Minnesota Historical Society Program Evaluation Handbook

For use by Society staff in conducting project- or program-level evaluation.

Compiled by Wilder Research, September, 2011

Minnesota Historical Society

To the staff of the Minnesota Historical Society,

Program evaluation can seem daunting, and maybe even a waste of precious time, energy, and other resources. And yet evaluation questions continually surface across programs, services and projects: How are things going? Are we making a difference? Are the participants benefiting? How can we demonstrate that this activity should be re-funded?

Program evaluation helps answer these questions. The data collected can be used to improve program services, document best practices, highlight program outcomes, and inform funders. It tells us what works and what does not work. It can increase a program's capacity to conduct a critical self-assessment and plan for the future in alignment with the Society's strategic plan. For all these reasons, the Society is committed to program evaluation as a valuable tool for strengthening the quality of our programs and to measure and improve outcomes for those we serve.

This handbook will move our work in this area forward by making available information about evaluation design, implementation, and analysis. It will be available to all staff as a reference tool. Please use it to enhance your evaluation skills and to improve your programs and services. Additionally, this handbook is a living document. It grew out of our own work on evaluation and will be updated as we deepen our understanding.

Some staff may be concerned that evaluation will lead to criticism of their programs or services that is not based on a full understanding of their goals and objectives or doesn't capture progress. As you develop your evaluation skills, keep in mind that the primary purpose of evaluation is to assist you in gathering the information you need to strengthen your programs and services, identify best practices, and communicate what the program or service is accomplishing. With this information, you become an even better advocate for your programs and services and more able to articulate their value.

In closing, we want to know what we do well now so that we may do it better in the future.

Pat Gaarder
Deputy Director, Programs

Table of Contents

Using Evaluation Results for Organizational Decision-Making

Minnesota Historical Society Philosophy and Practices

THE MINNESOTA HISTORICAL SOCIETY IS COMMITTED to using evaluation results for organizational decision-making. All evaluations conducted by the Society should be designed with use in mind. Uses of evaluation might include, for example, developing a program theory, refining the program implementation, clarifying program goals, or identifying user needs.

The Board of Directors, Director and Deputy Director, department directors and other leaders, program managers, and program staff must all have some input into the design of the evaluation and/or that the evaluation should be designed with these users in mind. In addition, all of these evaluation user groups must understand the criteria for program success. "Intended users are more likely to use evaluations if they understand and feel ownership of the evaluation process and findings" (Patton, 2002, "Utilization-focused evaluation," in Evaluation Models, Springer).

The evaluation methods described in the Minnesota Historical Society Institutional Evaluation Handbook are intended to provide useful tips and guidelines for conducting evaluations of all types. However, the methods program staff select to evaluate their program should be suited to the specific intended uses of the evaluation. Evaluation staff (the Evaluation Coordinator and program evaluation lead, et al.) should use an "active-reactive-adaptive" approach to designing the evaluation.

- Active is deliberately identifying intended users and focusing useful questions
- Reactive is listening to intended users and responding to what they know and believe about the particular situation in which the evaluation happens
- Adaptive is altering evaluation questions and designs in light of increased understanding of the situation and changing conditions

Active-reactive-adaptive evaluators are genuinely immersed in the challenges of each new setting and authentically responsive to the intended users of each new evaluation. This approach should be used in all phases of evaluation. **Evaluation results should be "just in time" and "just enough"!**

The Minnesota Historical Society has staff and leaders who are committed to doing good program evaluation. According to Michael Patton and other central figures in the field of program evaluation, the presence of someone who cares about the evaluation is a key factor in whether the evaluation results are used as intended. "Nothing makes a larger difference in the use of evaluations than the *personal factor*—the interest of officials in learning from the evaluation and the desire of the evaluator to get attention for what he [sic] knows" (Chronbach & Associates, 1980, quoted in Patton, M. 2002.)

THE PROCESS DESCRIBED BELOW WILL HELP staff at all levels of the organization use evaluation results to make decisions that will strengthen the impact of the Minnesota Historical Society. The Minnesota Historical Society will use the following processes to ensure that evaluations are useful and used:

- Presentations to Executive Council of program evaluation process and results (for selected programs quarterly, based on interest) as well as dashboard results (annually)

- On-staff Evaluation Coordinator to provide technical assistance to program staff to complete high quality and appropriate evaluations; to ensure organizational evaluation standards are understood, accepted, and met; and to continually improve the process of program evaluation at the Society to meet the needs of all users

- Ongoing Evaluation Team (made up of existing and/or new members that are leaders for across the organization) that meets quarterly to discuss current program evaluation issues (identified by Evaluation Coordinator and/or program staff and managers), to annually review and improve the content and process for program evaluations, and to review and improve the institutional dashboard to align with the organization's current strategic mission and vision every 3-5 years (the Evaluation Coordinator will implement the decisions of the Evaluation Team)

- Incorporate program evaluation into performance evaluations for staff at all levels of the organization, which will hold staff accountable for a) doing the evaluations needed to effectively and efficiently evaluation programs in their area of responsibility, and b) using evaluation results for program improvement and decision-making (at the level appropriate for their position)

- Legacy Manager and Evaluation Coordinator will make specific recommendations to leadership on projects that should receive continued funding based on the outcomes demonstrated in evaluations

INTRODUCTION

THE MINNESOTA HISTORICAL SOCIETY PROGRAM EVALUATION HANDBOOK is a resource for all staff, with best practices in program evaluation adapted for the Society. The handbook includes a brief introduction to the evaluation field; general information for evaluation planning; implementation, and utilization, evaluation standards, templates and tools specific to the Minnesota Historical Society; and web resources.

Following the standardized, systematic program evaluation guidelines in this handbook throughout the organization will ensure the Society is using its limited resources efficiently and effectively toward achieving our mission of illuminating the past to light the future.

Evaluation will also help us better communicate the impact of individual programs and the Minnesota Historical Society as a whole to all of our stakeholders, including users/participants and target audiences, staff, volunteers, partners, donors, current and potential funders, the Minnesota Department of Education, the Minnesota Legislature, and all Minnesotans.

Note on Standards and Best Practices:

The American Association of Museums (AAM) and American Library Association (ALA) bothprovide standards, guidelines, and best practices that guide museum and library operations. The organizations define these terms in the following ways:

"Standards are generally accepted levels that all museums are expected to achieve." (AAM)

"Standards [are] policies which describe shared values and principles of performance for a library." (ALA)

"Best practices are commendable actions and philosophies that demonstrate an awareness of standards, solve problems and can be replicated. Museums may choose to emulate them if appropriate to their circumstances." (AAM)

"Guidelines consist of procedures that will prove useful in meeting the standards." (ALA)

While current AAM and ALA standards, guidelines, and best practices do not include recommendations for conducting program evaluation, the Society is committed to doing evaluation which supports the AAM and ALA standards, best practices, and guidelines for Minnesota Historical Society operations.

Defining Evaluation

What Is Evaluation?

- Evaluation is a systematic process—it is a planned and purposeful activity, not an afterthought.
- Evaluation involves collecting information related to questions or issues about organizations, programs, processes, society, policy, etc.
- Evaluation is a process that improves knowledge and decision-making about process improvements or whether to continue or expand programs.
- Evaluation asks questions about issues that come from everyday practices to help us better understand what we do, the impact of our actions, and how we should move forward with what we learn.

How Is Evaluation Different From Research?

The terms "evaluation" and "research" are sometimes used interchangeably. Both forms of inquiry may use similar data collection and analysis methods, but may differ in the following ways.

	EVALUATION	RESEARCH
PURPOSE	Produce useful information for program improvement and measure impact	Gain general scientific knowledge about a topic or program
QUESTION TYPE	Specific to the program	General
AUDIENCE	Program stakeholders	Academic or public

Why Conduct Evaluation?

Evaluation can help the Minnesota Historical Society:

- Make informed decisions at the program and organizational levels.
- Improve the public's experience with the Society.
- Manage projects by focusing a program at the beginning, keeping it on track, and making decisions throughout.
- Increase access to funding and fulfill funder requirements.
- Save time and money.
- Improve communication with stakeholders about the Society's impact.
- Market programs and services.
- Contribute to the field by documenting successes, avoiding failures, and discovering answers to critical questions.

A Brief History of Evaluation

How Did the Field of Evaluation Develop?

Modern evaluation in the United States can be traced back to the 1950s and 1960s during President Lyndon Johnson's War on Poverty and Great Society initiatives. Prior to this period, evaluation was conducted by social scientists at a small number of universities and organizations, focusing mainly on education assessment.

Resources began pouring into Great Society programs, but the complex problems the programs attempted to address did not disappear. Pressure to provide evidence of effectiveness increased and resulted in federal evaluation requirements. In later decades, federal funds became less widely and generously applied, with "targeted investment" becoming more popular. Evaluation was often used in order to aid how and where to invest.

The evaluation field also progressed during this period, growing in theory, methodology, and practice. Professional associations began to emerge in the 1970s and 1980s. Many public and private organizations also began developing internal evaluation units during this time.

Meant to illuminate the causes of social problems as well as the clear and specific means with which to fix such problems, early expectations of evaluation still hold fast today. Evaluation requirements attached to public funding continues to focus more on accountability ("Prove it works!") than quality ("Improve this work!").

The field of evaluation, however, has broadened to include participatory, collaborative, and learning-oriented evaluation, recognizing that few evaluations are value-free and many are politically-charged. National and international associations continue to develop. The tension between the development of evaluation as a field and the pressure to demonstrate the effectiveness of social programs continues pushing and pulling as the discipline grows and changes.

How Did the Field of Museum Evaluation Develop?

Although evaluation in museums was conducted as early as 1932, the museum evaluation field is generally considered an emerging discipline. Professional associations such as the Visitor Studies Association, founded in 1988, focus on better understanding visitors, and how to attract, educate and serve them in lifelong learning through informal learning environments, inclusive of museums, zoos, nature centers, visitor centers, historic sites, parks, and other informal learning settings.

Museum-specific theoretical approaches are rooted in the social sciences, with a strong emphasis in behavioral psychology. Today, approaches vary widely, influenced by anthropological, sociological, and other disciplinary practices. As the museum evaluation field continues developing, it experiences many of the same pushes and pulls as the evaluation field at large, related to broadening theory and methodology and documenting accountability and quality.

Evaluation Criteria and Ethics

All evaluations should incorporate the following concepts throughout all phases of the project, including planning, implementation, and use of evaluation results. Refer back to this section, as needed, throughout your evaluation to ensure the process is reflective and thoughtful.

When doing evaluation, remember that there is more than one way to do evaluation and that all evaluation work is value-laden. This is key to avoiding assumptions and achieving better balance in evaluation work.

Evaluations should meet the following criteria:

CRITERIA	DEFINITION
Systematic	Evaluation and data collection are conducted in a systematic, consistent way such that all comparable data are treated similarly.
Respectful	Evaluation work should respect the security, dignity, and self-worth of the respondents, program participants, clients, and other stakeholders with whom they interact.
Credible	Findings/conclusions should be believable to intended audience. If stakeholders perceive evaluation processes and data as legitimate, and recommendations as reasonable, findings/conclusions are more likely to be accepted.
Feasible	The evaluation should be able to be reasonably accomplished.
Reliable	Measurement procedures should produce similar results under the same conditions. The process must be consistent and repeatable.
Valid	• External validity: Findings are applicable to groups or contexts beyond those evaluated. The results can be generalized. • Measurement validity: What is intended to be measured is accurately measured. Outputs and outcomes are logically linked with measurable concepts. • Statistical validity (if quantitative measures are used): Appropriate statistical techniques are used to analyze data and support findings.
Cost-effective	Evaluation should be conducted at the lowest cost, without sacrificing quality.
Culturally appropriate	Evaluation should be inclusive, relevant, sensitive, and respectful across cultural groups.

Using Evaluation and Research to Develop Programs

When starting a new program or revising a program, look to existing studies for best practices. Information from prior Society evaluations and other evaluation and research conducted in your field of interest may be useful.

Conducting A Literature Review

A literature review is a compilation of the evaluation or research that has been published on a topic by professionals in the field. For either evaluation of program options or to understand best practices in carrying out programs, conducting a literature review may help you make decisions related to program development.

Do not limit your search to web search engines; it is often easier to find relevant, quality information from databases that index scholarly journals and books. You can access databases through public libraries and at the University of Minnesota.

When conducting a search:

- Familiarize yourself with the scholarly journal literature related to your subject area or field (most topics are interdisciplinary though so do not limit your search to only one field or subject database!).
- Identify keywords and phrases that describe your topic. Keep a list of possible synonyms.
- Select and search appropriate databases or search tools.

During your review, you may want to focus on what knowledge and ideas have been established on a particular topic, and what their strengths and weaknesses might be.

After you have gathered the existing research:

- Organize the information related to or around your evaluation or research question.
- Synthesize results into a summary of what is or is not known.
- Identify the gaps or areas of controversy.
- Formulate questions where further research is needed.

Overall, which program option has been the most successful, according to the literature? What best practices have you learned from the literature? You may use knowledge gained from the literature review to answer these questions and help you with decision-making in program development.

GUIDELINES FOR EVALUATING PROGRAMS AT THE MINNESOTA HISTORICAL SOCIETY

The remainder of this handbook provides detailed how-to instructions for staff that are doing evaluation for programs at the Minnesota Historical Society. These tips will help Society staff design and implement evaluation projects that are appropriate in scope, are feasible given the nature of the program, and produce useful and meaningful results.

Please note: The term "program" is used throughout the handbook and refers to any type of project, program, or activity that serves external users, including collections, exhibits, web sites, events, classes or field trips, and other services.

Statements in bold represent specific guidelines about the appropriate uses of evaluation at the Minnesota Historical Society. This standardization will allow the Society to roll up program-level evaluation results in key areas, to provide a picture of the impact and success of the organization as a whole.

The Minnesota Historical Society recommends that any program that meets one or more of the following criteria include evaluation on a regular, as needed, basis:

- **Funded by the Legacy Amendment Arts and Cultural Heritage Fund (ACHF)**
- **Funded by another source that requires program evaluation**
- **Budget of more than $25,000**
- **Ongoing for more than three months**
- **Provides direct service to or has contact with the general public**

The scope and extent of the evaluation should be determined by program managers (and, as needed, the department head), based on the stage the program is in, the resources available for evaluation, and other factors. The evaluation of individual programs should also be determined within the context of the department and program type to ensure evaluations conducted at the Society will support and enhance each other.

Minnesota Historical Society Evaluation Checklist

This Evaluation Checklist should be used by Society staff to ensure that all critical evaluation steps are completed. Each of these steps should be considered when planning a program evaluation, although not all of these steps will be needed for every evaluation.

PLANNING EVALUATION

___ Stakeholders identified

___ Program theory developed

___ Logic model template completed

___ Budget developed

___ Institutional dashboard measures identified

___ Evaluation plan template completed

IMPLEMENTING EVALUATION

___ Data collection tool(s) designed

___ Data collected

___ Data analyzed

USING EVALUATION RESULTS

___ Recommendations provided by evaluation lead

___ Recommendations used by program manager to improve program

___ Evaluation results used for grant writing and/or marketing purposes

___ Evaluation results used for broader organizational-level planning

___ Report(s) written

___ Report(s) posted for staff and other stakeholders

___ Report(s) posted publicly on Society web site

___ Presentation(s) developed

___ Findings presented to staff and other stakeholders

___ Findings presented publicly (e.g., event, meeting, etc.)

___ Results of methods presented to museum evaluation field (article, conference, etc.)

Phase 1: Planning Evaluation

The following steps should be completed before new programs are implemented, and should also be completed or updated for existing programs on an annual basis.

Remember to think about evaluation early in program development and implementation. The following steps will take you through the critical components of program evaluation in a systematic way, which will strengthen your program evaluation and generate more buy-in among key stakeholders. In addition, following these steps will ensure that program-level evaluation will be useful for organizational decision-making and communicating the impact of the Minnesota Historical Society's activities to stakeholders.

STEP 1: Identify Stakeholders

Stakeholders are individuals, internal and external to the Society, who have a vested interest in your program. To begin, answer the following questions:

- Who are they?
- What do they want to know?
- What actions or decisions may be made based on the information they provide?
- What kind of information would be most useful in shaping these actions or decisions?

Once you have listed the key stakeholders, think of ways to involve them in the evaluation process. Involve them early on and in meaningful ways. For example, you might convene an advisory group to review your logic model and help identify the key evaluation questions. Work with stakeholders to shape evaluation priorities and focus, being clear about the constraints (budget, contract requirements, etc.) that may limit your ability to incorporate all of their suggestions. You may wish to review preliminary results with your key stakeholders to help you interpret your findings. This ensures the evaluation meets stakeholders' needs and helps them better understand the results.

STEP 2: Develop A Program Theory

A program theory is a model of how a program should work that makes sense and is based on sound evidence. A good program theory logically and reasonably links program activities to one or more outcomes for participants.

Program theories can be captured in a series of "if-then" statements – IF something is done to, with, or for program participants, THEN theoretically something will change. For example, the collections department may have the following underlying theory: "IF historical records and other items are digitized, THEN participants, users, or visitors will have easier access and will use these items to learn more about history."

Program theories can be based on:

- Established theories or models of professional practice
- Previous research findings (from your own program or published literature)
- Expertise and experience of program staff
- Feedback and input from key program stakeholders

The purpose of this step is to understand and describe how the activities that are provided are believed to promote changes in participants. It is very important to go deeper than superficial assumptions about how activities link to outcomes. Instead, consider the available theories and research evidence that support these connections for your specific program:

- Using available resources (e.g., staff, equipment, materials), some kind of activities or services are provided.
- Program participants engage in these activities and react to what they experience.
- As a result of their involvement in these activities, program participants experience changes in their knowledge, attitudes, and skills.
- These knowledge and attitude changes promote changes in participants' behavior and practice.
- As a result of these behavior and practice changes, the program has an overall impact on the individuals served or on the broader community.

To develop a program theory, complete the following steps:

1. List the major activities or services that your program provides.
2. Review your list of activities and consider the following questions:
   - Do each of these activities refer to services provided directly to participants? Administrative functions of the program, such as hiring staff or preparing budgets, are an important part of providing community programming and should be reflected in your work plans. However, administrative activities that are not expected to lead directly to changes for participants should not be included in a program theory.
   - Does your list contain any redundancies (i.e., same basic activity described in several different ways)? If so, eliminate duplicate activities. In designing your evaluation, consider your core activities without redundancies.
   - Of those activities listed, which do you feel are most important in terms of either the potential for impact on the participants or the level of resources that are devoted to the activity?
3. Review the list of activities and prioritize them beginning with the most important activities provided by the program.
4. Develop a program theory by answering the following questions for each of the activities you prioritized.
   - IF the activity is provided, THEN what should be the result for participants?
   - WHY do you believe the activity will lead to this result? Is there an underlying formal theory or set of assumptions guiding this activity?
   - What evidence do you have that this activity will lead to this result (e.g., data from your own or other programs, published literature, etc.)?

STEP 3: Develop A Logic Model

A logic model is a visual representation of your program theory that illustrates the linkages between program components and outcomes. In addition to being a guiding tool for the program's evaluation, a logic model is also helpful for quickly communicating to a variety of audiences the core activities and outcomes of your program, how they relate to each other, and to the larger goals of the Society.

Your program's primary outcomes must support or contribute to the general institutional outcomes identified in the Society's logic model.

A logic model may not encompass every detail of the program or the evaluation. Instead, it should reflect your primary activities, and the most important outcomes you expect as a result of those activities. The logic model should demonstrate a connection between the program's activities and the intended outcomes. However, one activity could lead to multiple outcomes or multiple activities could lead to only one outcome.

A logic model usually involves the following elements — inputs, activities, outputs, and outcomes — displayed in a linear (left to right) flow chart:

INPUTS →	ACTIVITIES →	OUTPUTS →	OUTCOMES
The resources and materials that go into the program. e.g., funding, staff or volunteers, facility, supplies, etc.	The major services provided directly to the participant.	The amount of service provided, most often expressed in numbers. e.g., number of people who participate in an activity or hours of services received.	Actual impact (the change that results), which can be short-term, intermediate, or long term. • Short-term: Knowledge, attitudes, skills • Intermediate: Behavioral changes • Long-term: Community impacts, global changes, etc.

To start your logic model, clarify the linkages between your activities and intended outcomes. For each activity, use your program theory to list the immediate, intermediate, and long-term outcomes. Consider the order in which outcomes should appear. What should be the first changes experienced by participants? How should these initial changes promote other, more long-term changes? It is generally unlikely that behavior changes will be the first changes experienced by participants. According to the theory of action, behavior and practice changes are usually preceded by changes in knowledge, attitudes, or skills. Community impacts usually cannot occur unless individuals in that community have changed their behaviors or practice.

Next, add inputs and outputs. List your major program inputs, the resources and raw materials that go into a program. Anything a program uses to accomplish its purposes is an input. Then add outputs, anything the program produces. Outputs are different from outcomes. While outcomes describe the actual impact of a program on participants, outputs simply describe the amount of service that was provided to participants. Outputs are most often expressed in terms of numbers or amounts, such as the number of people who participated in an activity or the amount of service that was received.

Once you have all the elements of your logic model developed, use the Minnesota Historical Society Logic Model Template in the Appendix to document it in a standard template. You can use the logic model for evaluation monitoring and for future reference.

Once the logic model is drafted, review and revise. Getting input from stakeholders is recommended, but not always feasible. At a minimum, your colleagues who are directly involved

in the implementation and/or evaluation of the program should be given an opportunity to provide feedback on your logic model. Other primary stakeholder groups to consider getting feedback from include: other Society staff who are less involved in the program, program users or participants (if applicable), volunteers, members, and funders.

Engaging stakeholders will increase buy-in to the evaluation process and result in a better logic model. Getting stakeholders' input on your logic model can help identify inaccurate assumptions and build consensus about the program's intended outcomes. When discussing the logic model with stakeholders, the following questions may be useful:

- Does the logic model include all of the most important activities and services?
- Are outcomes clear and realistic? Do they represent meaningful changes?
- Are the connections between components logical? What evidence supports the connections?

A logic model should be a living document that is reviewed and revised regularly to capture changes in resources and activities, as well as new evidence about what outcomes are reasonable to expect from your activities.

STEP 4: Determine Resources Available for Evaluation

While evaluation does not need to be expensive, time and financial resources are needed to plan the evaluation, collect the right information, and use the results effectively. A guideline is to allocate about 10 percent of the total program budget to evaluation, including both staff time and expenses. When seeking funds for new programs, be sure to include evaluation in your budget to potential funders. The budget will help to determine the appropriate scope for the evaluation. Answers to the following questions may help to maximize limited resources:

- Does the program collaborate with other departments or organizations? If not, is there potential for collaboration? In this way, there could be opportunities to conduct a shared evaluation.
- What resources can be used at little or no cost? Some examples include volunteers, donated materials, etc.
- How can evaluation be more strategic? Prioritize the most important issues and select the most cost-effective techniques.

It is important to keep in mind the resources that are available for evaluation and to design the evaluation scope and activities to be feasible given those resources.

STEP 5: Identify Appropriate Institutional Dashboard Indicators for Your Program

An institutional dashboard is a set of indicators that provides a snapshot for stakeholders about the organization's outcomes and impact. To report these outcomes at the organizational level, individual programs must all measure these concepts using the same methods and the same survey questions. The Minnesota Historical Society has committed to annually reporting progress and outcomes on key measures of success. The dashboard indicators correspond directly to the outcomes in the organizational logic model.

This section of this handbook describes the Minnesota Historical Society's dashboard indi-

cators and provides instructions for how to identify which dashboard indicators to include in your program's evaluation. The Appendix includes the actual survey questions you should include on evaluation tools so these outcomes are measured consistently across the Society.

Not all of the dashboard indicators are relevant or appropriate to measure for all programs. Therefore, **the dashboard indicators that you select for your program should be determined based on which of the Society's key organizational outcomes is a part of your program's logic model.**

Several of the dashboard indicators are applicable to all programs that serve the public. **All programs that serve the public should include the first three indicators on the list below. Programs that have ongoing involvement with users (i.e., more than a one-time visit to a site or exhibit) should include one or more of the remaining items, depending on the program's logic model.**

REQUIRED INDICATORS:

- **Positive experience:** User/visitor/participant rating of their overall experience AND likelihood that they will recommend the Minnesota Historical Society program to their friends and family ("Net Promoter Score")
- **Appreciation for history:** User/visitor/participant rating of their interest in history
- **Net Promoter Score:** Likelihood users/visitors/participants will recommend MHS to their friends and family (score is computed by counting those who rate a 9 or 10—"promoters"—minus those who rate a 6 or lower—"detractors")

OPTIONAL INDICATORS:

- **Organizational value:** User/visitor/participant rating of the Minnesota Historical Society as a valuable resource to Minnesota
- **Program relevance:** User/visitor/participant rating of the extent to which the program connected the past to their present and future
- **Perceived learning:** User/visitor/participant rating of the amount they learned
- **Fostering ongoing learning:** User/visitor/participant plans to pursue other opportunities to learn about Minnesota history

These additional indicators are relevant for certain user groups/stakeholders. **Programs that target the following user groups should measure these indicators:**

- **Students:** Learning (the 4 Cs—communication, collaboration, critical thinking, and creativity)
- **Teachers:** View the Minnesota Historical Society as a valuable resource for teaching
- **Partner organizations:** View the Minnesota Historical Society as an effective partner in preserving, protecting, and promoting historic information, artifacts, and places
- **Researchers and professional historians:** View the Minnesota Historical Society as a valuable resource for conducting historical research

STEP 6: Develop An Evaluation Plan

Minnesota Historical Society programs that are required to complete a program evaluation must have a completed evaluation plan that should be submitted to the Evaluation Coordinator within one month of program initiation.

An evaluation plan details the practices and procedures you and your colleagues will take to successfully conduct your evaluation. The evaluation plan should include enough detail so all staff understand roles and responsibilities for each step of the program evaluation.

Involving your stakeholders in developing the evaluation plan increases engagement, consensus, and creates a better evaluation process and product overall. Multiple revisions of the evaluation plan may be needed.

The following subsections provide instructions for staff who are completing the Minnesota Historical Society Evaluation Plan Template, which is provided in the Appendix.

Program name: Name of the program being evaluated.

Department(s): Name of the department(s) responsible for the program.

Evaluation lead(s): Society staff member(s) who are responsible for ensuring the program evaluation is completed. This should typically be the program/project manager or the person who is responsible for implementing the program. The evaluation lead is responsible for implementing the evaluation plan, reporting the results of the evaluation for the institutional dashboard, and completing any required external reporting.

Program description: A very brief (1-2 sentence) description of the program, including key activities and intended outcomes from your logic model.

Program participants/Target audience(s): The target audience the program serves. Please note which of the four key user groups identified in the Minnesota Historical Society's strategic plan (students in grades 4-12, teachers, older adults age 50+, and young adults) are targeted audiences for this program.

Program intensity: Describe how long the target audience experiences the program. For example, is the program a one-time, one-hour walk through a museum exhibit, one-day training, program with weekly meetings for a month, or a semester-long, twenty-hour per week internship?

External evaluation requirements: A brief description of any external evaluation requirements, including who the requirements are from (funder) and the timing and specific type of evaluation required.

Evaluation resources: List the resources (staff time, financial, etc.) you have available to conduct the evaluation, as identified in Step 2.

Key evaluation stakeholders: A brief description of evaluation stakeholders. Note any concerns or challenges that have been identified and the process used to involve stakeholders in the evaluation design.

Reporting period: Note the frequency and timing of when you anticipate conducting the evaluation and reporting your evaluation results. In general, on-going programs should be evaluated annually. **For time-limited programs, the evaluation timeline may go longer**

than the length of the program to give you time to analyze and write-up results after the program is completed. For example, a summer program may run from June to July, but the evaluation may run from May to August to provide time for evaluation planning on the front-end, and analysis on the back-end. Any required interim reporting dates should also be noted. Time-limited programs that end during the Minnesota Historical Society's annual reporting period (July – June) are required to report evaluation results upon completion of the evaluation, no later than one month after the end of the fiscal year (July 30). Ongoing programs are also required to report evaluation results by one month after the end of the fiscal year (July 30). Programs may also have other evaluation reporting requirements or programs may have a need for information that would necessitate a different reporting timeline.

Evaluation stage(s): Note which stage(s) of evaluation are included in the current plan. Evaluation may be conducted at several points throughout the development of a program.

- Front-end evaluation should be implemented during the planning and conceptual design phase of a program BEFORE program development. This stage of evaluation is appropriate before the program design is finalized and can be used to help inform program development, such as learning more about your audience, for example.
- Prototyping includes any type of evaluation of programs in the process of being developed and not yet available for public use. The purpose of evaluation at this phase is to assess impact and provide information for improvement.
- Process evaluation assesses areas of strength and areas for improvement in terms of the process of implementing the program. Not focusing yet on the program outcomes or results, this type of evaluation should be used once a program is fully operating but is new or in transition. It is particularly useful to evaluate process when working with partner organizations and/or if the implementation of the program was particularly challenging and/or if new methods are being used to implement the program.
- Remedial evaluation refers to the period of time in which an overall program is studied after installation, often during a "soft opening." Physical elements are usually the focus of this phase (e.g., exhibit traffic flow, web site usability, etc.) and may also include specific studies of how visitors are using individual components.
- Summative evaluation measures participant outcomes; therefore, it should be used once the program is considered fully operational and established. The purpose of this phase of evaluation is to assess effectiveness or impact of the program.
- Other methods of evaluation may include needs assessments or different types of exploratory research (e.g., benchmarking). See also the appendix on economics-based research.

Primary data collection method(s): While not exhaustive, these are some of the most frequently used methods for collecting evaluation data. Choosing which method is best suited to use depends on your audience, the extent they interact with you, and which dashboard indicators and other outcomes you are measuring. Each data collection method has strengths and weaknesses; choosing multiple methods and collecting both qualitative and quantitative data will strengthen your evaluation. See the chart on the next page.

METHOD	USE TO…	DETAILS	EXAMPLES
Administrative records	Gather information about outputs.	This is a relatively simple and cost-effective way to learn more about programs. Information must be updated regularly in order to be both accurate and meaningful.	• # of visitors • # of publications • # of items preserved • Volume of service • Intensity of service • Reach of program (materials distribution)
Document review	Examine internal documents to provide insight about a program's context, development, and change.	Document review is also low in cost, but should not be the only method used in an evaluation. Documents may be incomplete or missing.	• Mission statements • Organizational charts • Annual reports • Funding proposals • Promotional literature
Observation	Provide a better understanding of what is or is not happening, instead of relying only on what participants are reporting.	Participants may or may not be aware they are being observed during unobtrusive observation. Participatory observation (i.e., making the participant aware of being observed and asking questions throughout the observation period), on the other hand, may also be beneficial in understanding the experience.	• Unobtrusive observation • Participatory observation
Focus groups	Understand a range of opinions on a topic and/or explore in-depth reactions to a particular concept (e.g. helpful if your questions are open-ended).	This method does not produce results that are representative in a statistical sense of all users of a program. Participants are generally selected because they share certain characteristics which make their opinions particularly relevant to the study. The interaction in a focus group may generate ideas individuals would not have thought of on their own.	• Traditional focus groups (6–12 participants) • Mini focus groups (4–5 participants) • Online focus groups
Interviews	Provide rich, contextual information about a person's experience with a particular program or opinions about a specific concept.	This method typically does not produce representative results (i.e., the results should not be generalized to all program users), unless careful sampling techniques are used. Using this method also requires a trained interviewer to minimize bias.	• Informal conversational interviews • Guided semi-structured interviews • Standardized open-ended interviews
Surveys	Obtain information about program users and their experiences.	Surveys are a very useful and relatively easy technique for obtaining information. They should be kept relatively brief (aim for no more than 1 page, double-sided). They can include open and/or close-ended questions, but in general surveys do not produce as much in-depth contextual information (compared with focus groups and interviews). It is important to use appropriate sampling techniques and methods to ensure a good response rate.	• Self-administered • Staff/volunteer administered • In-person • Mail • Web

Sampling strategy: A sample is a subset of the entire "population" of users of a particular program. Sampling selects a portion of the population to represent the whole population. Evaluations of more intensive programs (i.e., those that involve more than a one-time visit to a site or exhibit) will likely include all users, so sampling is not needed. However, for programs with a large reach, such as the Society web site or museum exhibits, it is not necessary, and difficult, to include all users in your evaluation. Instead, you should collect data from a sample of the users that are representative of all users. In determining who to sample, consider what demographics are needed for the sample to be representative of the target population. Whose voices need to be heard for the evaluation to be inclusive?

Sites, exhibits, collections, stores, etc., may want to produce evaluation data that are representative of all users. In this case, a "continuous random sampling" technique can be used. In "continuous random sampling," the data collector draws an imaginary line on the floor. The first eligible visitor to cross the line is observed, surveyed, interviewed, etc. After data gathered from the visitor is complete, repeat the process. This procedure ensures we maintain a random sample that is representative of visitors. Dates (days of the week, times of day, and time of year) should also be considered in designing the sampling strategy and choosing sample size.

Measuring process: Any process issues/questions and process-related items on the program logic model should be inserted in this table on the evaluation plan. For each item, identify the source of the data. Make sure it is an appropriate source given the data collection methods selected above. Some of the more common process questions include:

- Motivation: Participants' reason(s) for using the program.
- Accessibility: Degree to which program met participants' needs, particularly related to cost, language, culture, location, timing, etc., and the degree to which any subsidization by the Society increased accessibility.
- Exhibit use and engagement: The way in which visitors use or interact with the exhibit, including stopping behavior, path, traffic, social interactions, use of interactive elements, noticing labels, reading labels, etc. The degree to which program users are involved in the program as intended.
- Collaborations/partnerships: The effectiveness of any collaborations or partnerships within the Society or with other organizations. Include the strengths of these partnerships and any challenges or barriers to effectiveness.

Also, include the specific measure, which is the actual survey, focus group, or interview question, to be used. The analysis method is the specific result to be reported from that measure. Some process measures may have quantitative results, but more often they are qualitative in nature. Then, briefly describe any potential challenges or concerns related to this process item, as identified by stakeholders in Step 1.

Once the period of evaluation is completed, the evaluation lead should insert any comments about changes that are made to the program as a result of the process evaluation results.

Measuring outputs: The outputs on the program logic model should each be inserted in the table in the evaluation plan. For each output, identify the source of the data. Make sure it is an appropriate source given the data collection methods you selected above. Some of the more common outputs include:

- Number: This could include the number of people who use or participate in the program; the number of artifacts obtained, preserved, cataloged, and/or made available to the public; the number of web users; the number of newspapers digitized, etc.
- Demographic characteristics of participants: This may include age, gender, race/ethnicity, primary language, income, education level, household/family type, home location, member status, museum visitor group, museum visitor history, etc.

Also, include the actual survey, focus group, interview question, or the specific field in an administrative database to be used. **The Minnesota Historical Society has standard questions for most demographic variables—see the Demographics Questions Template. Depending on your program it may be appropriate to only ask one or two of these demographic questions or to ask all of them.**

The analysis method is the specific results you will report. In this column, describe specifically how you plan to analyze and report the data.

Then, insert the target, which is the goal for the program on this item. These targets should be set with program staff and managers after consideration of budget, reach, capacity, etc. Any targets that are noted in original program proposals for funding should be referenced here.

Once the period of evaluation is completed, the evaluation lead should insert the actual evaluation results for each item.

Measuring dashboard indicators: The dashboard indicators you selected in Step 4 should each be inserted in this table. For every dashboard indicator you selected, identify the source of the data. Make sure it is an appropriate source given the data collection methods you selected above.

Also, select the specific measure you will use to gather data on this dashboard indicator. **The Minnesota Historical Society has standard questions for all of the dashboard indicators—see the Dashboard Indicators Questions Template.**

The analysis method is the specific result to be reported from that measure. **For the dashboard indicators, the Minnesota Historical Society has determined standard analysis and reporting methods. Once the evaluation is completed, the evaluation lead should insert the evaluation results for each item based on these analysis methods.**

Measuring other outcomes: Any other outcomes from your program logic model that are not dashboard indicators that you plan to measure should be inserted in the table in the evaluation plan. For each outcome, identify the source of the data. Make sure it is an appropriate source given the data collection methods you selected above. Good indicators are also SMART indicators:

- Specific – concrete, explicit, Who, What, and How
- Measureable – countable marker for success
- Action-oriented– focused on change
- Realistic – meaningful, but reasonable for resources
- Timely – explicit time parameters

Also, include the specific measure, which is the actual survey, focus group, or interview question; the specific field in an administrative database to be used; etc.

The analysis method is the specific result to be reported from that measure. In this column, describe specifically how you plan to report the data. Once the period of evaluation is completed, the evaluation lead should insert the actual evaluation results for each item.

Timeline: Complete the timeline table to illustrate the key steps in implementing the evaluation, who is leading that task, and the timeline for completion. Note any interdependencies; for example, data collection cannot begin until data collection tools are developed, etc.

STEP 7: Design Data Collection Tool(s)

Designing unique data collection tools may not be necessary with each project, as the Society has a number of templates to assist you in measuring the dashboard indicators (as described above). However, you may adapt the templates if you would like unique evaluation tools or have project-specific questions that need to be added.

A general rule is that if something is not in your evaluation plan, it should not be included on the tool. Developing a thoughtful and thorough evaluation plan will ensure that your data collection tools are efficient and are measuring what is most important for you to know.

COMMON ISSUES TO AVOID WHEN DEVELOPING DATA COLLECTION TOOLS:
• Avoid leading or loaded questions
• Avoid questions that ask for two or more pieces of information (often referred to as "double-barreled questions")
• Avoid jargon, slang, abbreviations; instead, use words respondents will understand
• Before collecting data, it is important to "pilot" or test the tool and gather feedback to see if questions are being interpreted accurately and how the tool works. This is particularly important if you are using any new or previously untested questions or tools.

Phase 2: Implementing Evaluation

Once you have carefully planned your program evaluation, implementation should be smooth. Just follow the steps on your evaluation plan. The following sections provide tips for every step of the process.

STEP 1: Collect Data

Now it is time to gather the information you identified in your evaluation plan. Before you can approach users to participate in your evaluation, there are several considerations:

INCENTIVES

If possible and ethical, include incentives (such as a small prize, museum tickets, gift shop certificate, etc.), to increase the response rate and thank participants for their input.

INFORMED CONSENT AND CONFIDENTIALITY STANDARDS

Everyone who participates in the evaluation should do so willingly. In general, people participating in any research project, including a program evaluation, have the right to:

- Choose whether or not they want to participate without penalties (e.g., participation in the evaluation should not be a mandatory requirement for program participation).
- Withdraw from the evaluation at any time, even if they previously agreed to participate.
- Refuse to complete any part of the evaluation including refusing to answer any questions.

The word "informed" is important – in addition to choosing whether or not to participate in the evaluation, people have the right to understand all implications of participating. To ensure that potential participants can make an informed decision regarding their involvement, you should:

- Provide potential participants with information about the evaluation, including why it is being done, what you are asking them to do, how you will you use the information, and how long it will take.
- Describe both the potential benefits of participation and any foreseeable risks, including possible discomfort due to participation.
- Share this information using language all participants can understand – avoid jargon and translate if needed.
- Allow the participant the opportunity to ask any questions about the evaluation.

When working with children, parental consent may be required, except in cases of minimally invasive evaluation methods such as short surveys or prototyping participation. Participants do not need to sign a consent form if they are adults capable of making decisions and will not be put at significant risk by participating in the evaluation. For example, if you want participants to fill out an optional anonymous survey asking them if they were satisfied with specific elements of a program, the fact that they complete and return the survey can be used as their consent. Signed consent forms may be necessary in other situations, however, especially if you plan to:

- Collect very personal or sensitive information.
- Use the results for purposes other than program improvement, such as publication,

training activities, or participation in a larger research project.
- Gather information about participants from other third parties, such as program staff, teachers, family members, or others.
- Require significant time or effort on the part of participants, such as asking them to participate in multiple or time-consuming interviews, and especially for children (in which case parent consent should be obtained).

If you are unsure as to whether or not consent is necessary, please consult your supervisor.

It is not always possible for evaluations to be conducted anonymously, without collecting identifying information such as a participant's name or phone number. However, all information gathered should be considered confidential and should not be shared with others. To ensure confidentiality, consider these strategies:

- Collect data in a private location where surveys cannot be seen and interviews cannot be overheard.
- Do not discuss information about individual participants with other people. Findings should generally only be discussed at an aggregate level or with identifying information disguised.
- Keep completed surveys or interviews in a secure location where they cannot be seen by other people.
- Securely dispose of completed evaluation materials when they are no longer needed.

You may encounter situations in which you believe that it is important to disclose confidential information. To the extent possible, consider in advance the types of disclosures that may be needed and develop a plan to handle these situations.

DATA COLLECTION STANDARDS

This section gives general data collection tips, as well as more specific tips for different data collection strategies. See chart on next page.

GENERAL TIPS

- Encourage participation. Explain the purpose of the survey, interview, or focus group, why it is important, and how you will use the information.
- Before collecting data, make sure the method and the instrument is understood by all staff or volunteers who are collecting data. Do not hesitate to ask questions if anything is unclear!
- Take as much time as you need to review the data collection instrument after data has been collected. Make sure each question is answered, legible, and that the instrument is numbered (if applicable). Double-check. This will save time later in data entry and analysis.
- Always thank the respondent(s) for their time. Remember to leave time at the end to answer any questions that respondent(s) may have.

FOCUS GROUP TIPS

- Prepare participants prior to the focus group. Participants should have information about the objectives of the focus group before it happens.
- Create a comfortable environment—this helps participants speak openly. Some things to foster comfort: seat them in a circle, provide food, allow time for small talk at the beginning, and start with a "simple" question to ease participants into the process.
- Remember to review ground rules—everyone's ideas are important, everyone has a right to speak, there are no right or wrong answers, negative comments are useful in gaining insight about the topic, all comments are confidential, and only summarized information will be communicated.
- Start with broad, general questions and gradually get to more specific questions, ending with the most important questions. The first questions will help the flow of conversation, trigger memories, and help with "harder" questions later on.
- Summarize key points at the end. Take the key points and repeat them back to the group; ask if these points are an accurate assessment of their answers.
- Additional tips for encouraging discussion include: Asking participants to think about and write down their responses before sharing, giving reading material for participants to reflect upon, noting responses on a flipchart or a whiteboard, etc.

INTERVIEWING TIPS

- Make sure your initial contact is positive. Clearly explain the purpose of the interview and how long the interview will take. Give the respondent an opportunity to ask questions and discuss any concerns they may have. It is okay to spend a little time making small talk before you start the interview. This will help you and the respondent get comfortable and build rapport.
- Interviewing one person in a visitor group can sometimes be challenging, especially when others within the group want to participate. Be flexible in the approach. Some group members may wander away and ask the interviewed person to catch up or the interviewed person may want the data collector to follow the group as the interview is being conducted. Be conscientious of these dynamics and be accommodating. Frequently, group members (particularly young children) will also want to listen or participate. Be as inclusive as possible but only write down responses from the one person. Focus on these answers.
- Remember to listen and don't correct the respondent if (s)he gives an incorrect response. Along the same vein, the data collector may be asked many questions. Feel free to say "I don't know." The data collector is not expected to be an expert.
- If a response is unclear, does not address the question, or more detail is needed, probe. Examples of probing questions include: *Is there anything else? • Could you talk a little bit more about that? • Why do you feel that way? • Would you explain further? • Would you give me an example of what you mean?*

INTERVIEWING TIPS

- Make sure your initial contact is positive. Clearly explain the purpose of the interview and how long the interview will take. Give the respondent an opportunity to ask questions and discuss any concerns they may have. It is okay to spend a little time making small talk before you start the interview. This will help you and the respondent get comfortable and build rapport.

- Interviewing one person in a visitor group can sometimes be challenging, especially when others within the group want to participate. Be flexible in the approach. Some group members may wander away and ask the interviewed person to catch up or the interviewed person may want the data collector to follow the group as the interview is being conducted. Be conscientious of these dynamics and be accommodating. Frequently, group members (particularly young children) will also want to listen or participate. Be as inclusive as possible but only write down responses from the one person. Focus on these answers.

- Remember to listen and don't correct the respondent if (s)he gives an incorrect response. Along the same vein, the data collector may be asked many questions. Feel free to say "I don't know." The data collector is not expected to be an expert.

- If a response is unclear, does not address the question, or more detail is needed, probe. Examples of probing questions include:
  - *Is there anything else?*
  - *Could you talk a little bit more about that?*
  - *Would you explain further?*
  - *Would you give me an example of what you mean?*
  - *Why do you feel that way?*

More examples:

SURVEY QUESTION	RESPONSE	GOOD PROBE	WRONG!! LEADING PROBE
What did you like best about the program?	I don't know.	Whatever you think is fine.	Didn't you like the staff?
How many classes did you participate in?	Oh, about 10 or 12 classes.	Would that be closer to 10 or to 12?	Interviewer writes 11.
How would you improve the program?	Everything!	Please be more specific?	You mean you didn't like anything?
What is your race or ethnicity?	Well, I'm a mixture.	A mixture?	Are you part Black or Hispanic?

- If a respondent chooses more than one answer to a close-ended question that asks them to choose only one answer, request that the respondent pick the best response. If respondents still refuse to choose one (such as choosing both "Very Likely" and Likely"), the lesser of the two responses should be marked ("Likely"). The same standard is applied to data entry and analysis.

- If a potential respondent refuses to participate, clarify why they are declining in order to record the information on the Refusal Log. This is important for data collection as well as for data analysis, in order to calculate an accurate response rate. Further, below are some of the most common reasons for refusal and examples of how data collectors can reply:
  - » "We don't have much time."
    - Say: "This should just take about ___ minutes of your time. It would really help us better understand ___ so we can improve ___."
  - » "I'm tired of interviews."
    - Say: "I'm sure you get asked to do interviews all the time. This one will help us understand how to meet needs and interests of people like you. We would really love your input."
  - » "I'm not the type you want."
    - Say: "We want to hear from everyone. This way we can better understand all our ___ to improve ___."
  - » "My English is not good."
    - Say: "That's not a problem. I can go over any questions that are not clear. We want to hear from everyone. This way we can improve our ___ for everybody."

- If a respondent ever makes a data collector feel uncomfortable, the data collector should politely end the interview and leave the area. If a respondent is angry or upset, find a manager, and call security if needed.

OBSERVATION TIPS

- As described in the section above on sampling strategies, remember to use a continuous random sample. It is easy to forget this and observe visitors who are already there before the data collector arrives or who are already there when the data collector is finished collecting data from a previous visitor.

- "Stops" are often key in observation of museum visitors. A "stop" is defined as "a visitor stopping with both feet planted on the floor and head or eyes pointing in the direction of the element for three seconds or more" (Serrell 1998).

- Recording time: Write out times in hours, minutes, and seconds. For example, if a visitor is observed for half a minute, write "00:00:30".

- When recording behaviors, time and track only ONE person at a time. This allows the data collector to focus on their behavior with an exhibit and/or with other visitors. Trying to time and track multiple people is distracting and the data collector may miss out on key behaviors.

STEP 2: Analyze Data

Once data are collected, you need to organize and analyze the information, consulting with other staff members when possible. Your analysis strategy should match your data and evaluation questions. Although there are many forms of analyses, the following text provides a brief overview.

You will first need to prepare your data by entering the information.

PREPARATION FOR DATA ENTRY

Data entry is often thought of as a time-consuming process, but there are steps you can take to make the process more efficient. You will save time in the long run if you take time up front to prepare for data entry. Below are some general steps to help you get started:

- Assign an ID number to each form or survey to be entered, and write the number at the top of each survey. These can be numbers such as 1, 2, 3, 4, but each number should be used only once, even across different batches of surveys. This will make it much easier to go back and re-enter data if you realize you have made a mistake.
- In some instances, like surveys you collect on a recurring basis from the same people (such as teachers), you may wish to continue to assign the same ID for each respondent. You will need to maintain a master list that you can reference for assigning and tracking these ID numbers in the future.
- Schedule a large enough block of time to enter an entire batch of data at once. The time needed for this will vary depending on the length of the survey and the number of participants, but it is best to enter all of the information at the same time if possible. This will minimize the chance that you enter the same survey twice, or forget to enter any remaining surveys.
- Before you begin entering the data, take time to go through each completed survey and identify questionable responses. See the section on making data entry decisions to identify some common unexpected responses that survey respondents make, and take note of the tips for working with difficult or confusing surveys. By taking time up front to identify potential problem areas, you can make consistent decisions about what you plan to enter in each situation, and you will save time once you begin entering the data.

MAKING DATA ENTRY DECISIONS

No matter how clearly a survey is written, there will be some survey respondents who do unexpected things. For example, respondents may choose multiple answers even when asked to choose only one, they may skip questions, or it may just be difficult to understand their intended response. The following are some common issues that you may discover and some ideas for navigating those difficult surveys. Once you have made a decision about how to treat a particular issue, make note of it in a separate document, or even in the margin of these instructions, to reference later and help ensure consistency in your decision-making process.

- If data are missing or unintelligible on a staff completed form, try your best to fill in the information by following up with the staff member who filled it out. If you are unable to obtain the missing information, or if the form was completed by a parent, just leave the space blank in the database. You should not try to guess what the respondent might have been thinking.
- A participant may respond to a numerical question with a range of numbers (e.g., "1 or 2" or "5–7") or a vague reference (e.g., "a couple" or "several") instead of a single number. In these cases, the response is too vague to translate into a single representa-

tive number, so you will simply leave this cell blank.

- Sometimes respondents will be unable to choose between two options such as "agree" and "strongly agree" and will select both! The lesser of the two responses should be entered ("agree").
- For those surveys in which an 'other' category is possible, you will have to decide how to treat these answers. Sometimes respondents choose 'other' but then provide an answer that closely aligns with one of your response categories. See the example below:

The highest level of schooling you've completed:

❑ Some high school	❑ Bachelor's degree
❑ High school graduate	❑ Some graduate school
❑ Some college/technical school	❑ Advanced degree
❑ Associate's degree/certificate	❑ Other: *Graduated high school*

In this situation, you would probably choose to recode the response "Other" to "High school graduate," as the response is reworded, but with the same definition.

After data entry, remember to:

- Review and correct data entry before analysis. You or someone else should check the data that has been entered. Try to check every 5 cases. If you have a large data set, try to check 10% of cases. If there are many discrepancies, you may need to check every case.
- Leave enough time and money for analysis—it is easy to focus on data collection and not leave enough time to analyze results.
- Identify the appropriate statistics for each question—get consultation if needed.
- Do not use the word "significant" when describing your findings unless the data have been tested and found to be statistically "significant".

QUANTITATIVE DATA

Quantitative data are information collected in numerical form (counts, percentages, rating scales). Close-ended survey questions, which limit responses to predetermined categories (e.g., yes, no, agree, disagree) may be given a numerical value in order to be analyzed quantitatively.

You may analyze your data using:

	DEFINITION	EXAMPLES
Descriptive analysis	These are simple statistics that can help you summarize your data and identify key findings, reducing your raw data to an understandable level.	• Frequency distributions: Counts that show how many participants fall into various categories (e.g. how many said they "agree" or "disagree") • Central tendency: The averaged response, also called a "typical score" • Variability: Amount of variation or disagreement in your results
Inferential analysis*	This analysis helps to determine whether results are meaningful/generalizable. In statistical terms, this is referred to as "statistical significance". Many statistical tests can be used.	• Chi-square tests • Correlations • T-tests • Analyses of variance

*Do not worry too much about using inferential analysis to obtain statistically significant results. **Statistical significance does not equal practical significance.** Statistically significant results do not necessarily lead to program improvement. Evaluation findings that are not statistically significant or that do not even test for significance may still be useful for program improvement.

QUALITATIVE DATA

Non-numerical information is qualitative. These include responses gathered through interviews, observation, focus groups, written documents or journals, or open-end survey questions. On its own, or in combination with quantitative information, qualitative data can provide rich, powerful findings.

Meaningful analysis of qualitative data can take time. The first step in analyzing qualitative is organizing your information. The simplest way is to transcribe your data into word processing software. Next, identify important elements. Because important information may be scattered throughout the document, choices must be made about which aspects matter to your evaluation. Focus on questions that you are trying to answer and the relevance of the responses to these questions. In that way, you can decide what information is fair to emphasize, minimize, or even leave out (if not relevant). You may highlight and note important content, segment and sort the data, rearranging them into logical orders or groups.

Identifying common themes in your data is also important. Referred to as "coding," a set of codes or themes may be developed before, during, or after collection of information. Choosing whether to use predetermined codes ("closed coding") or to develop codes after reviewing the data ("open coding") depends on your evaluation questions.

GENERAL GUIDELINES
• Use closed-coding if the responses are meant to answer questions about specific program goals or are based on particular theories you have about how your program works.
• Use open-coding if you are seeking to learn more or new information from responses, seeing what ideas may emerge.
• Coding can often combine both approaches, beginning with predetermined categories that you then apply and modify as you work, taking care to be open to what the data tells you.

As part of the coding process, develop a code book (a list of codes to apply to the data). Pay particular attention to how your codes relate to each other in terms of hierarchy and scope. Review and update your code book throughout the process. Create:

- **Code definitions:** Short identifying labels that clearly describe code meaning.
- **Decision rules:** Clear, commonly understood parameters for applying codes.

After you have processed your data, you need to understand what the data means. Following are some common methods, presented in order of increasing sophistication. The complexity of the data and the depth and scope of your evaluation questions will determine which is the most appropriate.

	DEFINITION
Reduction	Analyze in a way that summarizes and synthesizes the content to make it more manageable and easy to use.
Quantify	Count the frequency of themes or code occurrences to more accurately understand the key findings.
Content analysis	Reduce and quantify elements in the data to focus on larger themes, patterns, and categories to get a more complete picture.
Model	Use reduction, quantification, and content analysis to build typologies, taxonomies, or archetypes for understanding phenomena or grouping the data.
Theorize	As models emerge, the analyst uses them and other relevant knowledge to build and test hypotheses and establish casual links or relationships.

You can also validate your qualitative data in some of the following ways:

	DEFINITION
Strength	Data point (e.g., interview respondent) provides descriptive and thoughtful detail and context and is an appropriate and reliable source.
Internal verification	Data are consistent within a single source (e.g., interview respondent provides consistent information throughout process).
Inter-source verification	Data are confirmed by multiple quality sources in your data (e.g., multiple respondents provide the same or consistent information from different points of view).
External verification	The qualitative data are consistent with other data including quantitative data, existing literature, etc.

RESPONSE RATE

If analyzing survey data, it is important to calculate your "response rate," the number of participants who responded to the survey divided by the number of people asked to take the survey. Express the response rate in the form of a percentage. In general, a higher response rate means your data is more reliable. See the Refusal Log Template for an example of a method that you can use to keep track of your response rate. To calculate your response rate:

- Find the total number of people who responded to the survey.
- Determine the total number of people asked to participate in the survey.
- Divide the number of responses by the total number of people asked. e.g. 100 people are asked to participate. 70 people respond. 70/100 = 70%. Your response rate is 70%.

Phase 3: Using Evaluation Results

THE FOLLOWING SECTION PROVIDES AN OVERVIEW about making recommendations and communicating your findings. Understanding these areas will help you use results for program improvement, funding, marketing, and broader organizational level planning.

Also remember the following cultural considerations throughout this phase:

- Be careful not to generalize data or presentation of data to the predominant culture ("one size fits all" approach limits applicability to culturally diverse groups).
- Do not exclude findings relevant to culturally diverse communities.
- Present information in accessible ways (consider multiple channels of communication, different languages [if possible], etc).

MAKE RECOMMENDATIONS

After you have summarized and identified key findings from your analysis, you need to interpret results and draw conclusions. These tasks involve stepping back to consider what the results mean and what they imply about your work. **The evaluation lead is responsible for making recommendations, if appropriate.**

Ask yourself:

- What patterns and themes emerge in the results?
- Are there any deviations from these patterns? If yes, what might explain these deviations?
- Do the results make sense?
- Are there any findings that are surprising? If so, how do you explain these results?
- Are these results significant from a statistical standpoint? Are they meaningful in a practical way?
- Do any interesting stories emerge from the responses?
- Do results suggest any recommendations for improving the program?
- Do the results lead to additional questions about the program? Do they suggest that additional data may need to be collected?

Also remember to watch for and resolve inconsistencies. In some cases, you may obtain contradictory information. Stakeholders may expect different results or disagree with one another. It may be challenging to determine accuracy of information, especially when comparing different viewpoints or perspectives. Remember that various stakeholders may have valid viewpoints that vary based on their unique experiences. Try to resolve these discrepancies and reflect them in your findings to the extent possible.

Once you have reached a consensus, consider the following questions to help you make concrete recommendations:

- What do you and other stakeholders need to know more about?
- What decisions do you feel you need to make, but need more information?
- What will you do with the answers to your questions?

USE RECOMMENDATIONS

Once you have determined appropriate recommendations, articulate in writing specifically

how you might implement them. Determining your priority uses early in the process will help you make more effective decisions. **The program manager is responsible for implementing recommendations for program improvement, when appropriate.**

The following questions are helpful in going about implementing recommendations. Think about and be clear about addressing them:

- What are the different issues that are likely to surface related to these decisions?
- What other factors may affect the decision-making process?
- How will we know if evaluation results and process are used as planned?

COMMUNICATION OF FINDINGS

Sharing findings with key audiences and stakeholders is beneficial for a number of reasons. Sharing your results informs the public and other stakeholders about what you have done, what you intend to do, and how you are using your research in a meaningful way. Report findings can also help build community relations, develop partnerships, and sustain or secure funding

Whomever your audience(s), remember that content is not always best shared in a long and complicated report. A few concise pages may have the most impact. Instead of producing a document describing a complex set of ideas, consider dividing the results into several smaller reports.

Consider the easiest and clearest way to present the information to your key stakeholders. For example:

- Write separate executive summaries or articles, targeting specific audiences/stakeholders.
- Share your results with the media through a press release, conference, interview, etc.
- Consider publishing key results via social networking sites.
- Make presentations to community partners or potential funders.
- Make a short video presenting results to use in forums/discussions.
- Share results with professional communities, academic journals, etc.

You can also communicate your message effectively by:

- Knowing your audience—what will impact and what might overwhelm
- Determining audience interest in "hard facts" or anecdotal narrative
- Avoiding jargon an acronyms
- Using visuals and including clear and concise writing

STANDARD REPORTING PRACTICES

If you have a lot of information to report, it can be easy for readers to lose track of your main findings and conclusions. Make your key findings stand out, so that your audience can easily find them and determine their significance and usefulness. The following are examples of standard reporting practices that will help your audiences understand and follow the findings.

1. Be as specific as possible about who is reporting and what they are reporting on. Try to use wording that is as similar to the question asked as possible. For instance:
   - *Fifty percent of participating web visitors "agree" or "strongly agree" that the web site is laid out in a way that is easy to navigate.*
   - *The majority (65%) of teachers participating in the Northern Lights training program participated for more than 18 months.*

2. Report results from largest to smallest for multiple questions related to one another. This is generally the easiest way to read this information. For instance:

- *Seventy percent of library users reported that they use the library for genealogical research. Sixty percent reported that they encourage their family and friends to use the library for genealogical research.*

3. Know when to report in percentages and when to report in numbers. In general, we recommend using percentages to report information for samples with more than 10 participants. For samples with less than 10 participants, use numbers. For instance:

- *Of the 12 partners who participated in the training group during the first reporting period, 78 percent report increased capacity to preserve and promote access to historic resources. Of the eight partners who participated in the second reporting period, seven report that they have increased capacity to preserve and promote access to historic resources.*

4. Keep tables and graphs as clear and simple as possible. This narrative and table might look like the following:

- *Most of the staff reports of participating teachers' behaviors indicate that teachers understand the new state social studies standards. Staff also reported that teachers are able to find new ways to teach Minnesota history to their students (see Table 1).*

TABLE 1: Understanding social studies standards and teaching Minnesota history—Staff report (N=20)

	STRONGLY DISAGREE	DISAGREE	AGREE	STRONGLY AGREE
Teachers find new ways to teach Minnesota history.	5%	25%	40%	30%
Teachers understand social studies standards.	10%	10%	30%	50%

REPORTING MULTIPLE CHOICE QUESTIONS

Because most of the questions on the surveys included in your evaluation plan use scales such as "Strongly agree," "Agree," "Disagree," and "Strongly disagree," the agreement and disagreement categories can be combined to simplify the reporting of findings. That is, "Strongly agree" and "Agree" numbers or percentages can be added together, while "Disagree" and "Strongly disagree" numbers or percentages can be added together. For example, if 40 percent of visitors "strongly agreed" and 50 percent "agreed" that a new exhibit at the Mill City Museum is enjoyable, you might report:

- *Ninety percent of participating visitors enjoy the new exhibit at the Mill City Museum.*

REPORTING CONTINUOUS DATA

Another form of data that may be reported is continuous data (e.g., months, years, hours, and numbers of participants). Here are some suggestions for reporting this type of information:

- *Program participants were involved with this program for between six months and five years (average = 2 years).*
- *Partners spent an average of 20 hours in training during the first reporting period.*

- *Sixty percent of teachers participated in the program for less than 40 hours, while 40 percent participated for more than 40 hours.*

REPORTING OPEN-ENDED QUESTIONS

Some surveys also include open-ended questions, such as questions eliciting suggestions for improvement or level of interest in participating in an activity. Often open-ended data is used to support other quantitative data. For instance:

- *Eighty percent of web visitors felt that the web site met their expectations. When asked what they found to be the most positive thing about the web site, one participant stated, "I learned about new resources for teaching Minnesota history" while another said, "I love being able to view artifacts from my own home."*

Open-ended data may be presented with a verbatim list of responses or you may choose to code the responses. However, you may decide that your open-ended responses are more appropriate for internal use than external reporting and not report them. It is up to you and the requirements for your reports.

Verbatim list Open-ended responses can be reported as stand-alone data in the form of a list of verbatim responses. This is especially useful if you have less than 10 responses to any particular question. If you list responses verbatim, be sure to 'de-identify' the responses by removing any words or phrases within the response that could identify the respondent and replacing the identifying information with a general reference in brackets. For instance:

- *Teachers were asked what one thing they would change about the program to make it more helpful to them. Their responses include:*
- *"I would make the training times more flexible."*
- *"I wish that [trainer] would allow more time for questions."*

BE OBJECTIVE

However you choose to report your evaluation findings, it is crucial that you report objectively, including both positive and negative findings. Here are some tips for ensuring your objectivity and increasing credibility with stakeholders:

- Do not use emotionally charged language when describing your findings, like 'very' or 'extremely'. This can make you sound like a program advocate, thus reducing your objectivity and credibility.
- Use disappointing results to guide recommendations for enhancing services or addressing implementation barriers, rather than dismissing or hiding them.
- Discuss limitations in terms of how information was collected, so that audiences can judge the degree of confidence to place in the results. Every evaluation study has limitations, and it is important to know what they are so stakeholders can consider the findings in context.

APPENDIX

Using Research and Evaluation Based in Economics

The following is a summary of various economics-based techniques that may be relevant for evaluating the financial impact of the Society's programs:

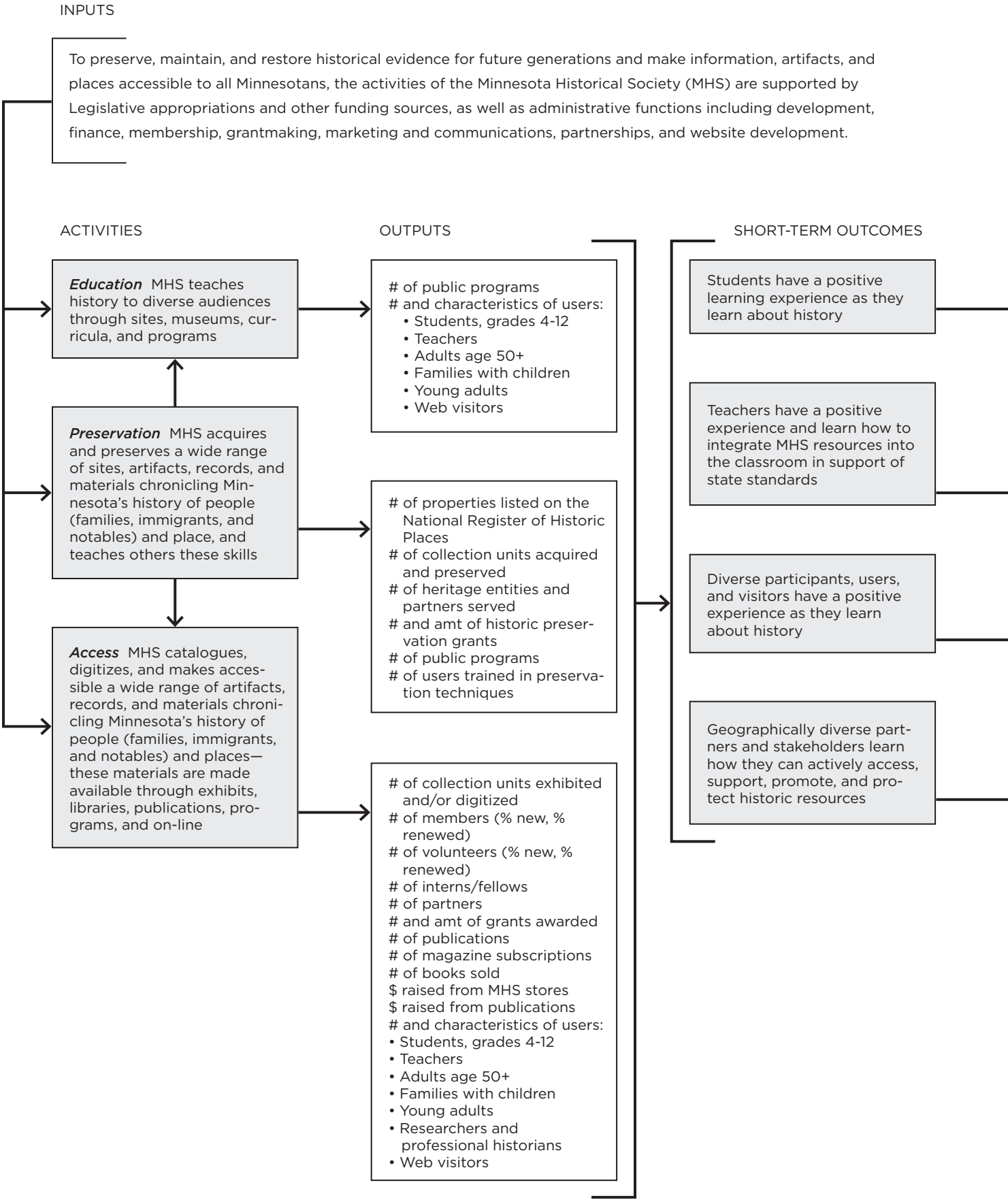
Cost-benefit analysis: Sometimes known as "running the numbers", a cost benefit analysis finds, quantifies, and adds all the positive and negative factors in dollar amounts. It subtracts the negative from the positive. The difference between the two indicates whether a planned program is affordable. It is important to include all the costs and all the benefits and to properly quantify them. Do not use this method if it is not appropriate to quantify costs and benefits in dollar amounts.

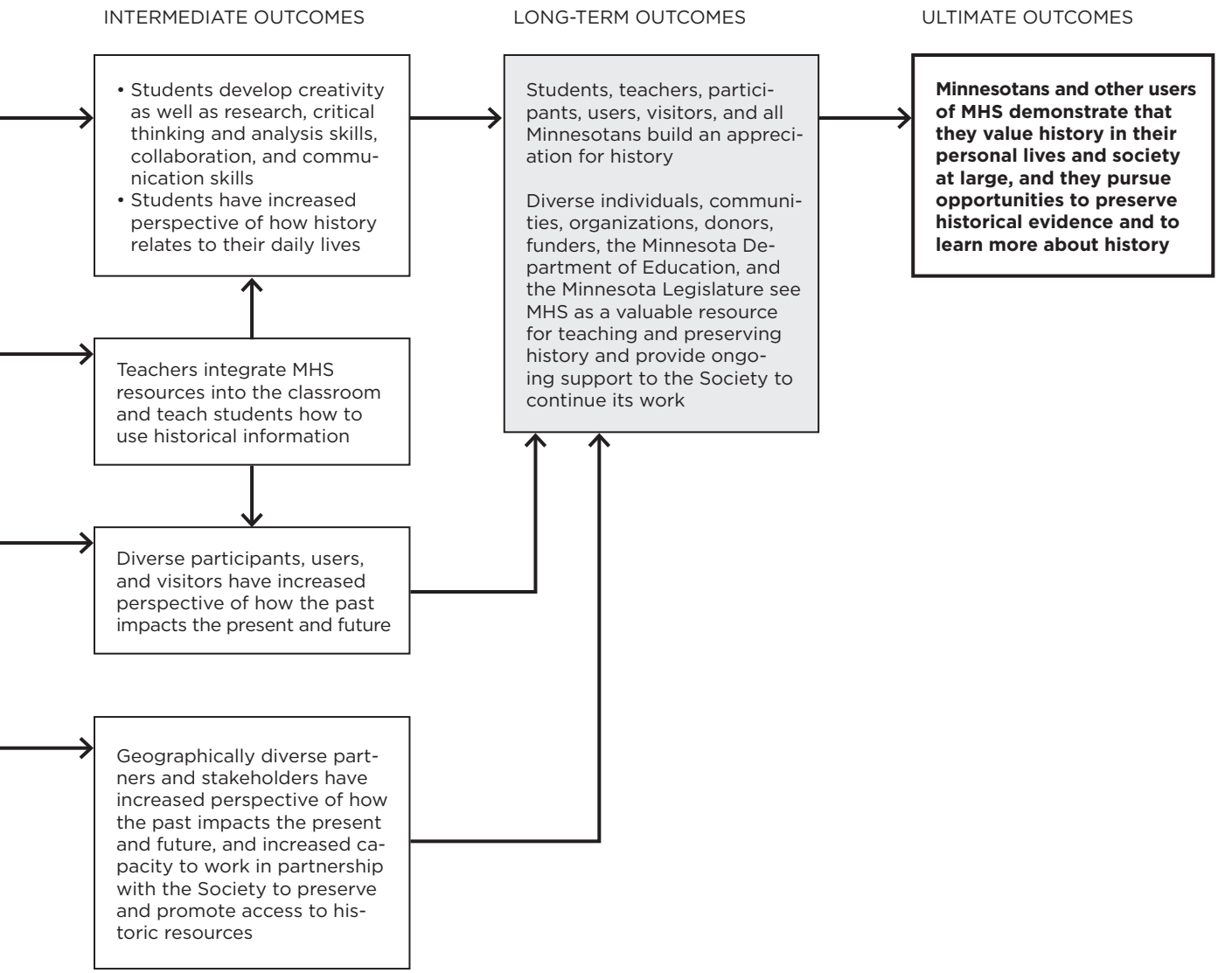
Cost-effectiveness analysis: This compares the relative costs and outcomes of two or more courses of action. It does not assign dollar values to measure effect (which may be social in nature). It is often expressed in terms of ratio. This method is frequently used in the field of health services. This method may be more appropriate than cost-benefit analysis if one or more variables may not be quantified into dollar amounts.

Return on investment: This is the ratio of money gained or lost on an investment relative to the amount of money invested. The technique calculates all the costs associated with an investment, estimates or calculates returns, establishes a timeline for costs and returns, and calculates annualized return of investment. This is usually expressed as a percentage.

Social return on investment: This method is a principles-based method for measuring "extra-financial value" (environmental, social, etcetera) relative to resources invested. It is used to evaluate impact on stakeholders, identify ways to improve performance, and enhance the performance of investments.

Minnesota Historical Society Logic Model REVISED June 20, 2011




INTERMEDIATE OUTCOMES
LONG-TERM OUTCOMES
ULTIMATE OUTCOMES
• Students develop creativity as well as research, critical thinking and analysis skills, collaboration, and communication skills
• Students have increased perspective of how history relates to their daily lives
Students, teachers, participants, users, visitors, and all Minnesotans build an appreciation for history
Diverse individuals, communities, organizations, donors, funders, the Minnesota Department of Education, and the Minnesota Legislature see MHS as a valuable resource for teaching and preserving history and provide ongoing support to the Society to continue its work
Minnesotans and other users of MHS demonstrate that they value history in their personal lives and society at large, and they pursue opportunities to preserve historical evidence and to learn more about history
Teachers integrate MHS resources into the classroom and teach students how to use historical information
Diverse participants, users, and visitors have increased perspective of how the past impacts the present and future
Geographically diverse partners and stakeholders have increased perspective of how the past impacts the present and future, and increased capacity to work in partnership with the Society to preserve and promote access to historic resources

Additional Resources

GENERAL EVALUATION

American Evaluation Association – Online Resources
http://www.eval.org/resources.asp

American Evaluation Association – Online Handbooks and Texts
http://www.eval.org/resources/onlinehbtxt.asp

EvaluATOD – Interactive Evaluation Tutorials
www.evaluATOD.org

EvaluationWiki
http://www.evaluationwiki.org

Wilder Research Evaluation Tips
http://wilder.org/searchresearch.0.html?no_cache=1&tx_ttnews[swords]=tips2008

MORE ON LOGIC MODELS

University of Idaho – Logic Models and How to Build Them
http://www.uidaho.edu/extension/LogicModel.pdf

InSites – Everything You Want To Know about Logic Models
http://www.insites.org/documents/logmod.htm

MUSEUM EVALUATION

Inspiring Learning for All – Research Methods and Guidelines
http://www.inspiringlearningforall.gov.uk/resources/research.html

Inspiring Learning for All – Resources
http://www.inspiringlearningforall.gov.uk/resources/resources.html

Institute for Learning Innovation – Resources
http://www.ilinet.org/display/Resources/Home

Randi Korn and Associates, Inc – Resources
http://www.randikorn.com/resources

Visitor Studies Association – Resources
http://www.visitorstudies.org/resources

Visitor Studies Group – Resources
http://www.visitors.org.uk/node/33/

Minnesota Historical Society institutional Dashboard

The Minnesota Historical Society has committed to annually reporting progress and outcomes on key measures of success. These measures directly correspond to the organizational logic model.

OUTPUTS

The first set of indicators are "outputs" or counts of activities or things you did. Numbers to be obtained from administrative records and/or program tracking:

# of collection units acquired; preserved; digitized, exhibited, or made accessible; and/or used
and amount of grants given by MHS
and amount of grants received by MHS
of public programs offered
# of publications
revenue of MHS (stores, publications, sites, and museums)
# of heritage entities served
# of partners
# of members (including % who are new members and % who are retained members)
# of volunteers (including % who are new volunteers and % who are retained volunteers)
# of interns/fellows
# of users (including % who are new users and % who are repeat users) • Total • Students • Teachers • Seniors • Families • Young Adults • Researchers and professional historians • Schools and school districts (public, private, and charter)
# of web visitors (including % who are new visitors and % who are repeat visitors)
# of social media followers (Facebook, Twitter, Pinterest)

Next, the following indicators directly correspond to the outcomes in the organizational logic model and will be measured with user ratings (surveys, etc.). All programs that serve end-users should include the required indicators on the list below in their evaluation plan. Programs that have ongoing involvement with users (i.e., more than a one-time visit to a site or exhibit) should include one or more of the optional outcomes, depending on the program logic model.

Required Outcomes: Impact of MHS programs on users/visitors/participants—to be measured with surveys

POSITIVE EXPERIENCE: User/visitor/participant rating of their overall experience

Q) Overall, how would you rate your experience with [INSERT NAME OF PROGRAM]?

❑ Excellent ❑ Very good ❑ Good ❑ Fair ❑ Poor

INTEREST IN HISTORY: User/visitor/participant rating of their interest in history (use both questions)

Q How would you rate your interest in history before you (attended/participated in) [INSERT NAME OF PROGRAM]?

❑ Extremely interested ❑ Very interested ❑ Somewhat interested ❑ Not too interested ❑ Not at all interested

Q) How did your (participation in/attendance at) [INSERT NAME OF PROGRAM] affect your interest in history (if at all)?

❑ Much more interested ❑ More interested ❑ Slightly more interested ❑ Not more or less interested ❑ Less interested

NET PROMOTER SCORE: Likelihood users/visitors/participants will recommend MHS to their friends and family (score is computed by counting those who rate a 9 or 10 ("promoters") minus those who rate a 6 or lower ("detractors")

Q) How likely are you to recommend [INSERT NAME OF PROGRAM] to your friends and family?

Not at all Likely 0 1 2 3 4 5 6 7 8 9 10 Extremely Likely

Optional Outcomes: Impact of MHS programs on users/visitors/participants—to be measured with surveys

ORGANIZATIONAL VALUE: User/visitor/participant rating of the Minnesota Historical Society as a valuable resource to Minnesota

Q) How would you rate the value of Minnesota Historical Society as a resource for you?

❑ Excellent ❑ Very good ❑ Good ❑ Fair ❑ Poor

PROGRAM RELEVANCE: User/visitor/participant rating of the extent to which the program connected the past to their present and future

Q) To what extent did [NSERT NAME OF PROGRAM] connect history to things that are relevant for your life today and in the future?

❑ A great deal ❑ Quite a bit ❑ Some ❑ A little ❑ Not at all

ONGOING LEARNING: User/visitor/participant plans to pursue other opportunities at the Minnesota Historical Society to learn about history

Q) How likely are you to visit or use any of the following Minnesota Historical Society resources in the next 12 months? Check one box for each line.

	LIKELIHOOD YOU WILL USE IN THE NEXT 12 MONTHS				
MHS RESOURCE	Extremely likely	Very likely	Somewhat likely	Not too likely	Not at all likely
Public programs					
Historic sites					
Museums and exhibits					
Library and archives					
Web site					

In addition, there are a few indicators that are only relevant for certain user groups/stakeholders. Programs that target these user groups should include these indicators in their evaluation plan.

TEACHERS: View the Minnesota Historical Society as a valuable resource for teaching (use both questions)

Q) How would you rate the value of the Minnesota Historical Society as a resource for teaching?

❏ Extremely valuable ❏ Very valuable ❏ Somewhat valuable ❏ Not too valuable ❏ Not at all valuable

Q) How likely are you to recommend the Minnesota Historical Society to other teachers you know as a resource for teaching?

Not at all Likely 0 1 2 3 4 5 6 7 8 9 10 Extremely Likely

RESEARCHERS AND PROFESSIONAL HISTORIANS: View the Minnesota Historical Society as a valuable resource for conducting historical research

Q) How would you rate the value of the Minnesota Historical Society as a resource for conducting research?

❏ Extremely valuable ❏ Very valuable ❏ Somewhat valuable ❏ Not too valuable ❏ Not at all valuable

Q) How likely are you to recommend the Minnesota Historical Society to others in your field as a resource for conducting research?

Not at all Likely 0 1 2 3 4 5 6 7 8 9 10 Extremely Likely

PARTNER ORGANIZATIONS: View the Minnesota Historical Society as an effective partner in preserving, protecting, and promoting historic information, artifacts, and places (use both questions)

Q) How would you rate the value of the Minnesota Historical Society as a partner?

❏ Extremely valuable ❏ Very valuable ❏ Somewhat valuable ❏ Not too valuable ❏ Not at all valuable

Q) How likely are you to recommend to other organizations like yours that they partner with the Minnesota Historical Society?

Not at all Likely 0 1 2 3 4 5 6 7 8 9 10 Extremely Likely

Minnesota
Historical Society