluctance to examine program participants’ experiences in library programs reflected a hesitancy to cross a perceived user confidentiality barrier. It also reflected an inability to develop a logical structure to accommodate wide variation in service delivery.

As the effort to use this approach to OBE matured through a little over the first half of the decade, IMLS continued to invest substantial resources to this endeavor. This impacted relations with each SLAA partner, as a number of the states soon designated at least one staff person to manage the new OBE approach within their agency. Further, attempts were made to drill deeper into two specific areas involving substantive investment of Grants to States funds for project grants – early childhood reading and staff development. While analyses were done for guiding assessment and pilots launched in multiple sites around both subject areas, the endeavors were not brought to a level of refinement to support large-scale adoption.

Overall, the Grants to States reporting system evolved without a unified theory of change to cumulate results for the plethora of library programs supported with Grants to States funds across SLAAs and their local library partners. As a result, IMLS and its independent contractors for assessment of the SLAA five-year evaluations submitted in 2002 and 2007 found it extremely difficult to aggregate beyond individual projects to assess the performance of a state or a suite of libraries involved with the same program category (e.g., early childhood reading).

Evaluation became part of the vocabulary for IMLS and grantees, but the terms and evidence derived from its incorporation could not yet be used systematically to frame conceptual understandings or to guide high-level decisions. Instead, the SLAAs’ autonomy, combined with disparate evaluation and program planning experience and knowledge across the Grants to States universe, created a system of complex and vaguely understood elements, fragmented across a multitude of projects. This made monitoring funding expenditures by project category complex and evaluation of results beyond individual projects nearly impossible.

By the latter part of the decade, a change in the appointed IMLS director corresponded with a de-emphasis in program evaluation in Grants to States (and elsewhere in the agency). The third-party evaluation training contract ended with no substitute. The one IMLS evaluation officer was reassigned and no replacement was hired. Remaining IMLS staff had never been never fully integrated for using OBE. Eventually IMLS stopped investing in improving its information reporting systems on the grounds of concerns for efficiency and the promise of government-wide solutions (which have not yet been implemented into this program’s administration at the time of this writing).

Ultimately, as project-level OBE using logic models was integrated into Grants to States program administration, and pertinent new technical and administrative challenges emerged, funds and resources moved away. Little was learned about how Grants to States translated into impacts on the public.

Despite these shortcomings, this era is noteworthy for an unprecedented shift in the program culture with a deliberate linkage of planning with evaluation under the rubric of OBE with logic models. This linkage is important as it had influenced responsibilities within each SLAA as well as between IMLS and SLAAs. The adopted approach to OBE, despite its imperfections, gained a strong foothold in the Grants to States Program.

This legacy would shape the parameters for the *Measuring Success* initiative, the focus of the next section.

**Impetus for the *Measuring Success* Initiative**

By 2010, the political calculus across all levels of government had changed fundamentally. It became and remains politically risky to rely primarily on articulating the belief that fostering strong libraries nurtures a vibrant democracy as an argument for sustaining and expanding funding in this area. It has become imperative to move beyond this narrative by much better ability to demonstrate *how* public investments in library services result in concrete benefits to the public. In fact, while overall demand for library services has continued to increase, aggregate state funding for library services has shrunk (Institute for Museum and Library Services, 2011).

In alignment with the political realities that spur the new urgency to increase the evaluative capabilities in Grants to States, new IMLS leadership created an Office of Planning, Research, and Evaluation (OPRE). In addition to overseeing the agency’s new statistical research program, this office is charged with oversight and reporting on IMLS programmatic performance. To help meet these demands, new research and evaluation staff were hired, including a senior evaluation officer late in 2010 and a second evaluation officer in 2011. In early winter of 2011 IMLS’s new appointed director signaled that evaluation was a chief priority of her administration. The LSTA Grants to States program, as the agency’s largest single program, has been at the forefront of this new policy direction.

A new IMLS evaluation approach to Grants to States has required SLAA buy-in as partners to ground the content while preserving each state’s flexibility to address the unique circumstances in their jurisdiction. It has required a more consistent framework to: (1) track performance of activities and services in the LSTA Grants to States over time and across different programmatic areas; (2) enable a much stronger performance synthesis that allows for purposeful clustering across individual projects; and (3) better identify and foster of best practices and shared learning.

**Designing the Measuring Success Initiative**

Creating a new evaluation approach for this formula grant program has faced several key challenges. For better and worse, evaluation has been closely tied to logic models since the late 1990s, and severely decentralized at a project level. Key parties in IMLS and its SLAA grantees have had disparate understandings of evaluation and interpretations of definitions and application of project measures. Finally, any solution must be embedded in a new grant performance reporting system.

Following months of building internal consensus, a framework for restructuring the evaluation approach for Grants to States was shared with SLAA participants at their annual convening in March of 2011. This new OBE approach moved away from logic models.

Decisions were operationalized in two transparent principles. First, the development of the new evaluation approach would unfold iteratively and incrementally. Three overlapping phases to occupy an approximately 24-month period were delineated: design, pilot, and roll-out. Sequential steps within each phase would provide systematic points for reflection and adaptation as appropriate. Second, the process was designed to be participatory. States would drive the content while IMLS would take an active facilitating role. Additional external stakeholders, such as experts in methodology, would be brought in as circumstances warranted.

As the overriding vision for *Measuring Success* was communicated, attention turned towards the first phase in designing the new evaluation system. Four sequential steps were initially planned: (1) kick-off at the March 2011 meeting, (2) “backward logic mapping,” (3) “forward logic mapping,” and (4) creation of assessment frameworks. At this time, basic features of the new design are nearly complete and are the focus of the remainder of this section.

**Kick-Off**

The March 2011 annual meeting of SLAA staff was comprised predominantly of SLAA directors (chiefs), state grant coordinators, and other senior SLAA professionals. The event was structured to engage the SLAAs as partners in a process with which participants had some familiarity. SLAA participants were broken into four groups that reflected six of eight major priorities spelled out in the new federal legislation summarized in Table 1 below.[[5]](#footnote-5)

***Insert Table 1***

Each team contained about 25 SLAA participants and two IMLS facilitators. In beginning a new OBE approach that differed from logic models, the teams engaged in “backwards logic mapping.” Teams were assigned one or two of the federally authorized priorities for review and identification of key objectives. Participants next articulated and ranked the most important external opportunities outside of the program related to these objectives. The opportunities were vetted using a taxonomy that considered political, economic, sociocultural, and technological influences as distinct areas of external influence on the program (See Bryson, 2004, for details of the PEST taxonomy). In a subsequent activity, participants repeated the process for external programmatic barriers. Finally, participants began to list and prioritize strategies to address the opportunities and barriers they had now identified and ranked.

By day’s end, the meeting rooms were filled with flip-charts sprinkled with post-it notes and colored dots representing rankings of the opportunities, barriers, and strategies. The introduction of a single strategic planning tool, exercised in the same way across all of the SLAA/IMLS teams, moved participant preoccupation away from sensitive issues associated with state autonomy and changes in performance measurement and towards discussion of what the Grants to States program intended to achieve as a whole. It also helped build trust and common ground on the importance of moving forward to develop a new evaluation approach.

**Continued Backwards Logic Mapping**

Communication between IMLS and SLAA partners continued following the conference. *Measuring Success* “branding” occurred with the creation of a Wiki. By late spring, SLAA participants had self-selected into six teams, each roughly corresponding to a priority in the new federal legislation.[[6]](#footnote-6)

Webinars were held biweekly from late May through early July 2011. Teams completed backwards logic maps correlated to the six federal priorities. They identified and recommended strategies to attain objectives associated with their priority, either directly or indirectly through “capturing” an opportunity or barrier. Figure 1.3 illustrates a sample backwards logic map produced during that time.

***Insert Figure 1.3.***

The backward mapping efforts were most successful in identifying specific objectives and associated barriers and opportunities. They were less successful in articulating the logic by which a successful program could reach its objective. Sometimes the program and objective were considered identical. Other strategies focused on related efforts such as advocacy. While participants homed in on key opportunities and barriers that could affect achievement of a selected objective, it often remained unclear how a strategy or bundle of strategies could capture external opportunities or barriers to produce the desired result.

 **Forward Logic Mapping**

By July, many SLAA participants were interacting frequently in various teams, virtually and by telephone. They included chiefs, senior librarians, and mid-level professionals.

In furthering momentum for emerging communities of practice (another underlying goal of the process), the teams began to use “forward logic mapping” to build out the new evaluation scheme. SLAA and IMLS participants referred to these logic maps as “results chains.” Forward logic maps or results chains involved vetting a series of if/then statements. Participants began with a strategy articulated towards the end of the backwards logic map construction. They then postulated a sequence of events that could be expected culminating in a suite of outcomes changing the circumstances of some targeted segment of the public.

These results chains differed from traditional logic models in several major ways. Both terminology and concepts had changed. Participants ceased to think about change as a simple linear flow, connecting inputs, activities, outputs, outcomes, and impacts. Instead, they began with a strategy (and its associated sub-set of supporting activities), and proceeded to consider how the strategy could lead to a subsequent related, desired result. A result could lead to another result or to initiation of an additional strategy, or both. No limit was imposed on the number of strategies/results that could proceed a concluding result. Similarly, there was no limit on how many strategies/results might be needed to reach the next sequential point. This logic mapping accommodated many non-linear interactions and allowed for recursive changes as well.

Social network theory also was introduced to frame and validate emerging theories of change in the results chains (for more on social network theory in evaluation, see, Penual et al., 2006). Four sets of actors were identified: (1) SLAA*s*; (2) public libraries[[7]](#footnote-7); (3) *non-library partners*; and finally, (4) users, comprising distinct segments from the broad public. Figure 1.4 summarizes the underlying social network revealed across all of the results chains.

**Insert figure 1.4.**

As seen in Figure 1.4, the four sets of actors are closely interwoven. First, the SLAAs conduct needs assessments, develop long-term statewide plans, and extend available types of support to catalyze efforts among public libraries. Second, public libraries customize the planning to the circumstances of their communities, which subsequently leads to delivering pertinent services and activities to various segments of the local public. Third, outside partners bring expertise, capital and connections to assist SLAA efforts at a state level as well as public libraries in their local communities. Fourth, users (i.e., segments of the public) access and participate in the library services and activities. As teams refined their chains, the social network was refined further to stratify users into two groups: target users who initially participate in the library services and activities, and potential users who might opt to participate in the same services and activities as target users diffuse information about value they found.

By the end of this phase, each team had produced at least three results chains that corresponded with their perceptions of major objectives tied to their team’s federal priority. The six teams completed about 25 results chains over a six week period.

**Refining Results Chains**

As the initial iteration of the logic maps was finalized, this phase of SLAA teamwork ended. The initiative’s original design had assumed that the logic mapping would be complete enough that SLAA planning participants could: (1) recommend points on the results chains for national assessment; and (2) agree on construction and methods of assessment at these points using descriptive statistics and/or other appropriate qualitative tools. The first goal was complete by mid-August, but the second never occurred. Instead, after reflection, it became clear that further refinement of the logic was needed.

Chain refinement reflected the political reality of federal legislation governing Grants to States. Its programmatic priorities did not articulate specific outcomes expected for various segments of the US public and communities. The initial round of results chains exposed a schism in mapping to federal priorities and to intended changes in participant circumstances. A decision consequently was made to continue IMLS facilitation of the new system design for another several months using a smaller group of SLAA participants who had emerged as leaders. Seventeen mid-level professionals from sixteen SLAAs agreed to participate as technical advisors.

These technical advisors first convened in a webinar in late September 2011. A new configuration was presented for consideration. Table 2 shows how the twenty-five results chains created by the original SLA teams were organized into focal areas that reflected the types of library services that SLAAs typically support through Grants to States.

***Insert Table 2***

The technical advisors agreed on this classification scheme and moved to refine the original results chains better to delineate steps in the social network that lead to desired changes in different segments of the public (i.e., users). Ultimately, consensus emerged that staff and leadership development was best integrated into results chains for the other focal areas and best viewed as in interim outcome.

As the teams further refined the results chains by focal areas, they implicitly agreed to an emerging theory of change for informal learning as summarized in Figure 1.5.

***Insert Figure 1.5.***

As seen in this figure, program staff makes initial assessments of the potential target public in a community or catchment area. These assessments include not only the types of learning that can benefit specific segments of the public, but also external opportunities and barriers that are likely to influence public participation in the informal learning opportunity. Following assessment, efforts are made to directly enable informal learning for individuals and to address indirect opportunities and barriers that may influence the ability and/or desire to access an identified body of knowledge. In inducing individuals to participate in a library’s learning program, attention is focused on capturing external opportunities and barriers through incentives. Finally, assuming satisfactory experience and effective incentives, the theory assumes initial target users will share their experience with others. These other users would adopt the same benefits acquired by the target users through diffusion. The one complexity not illustrated in Figure 1.5 concerns that this programmatic action operates through a social network that interrelates an SLAA, public libraries, other potential partners, and a suite of distinct user groups.

The emerging theory of change moved participants away from both KAB and diffusion theory in making the causal link from acquiring knowledge to its subsequent application. As seen in Figure 1.5, causality is embedded in capturing external opportunities or threats as the program interacts with the users. In practical terms, capturing these opportunities and threats involve applying appropriate incentives such as social media innovations or, say, addressing childcare constraints. This insight moved the program’s theory of change further into social networking by building on insights of social marketing theorists (see Lefebvre, 2010). Put simply, acquisition of knowledge may not suffice for someone to act or change behavior, be it in deciding to participate in some library service/activity or in subsequently applying the learning acquired from that participation to another aspect of the participant’s life (e.g., applying and obtaining a job). Incentives matter. Further, diffusion is seen as an outgrowth of the process: as one individual experiences the benefits of participating in some library service or activity, spill-overs are more likely to occur by either the target user opting to participate in another library program or another group of users opting to participate in the initial library program.

**Creating an Assessment Framework**

By early November 2011, the technical advisors had finished the bulk of their work. The teams had refined results chains for five focal areas (including online databases but excluding civic engagement). All incorporated the same set of interrelated actors in a social network and articulated a unified theory of change associated with informal learning. Unlike logic models, the results chains relied extensively on non-linear relationships culminating in expected improvements for a plethora of groups in society.

As the new focal area results chains were finalized, the technical advisors moved to consider designing an assessment framework to measure whether the theories embedded in the results chain worked as intended, and if not, to be able to adapt and modify. Links on the results chains were reviewed to decide whether national assessment was merited. If the answer was yes, research questions were vetted around these links.

At this point, an unexpected challenge emerged. IMLS had always assumed that evaluation methods would not rely exclusively on statistical performance indicators. It expected that the nature of the question would drive the method, be it statistical metrics and/or other qualitative inquiry approaches. The senior author of this manuscript led methods selection in close collaboration with the director of IMLS’s Office for Planning, Research, and Evaluation. This task had been presumed to be relatively straightforward, given their education and experience, but it turned out otherwise.

If a new evaluation approach to collect data across disparate projects and connect to the interactive social network of the SLAA universe was to be created, an entirely new scheme that could scale across local, state, and national levels was needed. There was no prior precedent for such a system in library services in the US. Besides introducing new political and social uncertainties into the newly emerging evaluation approach, this lack of precedent created a methodological complexity. It was one issue to ask SLAAs to report about matters to which they had direct control, such as funds allotted for a particular program category, say, through a project grant over a given time interval. It was something else to collect data for initiatives over which they have no direct control, and in which no dominant culture for using data in programmatic decision making exists (e.g., changes in employment status for participants in a library-based employment training program). A one-size-fits-all data assessment model was not possible.

As a result, the *Measuring Success* design phase stopped again for reflection in late fall 2011. By this time, the political urgency of the initiative had increased again, as the state library directors (chiefs) were scheduled to meet at one of their periodic meetings. Discussions between IMLS senior leadership and these state library officials at this meeting about progress already made led to renewed support to continue the *Measuring Success* endeavor.

The proposed solution for the assessment scheme was developed by IMLS internally in winter 2012. The schism in the earlier evaluation approach with logic models used in Grants to States between projects and some higher unit of analysis was resolved in the new evaluation approach by targeting the focal areas in the results chains as foundations. Using prior input from SLAA participants, projects were classified into discrete activities and services corresponding to each of the focal areas. This scheme would focus on performance across the entire social network, with discrete assessments of an SLAA, a participating library, and some targeted set of users. Another priority simplified the logic to enable easier aggregation of data across individual projects.

The process for building out the solution continued to rely on incremental change to strengthen capacity and buy-in of the SLAAs. Initially, reporting would not look much different from the present, but information would be simplified and streamlined. Annual reporting would retain the essential need to monitor expenditures, particularly at project levels, but it would improve the capacity to report outcomes, at least to the extent they could be attributed to IMLS funding to SLAAs and their local libraries. This reform also would begin to allow for better aggregation of information across individual projects.

A gradual build out of the whole assessment framework for user-specific impact would happen by piloting through volunteer SLAAs. Building this pilot capacity would foster participation in new communities of practice among participating states and their network. As their capacity increased, diffusion of new practices to other states would be more possible. In addition, data would be collected and analyzed across clusters of projects that corresponded with various activities under each of the focal areas.

The overall strategy introduced an additional new wrinkle. Instead of being confined to an annual SLAA web-based reporting system linked to mandates for five-year plans and evaluations for data collection and reporting, IMLS would exercise leadership and resources to frame multi-site evaluations of selected activities and services within focal areas (e.g., evaluation of a suite of childhood reading projects across a suite of local libraries under the lifelong learning focal area).

**Next Steps**

At the time of writing, IMLS is vetting the design of the program’s new evaluation system with its SLAA partners. Unlike in 2011, the 2012 conversations are more grounded, encompassing the discussions between IMLS and its SLAA partners over the past year around the completion of the design phase of the *Measuring Success* initiative. The planned design contains a more logical structure for detailing the plethora of initiatives for which SLAA partners use Grants to States funds. The system is built on the non-linear unifying theories of change embedded in the results chains developed by IMLS and the SLAAs. It contains a more logical approach to measurement and assessment and enables individual projects to be systematically sorted into coherent clusters for higher-level assessments of particular library initiatives for each SLAA and the nation at large.

The plan also details next steps in the *Measuring Success* initiative, including a roll-out of a new performance reporting system and SLAA partners that have volunteered to pilot these innovations. The high level of collaboration using virtual and other technologies continues. Emphasis is placed on creating and nurturing new communities of practice, diffusion, reflection, and adaptation for continuous learning.

**Discussion**

The authors expected that the flaws in logic models would manifest as a major lesson for participants in this initiative. This is believed to be true. Logic models have demonstrated limits. The approach adopted in *Measuring Success,* using backwards and forwards logic maps, has provided an effective alternative to IMLS’s OBE architecture in allowing participants to better articulate the major types of library services and programs that libraries actually deliver using Grants to States funds. It allows for non-linear relationships. Participants have embraced dynamic environmental interactions, leading to an emerging theory of change that embeds a social network of interdependent actors.

The approach using the backward and forward logic maps also had another intentional benefit. Logic models presented a simple and attractive way of integrating planning with assessment when OBE was initially introduced to Grants to States in 1997. The logic maps introduced in the *Measuring Success* initiative significantly increased the strength and quality of the integration. The logic maps focused explicitly on issues surrounding planning in better specifying precisely what the various activities captured in Grants to States’ various focal areas intended to achieve in benefitting specific segments of the American public. Stakeholder acceptance of the new approach to outcome based evaluation (or, more broadly, results based management) resulted in large because of this greater capacity of logic mapping. Further, the increased quality of articulation of program intentions using the forward logic maps enabled the ability to develop a more sophisticated and clearer logic for assessment and measurement.

Nonetheless, logic models were only one challenge in changing an entire culture formed around prior program evaluation experiences. In particular, IMLS’s evaluation approach in Grants to States decentralized data collection and analysis in individual project grants without systematic links to support broadly useful collection and analysis of data for state and federal decision making. Whether fortuitous or otherwise, the forward and backward logic maps in the *Measuring Success* initiative enabled development of solutions that allow scaling of individual projects into clusters of similar activities using the derived focal areas as a foundation. This foundation has led to construction of a logical reporting structure that should greatly increase the effectiveness of overall monitoring of expended funds and performance for individual projects when clustered into their hierarchical groupings. We anticipate it will be both more effective and more persuasive in showing the value of libraries to the public.

The capacity of the new evaluation approach to function more effectively across multiple levels was directly related to the introduction of systems theory thinking. A structured social network linking a suite of actors is critical for evaluating the effectiveness of a wide array of library services supported through Grants to States. The same is true for the introduction of environmental opportunities and threats as action points in program planning and monitoring. These precepts are essential for understanding the realities of any community-based program, with those of a local library a core example. They are expected to yield huge future dividends when stakeholders begin to consider applying evidence emerging from this new evaluation approach. There is huge potential for generalizing this insight to a slew of other library service endeavors grounded in rich community relationships.

A major factor that will determine the effectiveness of this new evaluation approach will be its usefulness to various stakeholders (Patton, 2008). The process used in launching *Measuring Success* has increased the likelihood of such success. SLAA participants can see their fingerprints on the entire scheme as they drove the content, with active IMLS leadership and fulfillment. With greater ownership of the product, the SLAA partners and IMLS program staff are more likely to value it, apply it, and find it more useful.

Technical, administrative, and political uncertainties remain. Prior habits that have shaped understanding about what evaluation entails, particularly the emphasis on tracking expended funds and assessing outcomes for individual and widely disparate projects remains a core mindset among many SLAA and IMLS staff despite the progress made over the past year with the launching of the *Measuring Success* Initiative.

Change takes consistency, iteration, and time. The point of this case study is that change is historical. The changes made thus far in introducing a new OBE approach linking program planning with assessment would not have been possible if there had not been experience with a prior one using logic models. The prior approach reflected dominant professional thinking at the time of introduction 15 years ago. Correspondingly, even if progress continues in instituting a new evaluation approach to the Grants to States Program, there undoubtedly will be unforeseen technical, administrative, and political uncertainties that will arise, requiring further reflection and adaptation. These uncertainties will undoubtedly cause a future generation of program evaluators and library service professionals to decide to develop a new evaluation approach in response to changed circumstances.

Despite this note of caution, we emphasize that the process used in Measuring Success has enabled a concrete and compelling vision to emerge with a promise of better helping library executives and administrators better address external policy-maker concerns for accountability and results. This will help these key officials better manage library programs in a continuing era of heightened public demand and scarce budgets.

**Acknowledgement**

The authors appreciate the reviews of various drafts of the paper performed by their IMLS colleagues -- Carlos Manjarrez, Mary Chute, Laurie Brooks, James Lonergan, Michele Farrell and Timothy Owens for reviewing various drafts of the paper. They especially thank Karen Motylewski for her high level of professionalism in continually engaging with us, leading to a substantially improved argument through her editing multiple versions of this paper.

**Authors’ Note**

The views expressed are those of the authors and do not necessarily reflect the official position or policies of the Institute of Museum and Library Services.

**References**

Boston, J. M., Pallot, J. and Walsh, J. and Walsh, P (1996). *Public Management: The New Zealand Model*. Oxford, UK: Oxford University Press.

Bryson, J.B. 2004. *Strategic Planning for Public and Nonprofit Organizations: A Guide to Creating and Sustaining Organizations.* 3rd Edition. San Francisco, CA: Jossey Bass.

Farrior, M. (2005). Breakthrough *Strategies for Engaging the Public: Emerging Trends in Communications and Social Science*. Retrieved February 1, 2012 from <http://www.biodiverse.org/docs/publicationsandtipsheets/breakthroughstrategiesforengagingthepublic.pdf>

Gladwell, M. (2000). The *Tipping Point: How Little Things Can Make a Big Difference.* New York, NY: Little Brown and Company.

Harris, E. (2005). An Introduction to Theory of Change.  *Evaluation Exchange*. XI (2). Retrieved February 1, 2012 from <http://www.hfrp.org/evaluation/the-evaluation-exchange/issue-archive/evaluation-methodology/an-introduction-to-theory-of-change>.

Heath, C. and Heath, D. (2007). *Made to Stick: Why Some Ideas Survive and Others Stick*. New York, NY: Random House.

Hernandez, M. (2000). Using Logic Models and Program Model Theory to Build Outcome Accountability. *Education and Treatment of Children.*  23 (1), 23-41.

Institute for Museum and Library Services. (2011). Public *Libraries Survey: Fiscal Year 2009*. Retrieved February 5, 2012 at <http://www.imls.gov/assets/1/News/PLS2009.pdf>

Levebvre, R. C. (2010). *On Social Marketing and Social Change: Selected Readings 2005-2009*. United States: Createspace.

McDavid, J.C. and Huse, I. (2012, March). Legislator Uses of Public Performance Reports: Findings from a Five Year Study. *American Journal of Evaluation*. 33:1, 7-25.

Patton, M. Q. (2011). Developmental *Evaluation: Applying Complexity Concepts to Enhance Evaluation Innovation and Use*. New York, NY: Guilford Press.

Patton, M. Q. (2008). *Utilization-Focused Evaluation*. 4th Edition. Thousand Oaks, CA: Sage Publications.

Penual, W. R., Sussex, W. and Korbak, C. (2006). Investigating the Potential of Using Social Network Analysis in Educational Evaluation. *American Journal of Evaluation*. 27 (4), 437-451.

Rogers, E., I. (1962). Diffusion *of Innovations*. New York, NY: Free Press.

Rogers, P. J. (2008). Using Program Theory for Complicated and Complex Programme Evaluation. *The International Journal of Theory, Research and Practice*. 14 (1), 29-48.

Rogers, P. J. (2010, November). *Representing Simple, Complicated and Complex Aspects in Logics Models for Evaluation Quality*. Paper presented at the meeting of the American Evaluation Association. San Antonio, TX.

Rossi, P., Lipsey, M.W. and Freeman, H.E. (2004). *Evaluation: A Systematic Approach*, 7th Edition. Thousand Oaks, CA. Sage Publications.

Russ-Eft, D. and Preskill, H. (2009). *Evaluation in Organizations: A Systematic Approach to Enhancing Learning, Performance, and Change*. Second Edition. Philadelphia, PA: Basic Books.

Schelling, T.C. (1971). “Dynamic Models of Segregation.” *Journal of Mathematical Sociology*.

Schmitt, E. (2007, December 21). “The Theory of Change Primary.” *The American Prospect*. Retrieved February 1, 2012 from <http://prospect.org/article/tehory-change-primary>.

Weiss, C.H. (1995, Winter). “The Four I’s Of School Reform: How Interests, Ideology, Information and Institution Affect Teachers and Principals.” *Harvard Education Review*. 65:4, 571-593.

Weiss, C.H. 1972. *Evaluating Action Programs: Readings in Social Action and Education*. Boston, MA: Allyn and Bacon.

1. A theory of change is a way to make assumptions about how some factors lead to some type of social (or other) change which is subject to assessment for evaluating its relative effectiveness. See, for instance, Schmitt, 2007 and Harris (2005). [↑](#footnote-ref-1)
2. Despite a wealth of literature in the public health and social psychology fields in the 1980’s, not much empirical testing of the model has occurred. Many scholars have either qualified the model to better specify types of knowledge, attitudes, and behaviors involved or modified links. For a nice synopsis of the model and its shortfalls, see Farrior (2005). [↑](#footnote-ref-2)
3. By societal change, we are simplifying multiple units of change that extend beyond individuals, such as distinct groups, communities, and so forth. [↑](#footnote-ref-3)
4. Those with experience in organizational politics may have seen this referenced as the “80:20” rule. This means that change may never reach 20% of a target population. [↑](#footnote-ref-4)
5. The IMLS team working on *Measuring Success* agreed that the fourth priority, partnerships, would be integrated in other priorities. It also was decided to omit the last priority, since it allows the SLAAs to meet other non-specified state library needs with Grants to States funds. [↑](#footnote-ref-5)
6. Many team members participating in the Webinar process had not participated in the March 2011 conference. [↑](#footnote-ref-6)
7. Many states give project grants to different types of libraries. For the purposes of this paper, however, the discussion is confined to public libraries, since it is this type of library that has been the predominant focus of discussion during the Measuring Success initiative. [↑](#footnote-ref-7)