Wilder Research

Program evaluation capacity building at the Minnesota Historical Society

Summary of the project and recommended next steps

In October 2010, the Minnesota Historical Society (the Society) contracted with Wilder Research to assist with institutional evaluation capacity building. The purpose of this project was to develop comprehensive program evaluation capacity and to strengthen evaluation activities within the Society. This report summarizes the activities that were completed as a part of this project and provides recommendations for how the Society should move forward to continue to strengthen program evaluation in the organization.

Approach

The first phase of this project was an organizational assessment. The goals of the assessment were to better understand current evaluation practices at the Society, existing evaluation skills and capacities among staff, gaps in evaluation skills or resources, staff attitudes toward evaluation, and barriers to conducting high quality, systematic evaluation of Society programs and the organization as a whole.

Wilder conducted a web survey with 111 Society staff members and interviews with a subset of 38 staff members in December 2010. These staff were identified by the Society's Evaluation Team as individuals who have some role in program evaluation.

Following the assessment, Wilder Research worked with the Society on other activities designed to build organizational evaluation capacity. These tasks included developing an organizational logic model and institutional dashboard of success indicators. Wilder also developed an Evaluation Handbook that provides recommendations for which programs should be evaluated, the methods that should be used for conducting evaluations, and how leadership should use evaluation results for decision-making.

Logic model, dashboard, and Evaluation Handbook

Wilder Research worked with the Society to develop a logic model and a corresponding dashboard of success indicators. These were developed based on the results of the assessment, meetings with the Evaluation Team, a review of the literature and logic models and dashboards from similar organizations, and a review of the work completed by previous consultants to the Society. They were refined based on feedback from staff via several in-person meetings and formal polling that was conducted at a Leadership Council meeting in June 2011.

The logic model is an overview of the Society's key activities and intended outcomes. It provides guidance regarding the most important things to measure in program evaluation. It ensures that all programs at the Society are working toward and tracking progress on the same ultimate goals. The organizational logic model also serves as a template for program-specific or project-specific logic models at the Society. This will ensure that program-level findings can be rolled up to the institutional level.

Wilder also developed a dashboard of key success indicators based on the organizational logic model. The dashboard is a tool to communicate with stakeholders about the impact of the Society. We recommend that this dashboard be updated annually.

The dashboard results can and should be used internally with staff and leadership, and can also be communicated to external stakeholders via print and online sources. We recommend using any and all of these options, once data gathering and reporting has had some time to roll out.

In the assessment process, staff reported a need for existing evaluation tools and resources so they have consistency for every program evaluation. To that end, Wilder developed standard measures for surveys and other evaluation tools that can be used to track progress on key indicators.

The consistent use of one good survey question across programs to measure one key outcome for the organization will introduce efficiencies and will allow the Society to roll up program-level evaluation results to the organizational level. This is useful for communicating the impact of the organization as a whole and for comparing "apples-to-apples" when making internal decisions about programming.

However, fewer than half of the Leadership Council members feel that it will be "very easy" (21%) or "somewhat easy" (21%) for them to incorporate one or more of the recommended dashboard indicators into their program evaluation(s). One-third (32%) think it will be "somewhat hard" and 15 percent think it will be "very hard."

In addition, fewer than one-third of the Leadership Council members feel that it will be "very easy" (6%) or "somewhat easy" (24%) for their work unit to comply with the evaluation requirements described in the Handbook. Half of Leadership Council members (50%) think it will be "somewhat hard" and 12 percent think it will be "very hard."

More work is needed to identify and ameliorate the specific challenges to successfully implementing the dashboard indicators and program evaluation requirements described in the Handbook.

Staff trainings

In addition to developing the documents described above, Wilder Research trained Society staff members on various evaluation skills. A total of 70 Society staff members participated in one or more of the training sessions provided by Wilder from March through June 2011. The webinars have been made available to the Society for ongoing use as needed. Sessions 1, 2, and 3 were part of the series that Wilder recommended as "Evaluation 101" for any staff who are responsible for completing program evaluation at the Society. The other sessions provided training on more specialized techniques that were only relevant for a subset of Society staff.

Session #1: Designing a logic model. This session on the purposes of logic modeling and how to draft a logic model was attended by 46 people (35 in-person, 10 web).

Session #2: Designing and implementing an evaluation plan. In this session, participants learned about designing and using evaluation tools to gather valid evaluation data and how to prepare an evaluation plan. It was attended by 40 people (32 in-person, 8 web).

Session #3: Analyzing and reporting evaluation results. During this session, participants learned guidelines for organizing and analyzing quantitative and qualitative data from program-level evaluations, as well as tips to avoid common analysis and interpretation mistakes, and strategies for presenting evaluation results to a variety of audiences. It was attended by 23 people (20 in-person, 3 web).

Evaluating partnerships. This session, with 22 participants, explored different types of partnerships and how to evaluate the effectiveness of the partnerships.

Evaluating web media. Wilder partnered with the Society's IT Department for this session on using website data for formative, summative, and outcomes evaluation. The 17 participants learned about the specific tools the Society uses to evaluate its web presence, the types of web analytics data available to program staff, and tips for interpreting and using website data for evaluation.

Introduction to return-on-investment. This session provided the 15 participants a broad overview of how concepts from the field of economics can be applied to evaluation, and helped participants identify programs or areas of the Society that may be suitable for this type of evaluation.

Using evaluation results for fundraising. This session, with 12 participants, focused on how to use data in grant proposals, report evaluation back to funders, and communicate results to other audiences.

Office hours. Wilder consultants were also available to answer specific evaluation questions during three "office hour" sessions at the Minnesota History Center. Wilder Research provided 10 staff members with individual technical assistance on a variety of evaluation issues such as designing surveys, developing logic models, calculating response rates, analyzing and reporting focus group data, and completing semi-structured interviews.

Training Satisfaction Survey

Of the 70 attendees participating in at least one of Wilder's evaluation workshops, 44 (63%) completed an online evaluation survey in June 2011. The Society should consider this feedback when planning future evaluation capacity building efforts and when implementing the Society's new evaluation program.

Overall, the first two sessions in the Evaluation 101 series on logic modeling and evaluation plans received the most positive feedback from staff with 9 out of 10 learning at least some in each session, three out of four finding the information at least somewhat useful for their job at MHS, and 9 out of 10 being satisfied with those sessions.

The Communicating Evaluation Results to Funders was rated lowest, with just one-third of participants reporting learning some or a great deal, only one-quarter finding it useful to their current job, and 6 out of 10 reporting being satisfied with the training. This may be due to the fact that the training had to be rescheduled and many of the attendees were from the development office and thus were already experienced in grant writing and communicating with funders.

In general, attendees said the trainings would have been more useful had the Society's institutional guidelines and expectations been clearer, and an institutional leader (either from the board or management) had championed and appreciated the staff's capacity building efforts. Without these expectations clarified, some felt they learned about evaluation conceptually, but aren't clear on how it will apply to their work. Wilder fielded many questions from staff during trainings about what they were required to do for evaluation. At that time we were not able to answer many of their questions because the Handbook and dashboard have not yet been finalized and approved.

Attendees also expressed a need to break down silos and learn from the evaluation work others are doing at the Society. One person suggested pairing those staff currently doing quality evaluations with others who are just learning evaluation skills.

Though Wilder worked to incorporate examples that were appropriate for the Society, given the breadth of what the Society does, some staff felt the examples did not well represent their work, or wanted more one-on-one assistance. Despite this interest in getting more help with their specific programs, few people attended Wilder's office hours sessions or contacted Wilder about getting additional assistance. Others noted they were explicitly asked to attend, but the training did not seem to fit with their work. This was particularly noted by staff in internal operations (e.g., human resources, marketing), which is not surprising given the trainings were geared toward staff serving external users.

Role of the Evaluation Coordinator

In January 2011, Wilder Research recommended that the Minnesota Historical Society hire an in-house Evaluation Coordinator. Steps are currently being taken to create and fill this position in summer 2011. We have a few suggestions regarding the key focus areas for this position.

First, the Evaluation Coordinator should work with the Legacy Manager, the communications team, the Evaluation Team, and leadership to roll out the Evaluation Handbook, logic model, and dashboard. It may be preferable to wait to finalize these documents until the Evaluation Coordinator is in place so this person can have some input into the content and design. In particular, the Evaluation Coordinator should work with the Evaluation team to further clarify some of the items in the dashboard.

Second, the Evaluation Coordinator should work with the Legacy Manager to identify which programs will be required to implement a standard evaluation plan and measure dashboard indicators. The Evaluation team as well as division and department directors should be involved in this process to ensure that limited evaluation resources are prioritized across the organization. The Evaluation Coordinator should then assist program staff to conduct their program evaluation according to the steps described in the Handbook.

Third, the Evaluation Coordinator will be responsible for compiling data and reporting on the dashboard indicators. We recommend that the Evaluation Coordinator prepare a first draft (for internal eyes only) of the dashboard in January 2012 based on evaluations that are completed in the first half of Fiscal Year 2012. The Evaluation Team and the Legacy Manager should then work with the Evaluation Coordinator to refine the reporting of the dashboard indicators. The communications and web teams may also need to be involved, depending on what formats the Society decides to publish the results.

Finally, we recommend that the Evaluation Coordinator and the Legacy Manager be responsible for ensuring that the Executive Council and Society leadership receive regular updates (in their preferred format) on program-level and organization-level evaluation results.

Clarifying expectations

One of the most commonly mentioned concerns about evaluation among Society staff is that expectations from leadership are not clear. In June 2011, Wilder Research met with your Leadership Council and asked them: "How clear do you feel expectations from leadership are regarding how you should be doing program evaluation?" None of the 32 staff who answered this question feel that expectations are "very clear," whereas 38 percent feel that expectations are "somewhat clear," 50 percent feel they are "not too clear," and 13 percent feel they are "very unclear and/or conflicting."

Wilder Recommends that a clear and concise message should be conveyed to all staff who are responsible for program evaluation activities with specific guidance about which staff and/or programs are required to adhere to the standards that are set out in the Evaluation Handbook. This message should also include information about when and how program evaluation results will be used by leadership. The Society may choose to modify the requirements before rolling out to all staff.

Further, we believe that the roll out of these evaluation documents should be paired with in-person conversations with program managers to ensure that they understand the expectations and to address any barriers these staff may have in implementing these requirements. We suggest that the Evaluation Team along with the Deputy Directors and Director should lead this process.

The methods and standards described in the Evaluation Handbook are not intended to create a rigid, one-size-fits-all approach to evaluation. Rather, the Handbook provides a framework, but the staff who are responsible for program evaluation at the Society, and their supervisors, should use their best judgment about what should be evaluated and when.

The most effective way to make these decisions is to identify at the outset how every piece of information gathered in the evaluation is going to be used – if the program evaluation lead can't think of a use for something up front, then it should not be included in the evaluation! Also, the evaluation lead should consider all stakeholder groups and their needs for the evaluation to ensure that the information gathered will get used.

Using evaluation for organizational decision-making

Wilder recommends that the Society complete the following activities to ensure that evaluation results are used effectively for decision-making:

- Present program evaluations and dashboard results to Executive Council (Evaluation Coordinator and the program's evaluation lead)
- Hold quarterly meetings with the Evaluation Team to discuss current program evaluation issues, review and improve the content and process for program evaluations, and review and improve the institutional dashboard to align with the organization's current strategic mission and vision every 3-5 years (Evaluation Coordinator)

- Recommend Legacy-funded projects for continued funding based on outcomes demonstrated in evaluations (Legacy Manager and Evaluation Coordinator)

All staff who conduct program evaluation, and their managers, should use their limited evaluation resources efficiently by only measuring things that will be used by program managers and leadership to make decisions about the program.

Longer-term recommendations

Organizational culture impacts the Society's ability to effectively establish, support, and sustain high quality, systematic evaluation. Therefore, we have identified several ways the Society can begin to make a culture shift. Many of the Society's staff are currently working on evaluation and other research using sophisticated, rigorous methods.

As the Society makes evaluation more systematic across the organization, leadership and staff will start *thinking* and *projecting an image to the external world* that "The Society has clear priorities, is systematically measuring its performance, and can demonstrate that it achieves the intended results in a cost-effective way" (from the Minnesota Historical Society Five-Year Vision Statement). The following are a few suggestions of ways the Society might begin to do this.

First, the Society should consider ways of strengthening connections with others in the field. Also, presenting at conferences, and particularly presenting the results of your research and evaluation efforts, could enhance the perception (both internally and externally) of the Society as an organization that is good at evaluation. The Society could also post resources such as tip sheets on its website for other organizations that are interested in evaluation.

Second, the Society should consider ways of gradually incorporating evaluation into staff performance reviews at all levels of the organization. Leadership should be held accountable for how evaluation is used in organizational decision-making.

Third, the Executive Council, Director, and Deputy Directors should also consider ways of better understanding, supporting, and using program evaluation. One way to initiate this could be to have quarterly presentations at board meetings about recently completed program evaluations, and an annual presentation on the dashboard. The Society should also consider other ways of breaking down silos among departments – one way to do this might be to provide presentations to staff about specific programs and evaluation results at staff meetings and/or Leadership Council meetings.

Fourth, the Society should consider ways of strengthening your infrastructure for evaluation. This might include, for example, investing in software programs that allow for more systematic gathering of user information, including the ability to track users across programs, sites, and time.

In addition, we recommend specific tracking of program users to align with the Society's target populations. However, the Society's ticketing system is not currently set up to gather data that will align with the categories of interest. This is also true for web visitors. It may require infrastructure changes as well as changing the way you have "always done things" to align administrative systems to gather output and outcome data that is relevant to demonstrate the Society's impact.

Wilder Research
Information. Insight. Impact.

451 Lexington Parkway North
Saint Paul, Minnesota 55104
651-280-2700
www.wilderresearch.org

AMHERST H. WILDER FOUNDATION
ESTABLISHED 1906
Here for good.

For more information

This summary presents highlights of the evaluation capacity building project conducted by Wilder Research for the Minnesota Historical Society. For more information, contact Nicole MartinRogers at Wilder Research, 651-280-2682.

Authors: Nicole MartinRogers, Jennifer Valorose, and Darcie Thomsen
JUNE 2011