Best Practices Guide to Program Evaluation

For Aquatic Educators

TAKE ME FISHING™
takemefishing.org

Recreational Boating & Fishing Foundation

Best Practices Guide to Program Evaluation

For Aquatic Educators

TAKE ME FISHING™
takemefishing.org

Recreational Boating & Fishing Foundation

500 Montgomery Street, Suite 300, Alexandria, VA 22314-1657
703/519-0013

Reprinted March 2011

Acknowledgments

The Recreational Boating & Fishing Foundation (RBFF) would like to thank several groups for contributing their time and expertise to the development of the Evaluation Guide.

Original Authors

The following people were provided the writing and editing for this guide:

Susan Marynowski, Lead Author

Christine Denny and **Peter Colverson,** Contributing Authors

Karen Hill, Editor

Review Panel

RBFF recruited practitioners and evaluation experts from state and federal agencies, academia, and the private sector to provide guidance throughout the development of this guide.

Janet Ady, National Conservation Training Center

Tony Fedler, Human Dimensions Consulting

Bill Hastie, David Heil and Associates Inc.

Nancy Herron, Texas Parks and Wildlife

Libby Hopkins, U.S. Fish and Wildlife Service

Tom Marcinkowski, Florida Institute of Technology

Martha Monroe, University of Florida

Jim Stewart, RBFF

Kristen Chambers, RBFF

Lauren Tyler, Florida Aquarium

Chad Tussing, South Dakota Game Fish and Parks

Michaela Zint, University of Michigan

Case Study Contributions

Practitioners from across the country were asked to submit case study information for the guide. RBFF appreciates the contributions made by various individuals and organizations in providing the authors with quotes, stories, and program information. Case study sources are noted throughout the guide.

Practitioner Interviews

Interviews were conducted with practitioners prior to beginning work on the guide. These interviews provided insights into what the Evaluation Guide should include.

Sam Faith, Missouri Department of Conservation

Laurel Anders, Pennsylvania Fish and Boat Commission

Barb Gigar, Iowa Department of Natural Resources

Dave Hagengruber, Montana Fish, Wildlife, and Parks

BeBe Harrison, South Carolina Department of Natural Resources

Bob Jones, Department of Natural Resources and Environmental Control

Paul Thomas, Florida Fish and Wildlife Conservation Commission

Doug Thornburg, Arizona Game and Fish

Kim Toulouse, Nevada Department of Wildlife

Amanda Wuestefeld, Indiana Department of Natural Resources

Contributing Literature

A diverse collection of evaluation literature was used in the research for this guide. The authors acknowledge the great contribution this body of literature made to the guide. The Resources section includes this literature.

About the Evaluation Guide

The Evaluation Guide is designed to assist practitioners of aquatic education programs with all levels of evaluation.

About the Evaluation Guide

The Evaluation Guide was developed as a companion to the *Best Practices Workbook for Boating, Fishing and Aquatic Resources Stewardship Education* and was developed to provide a thoughtful introduction to evaluation. The guide has benefited from the input of more than two dozen evaluation experts and aquatic education practitioners.

About RBFF

The mission of the Recreational Boating & Fishing Foundation is to implement an informed, consensus-based national outreach strategy that will increase participation in recreational angling and boating and thereby increase public awareness and appreciation of the need to protect, conserve, and restore the nation's aquatic natural resources.

In 2000, RBFF's task force on education developed guidelines for research-based boating and fishing education programs utilizing best professional practices to determine which processes provide the best experiences for conveying knowledge, developing skills, and changing attitudes and behaviors. Environmental and outdoor education professionals were commissioned to provide summaries of research and recommendations for the development of best practices for fishing, boating and aquatic stewardship education. This group also recommended basic practices for program planning, development, and implementation, professional development, program evaluation, and educational program research. The result of this collaboration is the 180-page *Best Practices Workbook for Boating, Fishing and Aquatic Resources Stewardship Education,* including a summary report and 11 supporting papers.

The Best Practices Workbook for Boating, Fishing and Aquatic Resources Stewardship Education and other related materials can be downloaded at http://www.rbff.org.

For more information about RBFF, go to www.rbff.org.

Disclaimer

Participation by Division of Federal Assistance staff and/or other U.S. Fish and Wildlife Service staff is in the spirit of collaboration and does not constitute an endorsement of this product by the Service or the Division of Federal Assistance.

Table of Contents

Chapter 3. Plan the Evaluation

Chapter 4. Design and Manage the Evaluation

Chapter 5. Create Useful Results from the Data

Chapter 6. Aquatic Education Evaluation Tools

Chapter 6. Aquatic Education Evaluation Tools *(continued)*

Preface

Who is Served by the Evaluation Guide?

This guide is intended for aquatic educators and others who design and implement all types of education, outreach, and stewardship programs. Case studies throughout the guide illustrate how evaluation strategies have been implemented in aquatic, boating, fishing, stewardship, and natural resources education programs just like yours from across the country.

What is the Purpose of the Evaluation Guide?

This guide is intended to assist aquatic educators with planning, creating, conducting, and reporting the results of program evaluations. The guide is useful for a range of practitioners. Those with no prior evaluation experience can use the guide to learn about evaluation. Those with experience can use the guide to enhance their current skills and to share evaluation techniques with colleagues and volunteers. Incorporating evaluation into your aquatic education curricula is a necessary component of operating a successful program, and it can be very rewarding. Understanding and implementing the appropriate evaluation strategy will help your program to demonstrate effectiveness and will further ensure its sustainability.

We hope you find this guide useful as you explore ways to enhance and improve your programs.

What is Included in the Evaluation Guide?

The guide includes evaluation methods and practices, as well as the tools you will need to perform your own program evaluations.

It takes you, step by step, through the entire evaluation process, from developing support within your organization and planning your evaluation, to selecting tools, managing the process, and turning evaluation data into results.

A summary of best practices is provided at the end of Chapter 2, 3, 4, and 5. A glossary and list of resources for further information and support are included at the end of the guide.

How to Use the Evaluation Guide

Quick Reference to the Evaluation Process

If you answer YES to any of these questions…	…you need:
Beginning the Evaluation Process	
• Are you completely new to evaluation?	Chapter 1
• Do you need an introduction to evaluation?	Chapter 1
• Do you need help in determining the purpose of your evaluation?	Chapter 2 and 3
• Are you unsure of how to develop support or capacity for evaluation within your organization?	Chapter 2
• Do you want guidance in creating an evaluation plan?	Chapter 3
• Do you need help in using the Logic Model to integrate evaluation throughout your program?	Chapter 3
• Do you need help in reviewing your program goals and objectives?	Chapter 3
• Do you need guidance in selecting appropriate evaluation tools?	Chapter 4 and 6
Managing the Evaluation Process	
• Do you need support for managing the evaluation process?	Chapter 4
• Are you unsure of how to develop support or capacity for evaluation within your organization?	Chapter 2
• Do you need help in using the Logic Model to integrate evaluation throughout your programs?	Chapter 3
• Do you need to work with an outside contractor?	Chapter 4
• Are you looking for detailed descriptions of evaluation tools?	Chapter 6
• Are you wondering how to reach coherent conclusions based on evaluation results?	Chapter 5
Interpreting, Reporting, and Using Evaluation Results	
• Do you need support for collecting and analyzing evaluation data?	Chapter 5
• Are you wondering how to reach coherent conclusions based on evaluation results?	Chapter 5
• Do you need help in developing strong recommended actions based on evaluation results?	Chapter 5
• Are you ready to communicate evaluation results and recommendations to various audiences?	Chapter 5
• Are you wondering how to make your evaluation results useful?	Chapter 5
• Do you know how to monitor changes that follow from the use of evaluation results?	Chapter 5

How to Use the Evaluation Guide *(continued)*

Quick Reference by Program Stage

If you answer YES to any of these questions…	…use this type:
Evaluation for New Programs (Planning or Pilot Testing Stage) • Is this a new program? • Do you need some good ideas for a new program? • Are you wondering what kind of program would best serve a target audience? • Are you wondering what kind of information your audience needs? • Do you want to pilot test a new program idea on a target audience? • Do you want to pilot test an existing program with a new audience or in a new area?	Planning Evaluation★
Evaluation for Existing Programs (Implementation/Delivery Stage) • Does your existing program have a problem that you can't solve? • Are you wondering who is participating in an existing program? • Are you wondering what level of service is being provided by an existing program? • Are you wondering about the results of a preliminary round of your program, say, after the completion of one event or training class? • Has your program just been modified and you want to know how those modifications are working? • Has your program just been adapted for a new audience, a new setting, a new problem, or a new behavior? • Do you want to evaluate the progress of an ongoing program?	Formative Evaluation★★
Evaluation for Long-Term Programs (Ongoing/Results Stage) • Do you have an ongoing program? • Do you want to evaluate the outcomes or impacts of an ongoing program? • Are you wondering how well an ongoing program is meeting objectives?	Summative Evaluation★★★
Evaluation for Completed Programs (Results Stage) • Is your program complete? • Are you wondering if your program is achieving the desired results? • Have you measured the long-term impacts of your program and the progress in meeting the ultimate goal?	Summative Evaluation★★★

★ See Chapter 1 for a discussion of planning, formative, and summative evaluation types.

★★ See Chapter 3 to plan the evaluation (especially the Logic Model).

★★★ See Chapter 4 to select tools according to program stage (especially Table 4.4 and 4.5).

Chapter 1

The Landscape and Language of Evaluation

What is Evaluation?

Evaluation Defined

Evaluation is the systematic collection and scrutiny of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming. As it relates to program management, evaluation involves the collection of data that is then transformed into useful results to inform decisions. It is important to extract the "lessons learned" from experiences so that you can develop solutions to program or organizational problems.

Evaluation can help you determine how well a program is working or whether the results of the program are meeting certain criteria. Evaluation can also be used to gather information to help design and improve programs. In the long run, the evaluation effort will help to improve program operations and outcomes.

The general goal of most evaluations is to provide useful feedback to a variety of audiences, including program staff members, organizational administrators, program participants, sponsors, and other stakeholders. Feedback is considered most useful when it informs decisions pertaining to program improvements, budget, future program design choices, or long-term policy directions.

To perform a systematic evaluation of an aquatic education program, think about your program participants: Who are they, and what do you want them to accomplish as a result of your program? What questions do you want answered by the evaluation? Through the evaluation process, you will decide what information you need to answer those questions. For example, if you are evaluating a teacher training program, you may want to know how teachers are using the materials and what students are learning. You can collect this information with various evaluation tools, such as teacher surveys, classroom observations, or assessments of critical thinking skills.

You also need to think about how evaluation can be integrated throughout the many stages of your aquatic education program, from design to completion. For example, to find out about your audience before designing a "Teen Fishing" program, you can meet with a group of teenagers to gather information about their knowledge levels and previous experience, as well as to identify key motivating factors for their involvement in fishing.

This guide will help you work through all of the steps of evaluation! Evaluation includes a broad spectrum of activities involved in collecting data and transforming it into useful results. The guide explores the various evaluation approaches that can be used with aquatic and natural resource education and outreach programs of all types.

Benefits of Program Evaluation

Before we begin to describe the evaluation process and tools, let's address why you might want to evaluate. In case you, or your supervisors, are not convinced of the potential value of evaluation for your aquatic education programs, this information – along with the information in Chapter 2 – will help you see the value and benefits of evaluation.

Evaluation can determine if a program is meeting its objectives, distinguish the program's outcomes and impacts, and provide concrete information for program improvement.

You can expect the results of your program evaluation to help you:

- Improve program design, implementation, and effectiveness: With evaluation, you can say with confidence that the proposed program changes are based on an unbiased evaluation of actual results and outcomes.
- **Demonstrate your program's support of the organization's mission:** With good evaluation results in hand, administrators will better understand your program's ability to support the organization's mission.

- **Justify the costs of your program:** Only by thorough evaluation and cost-benefit analysis can you make the case of your program's value and challenge budget cuts to it.
- **Determine program strengths and weaknesses:** Evaluation shows you how well you are meeting objectives and the areas that need improvement, so that you can make modifications to improve or retool the program.
- **Measure and explain program performance, outcomes, and impacts:** With your evaluation results in hand, you can explain the program's results (target audience knowledge, attitudes, or behaviors) and impacts (e.g., on natural resources).
- **Reveal program successes to supporters, funders, and stakeholders:** Evaluation will generate the evidence you need to gain more support for the program.
- **Validate or discover effective programming methods:** Evaluation of existing aquatic education programs provides information and ideas for future programming strategies.
- **Share information about what works with colleagues and similar organizations:** Valuable information about program effectiveness can be shared with other public and private organizations. The program can serve as a model for organizations in similar situations.

In summary, programs that incorporate evaluation are enhanced by unbiased information about their design or performance. Evaluation can provide evidence that a program is effective, and demonstrate positive outcomes to funding organizations, administrators, and the community. Evaluation helps improve program effectiveness and creates opportunities to share unbiased information with partner organizations.

Case Study: Documenting Results in Delaware

An example of overall evaluation benefits

Delaware manages a statewide *Adopt-A-Wetland* program that links community groups with local wetlands. The aquatic education staff provides training workshops, technical support, and loaner kits for carrying out activities. A recent annual mailed survey of 32 *Adopt-A-Wetland* groups showed that, in the previous year, the groups:

- had made a total of 872 site visits, with an average 2.5 hours per visit
- had involved 7,695 participants
- had spent their time on a variety of activities to improve and enjoy their wetlands:
  - educational activities (53 percent)
  - clean-up activities (46 percent)
  - recreational activities (38 percent)
  - biological surveys (17 percent)
  - wildlife projects (12 percent)
  - water testing (7 percent)
  - restoration planting projects (2 percent)

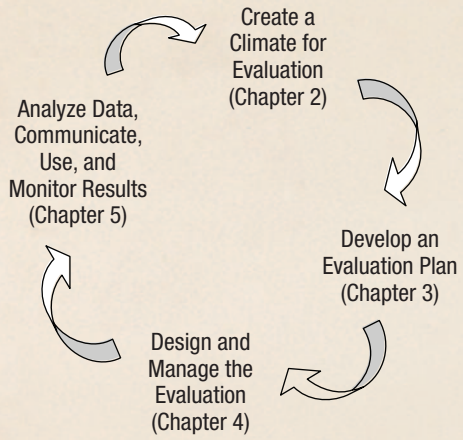
The agency does a survey every year to document both the level of involvement of citizen groups and the actions they have taken to protect and restore their local wetlands.

Source: Gary Kreamer, Delaware Division of Fish and Wildlife

The Evaluation Cycle

There are steps involved in any evaluation process. Each time you move through the evaluation cycle, you end up at the beginning again. There you start over by asking the next set of important questions that will guide the success of your aquatic education programs.

The main idea here is that evaluation is a repetitive feedback cycle, with learning taking place through each repetition. Evaluation can begin at any point in the cycle, depending on the stage of the program and the existing evaluation practices.



Create a Climate for Evaluation (Chapter 2): As you begin to determine the purpose of the evaluation, you will also want to build support for the evaluation within your organization. What information is needed from the evaluation to update and improve the program? What information do you need to collect to build support for your program? How does your program support the organizational mission? How might the evaluation results influence decisions in the organization? Chapter 2 provides ideas to create a welcoming climate for evaluation in your organization.

Develop an Evaluation Plan (Chapter 3): The next stage in the evaluation process is to plan the evaluation, including the development of processes to understand your target audience, the development of meaningful program objectives, and the selection of appropriate indicators to answer your evaluation questions. Of special importance to evaluation planning is use of the Logic Model framework to understand program design and outcomes/impacts. The Logic Model, featured in Chapter 3, allows you to integrate evaluation throughout the life of your program.

Design and Manage the Evaluation (Chapter 4): Once you have an evaluation plan in place, you can begin to manage the evaluation, including the selection of evaluation tools and working with evaluation contractors who can provide critical evaluation support to your program.

Analyze Data, Communicate, Use, and Monitor Results (Chapter 5): Once you collect all of the data, you can develop feedback on your program for interested stakeholders. Chapter 5 describes the process of analyzing data collected during your evaluation and working with statisticians or analysts to obtain the information you need from this data. More importantly, Chapter 5 discusses how to develop coherent conclusions and policy recommendations and communicate your evaluation results to key evaluation audiences. The evaluation results should be used both for program improvement and for organizational growth. Finally, you will monitor implementation of results based on the recommended actions from the evaluation. As you monitor the program and implement policy improvements, you will see the full benefits of your evaluation. Then you can start asking new questions as the evaluation cycle begins again.

Evaluation Basics

There are many different types of evaluation – a different type for each different step of program development and implementation. Evaluation can be integrated throughout the life of a program, or added at any stage in the program.

Planning, Formative, and Summative Evaluation

There are three major types of evaluation: **Planning, Formative,** and **Summative**. You can think of planning evaluation as what takes place before the program is designed, formative evaluation as what provides information to improve the program, and summative evaluation as what measures the effects of the program. The planning, formative, and summative evaluation approaches will reflect the evaluation's purpose, the program needs, and the evaluation questions to be answered.

Planning Evaluation asks "What is needed?"

Planning evaluation collects input and develops guidance before and during the design of an educational program. Planning evaluation considers program goals, objectives, strategies, and timelines. Planning evaluation also:

- asks whether the implementation plans are appropriate, necessary, and/or feasible;
- encourages program revisions before you are committed to the implementation process and allows program revisions if program development begins to diverge from previous plans;
- ensures that all team members, advisers, and stakeholders share a common vision of the program plan and of the evaluation plan;
- establishes the groundwork for future formative and summative evaluations by developing indicators and benchmarks.

Keep It Simple

Evaluation can be simplified and streamlined to best fit your needs and abilities. Look over the tool descriptions in Chapter 6 to find out when each tool should and shouldn't be used. Even something as basic as a program brainstorming session can be reported as a component of the program evaluation.

Types of Planning Evaluation

There are several different types of planning evaluation that you may hear about:

- Context evaluation defines the environment (cultural or physical) in which the program will be presented and diagnoses the problem the program seeks to solve.
- Needs assessment determines the needs of identified audiences or communities and establishes what messages or media might work best to meet the need. Needs assessment often involves focus groups or survey techniques.
- Input evaluation examines the budget, personnel, equipment, facilities, and other resources that are necessary and available to deliver the program.
- Feasibility/market analysis asks if the program is feasible and/or desirable, and whether the available inputs and ideas can be crafted into a real-world program that sells. It also assesses how likely the program is to be successful in light of other providers who may be offering similar programs.
- Baseline study measures the status quo, establishing a benchmark against which to judge future changes or program outcomes and impacts.

Formative Evaluation asks "How can it be made better?"

Formative evaluation keeps a focus on modifying or improving the program prior to and during implementation. In essence, formative evaluation results in information to help "form" a better program. If the formative evaluation is done as the program is being implemented, it is sometimes called midstream, interim, progress, or implementation evaluation. Formative evaluation also:

- examines the program design, technology, delivery, content, personnel, procedures, and inputs;
- helps to define the scope of a program and to identify appropriate goals and objectives;
- can be used to test ideas and strategies before a program is designed;
- can be done while the program is in progress, to determine if the program is on the right track, providing information for fixing weaknesses, correcting shortcomings, or dealing with unforeseen obstacles in program delivery; and
- provides definite information to create a well-designed and well-targeted program from the start.

Types of Formative Evaluation

There are several different types of formative evaluation:

- Pilot testing is a small study carried out prior to a large-scale study to try out a technique or procedure. You may pilot test an evaluation technique, and you may also pilot test a program component.
- Implementation evaluation or process evaluation looks at how the program is implemented and how the implementation processes might be improved. This usually takes place early in program implementation.
- Midstream evaluation takes place when you want to adjust a program that is already underway, and often leads to adjustments in program delivery or design.

Summative Evaluation asks "Are the objectives being met?"

When you look for evidence of the value or success of a program, you engage in summative evaluation. This type of evaluation is undertaken to measure the effects of a program. Summative evaluation is also called product, completion, or final evaluation. Summative evaluation also:

- looks at whether a program is meeting its stated objectives;
- provides information about whether a program reached the intended target audience and whether the participants found the program helpful or useful;
- seeks to determine whether the program itself caused the observed outcomes, whether there were secondary or unexpected program outcomes, and the relative costs and benefits of the program;
- provides ideas for future modifications or improvements in your programs;
- supplies unbiased information for discussing (defending, even!) the impacts, benefits, and cost-effectiveness of your program with administrators, funders, sponsors, community members, and other stakeholders; and "summarizes" a program by describing what happens after delivery of the program.

Types of Summative Evaluation

There are several different types of summative evaluation:

- Output evaluation considers the basic program outputs, such as number of programs delivered, number of program participants, and program costs.
- Outcome evaluation investigates the changes that occur as a result of the program and whether the program is having the intended effect. Outcome evaluation often measures progress toward program objectives, such as changes in knowledge, attitudes, skills, or behaviors.
- Impact evaluation seeks to measure broad and long-term program effects, such as long-term changes (intended or unintended) in ecological, social, economic, or community conditions.
- Cost-benefit analysis addresses questions of program efficiency by measuring outcomes in terms of their dollar costs and values.

Case Study: Redesign of a Watershed Education Program

An example of summative evaluation used to improve future programs

New Hampshire evaluated its Merrimack River Watershed Education Program, which was based on William Stapp's model of students testing water quality, analyzing data, and comparing results at a Student Congress. Teacher participation was dropping off, so the aquatic education staff dedicated one person to individually interview 60 percent of the teachers. They learned that teachers:

- unanimously saw the program fitting in with state curriculum standards;
- needed more flexibility in both the testing activities and training opportunities;
- suggested expanding the Fall testing period and eliminating the Student Congress (it was too hard to take five students out of class and get a substitute teacher);
- needed more support materials for the new parts of the curriculum;
- unanimously supported the organization's effort to make a stronger connection between water quality and wildlife;
- agreed that using fish and wildlife as the link between water quality and land-use practices would help students gain a more concrete understanding of cause-and-effect relationships in the watershed (this focus would also be excellent for teaching river ecology and demonstrating the interdependence of living and non-living parts of the river ecosystem); and
- needed more opportunities for students to actively contribute to resource-related activities in the watershed – to improve actual conditions for fish and wildlife.

The agency used the summative interview results and recommendations as feedback to reconfigure the program into three related but independent modules that teachers can use with their students in local watersheds. Teachers were delighted with the redesign and now tailor the program to fit their needs and constraints, while still helping the agency by teaching key concepts and skills related to watersheds and fisheries habitat conservation.

Source: Laura Ryder, New Hampshire Fish and Game Department

Quantitative and Qualitative Evaluation Data

Two types of evaluation data can be collected: Quantitative and Qualitative. Quantitative data evaluation uses numbers or ratings to define or measure program elements. Qualitative data evaluation uses descriptions and stories to gain deeper insight into a program. The types of quantitative and/or qualitative evaluation approaches you employ depend on the evaluation purpose, program needs, and evaluation questions to be answered. It is important to determine whether numbers or descriptions reflect what is needed from an evaluation. Although some evaluation tasks are better served by one or the other, it is generally best to gather both kinds of data to get more comprehensive answers to your evaluation questions.

Quantitative Evaluation Data

Quantitative evaluation results are expressed as numbers that measure or rate specific program features or outcomes. Quantitative data are useful for comparing or ranking, classifying, and judging the effects of a program. Because quantitative data are numerical, they are sometimes easier to understand and analyze. Quantitative data are most suited to:

- evaluate large-scale programs;
- generalize results to large populations;
- measure levels of knowledge, attitudes, beliefs, perceptions, or change in behavior;
- measure the amplitude of program outcomes or impacts, or causes and effects;
- determine if changes are statistically significant; and
- compare or rank features of various groups.

Quantitative data can be coded, categorized, and statistically analyzed. Statistical analysis of quantitative data provides the evaluation team with an idea of what factors are significant in the success of the program or activity. Evaluation results based on quantitative data may spark the attention and support of certain audiences, since it is often possible to present a graphic representation of the evaluation results. Using strictly quantitative data, however, limits you to numerical data and may hamper your need, to consider in-depth information about program features and impacts.

Tools for Quantitative Data

- surveys/interviews
- skills (performance) assessment
- content analysis
- observations (e.g., counting)
- license sales tracking
- website visit tracking
- stewardship monitoring

Qualitative Evaluation Data

Qualitative research collects detailed, verbal or narrative descriptions of program characteristics, cases, and settings. Qualitative data provide a "story" about the program or event. Qualitative data may be better able to address the question of "why" something happened the way it did. Qualitative evaluations typically use observation, interviews, discussions, and content analysis to collect information. Qualitative methods are appropriate to:

- collect descriptive information;
- understand attitudes, beliefs, and perceptions;
- recognize program outcomes or impacts;
- measure and understand behavioral change;
- judge the nature of causes and effects;
- understand complex issues and program context; and identify unintended or unexpected program outcomes.

Qualitative data can sometimes be time consuming to collect, analyze, and report. Open-ended answers or interviews must be reviewed and reported in a detailed way for them to add value to your evaluation process. Qualitative data may be coded according to topic and then summarized in a quantitative manner, but the information should also be summarized in a qualitative narrative form in the report. The extra attention is rewarded by a deeper and more comprehensive understanding of program outcomes and impacts.

Here is an example of coding qualitative data: A phone survey is conducted for a boater safety program. Several open-ended questions are designed to collect qualitative data. Answers to the question "Name something you learned that you did not know before" are recorded. The resulting qualitative information

provides a rich picture of what audience members learned from the program, some of which is unusual and unexpected. The answers are also coded into several categories. Answers that relate to *learning where to boat* are coded as "1," answers that related to *why boaters should reduce speed in some areas* are coded as "2," and so forth. The answers are presented in both narrative and graphic form in the final report, and several recommendations for program improvement and new program features are based on this qualitative information.

More information on coding and analyzing evaluation results can be found in Chapter 5.

Tools for Qualitative Data

- interviews/surveys (e.g., open-ended questions)
- focus groups or discussions
- brainstorming/nominal group technique
- citizen advisory group/public workshop
- observation techniques
- case study
- expert opinion/Delphi group
- content analysis

Case Study: Evaluating a Summer Teacher Training Course

An example of qualitative evaluation of knowledge and behavior outcomes

Vermont offers a week-long, 3-graduate credit summer teachers' course, Fish and Wildlife Management for Educators, focusing on terrestrial and aquatic ecology, fisheries management, wetlands, and socioeconomic issues affecting the state's natural resources. Besides the usual course satisfaction questionnaire, the agency mailed participants a survey with open-ended questions one month after the course. Results included:

- What are the most important concepts you learned in the course?
- Importance of forest and wildlife management, balancing societal and economic and ecological needs (34 percent).
- Importance of habitats, land and biodiversity for wildlife (25 percent).
- Based on the course presentations, what do you think are the agency's most important responsibilities?
- Fish and wildlife population and habitat management (80 percent).
- Fish and wildlife education (74 percent).

Based on the results, the agency made changes to the curriculum. The agency was also able to show that the training course was meeting its objectives in increasing teacher knowledge of key concepts and encouraging use of the concepts in classroom teaching.

Source: Mark Scott, Vermont Fish and Wildlife Department

Using both Quantitative and Qualitative Data

Due to their open-ended nature, qualitative questions can provide insights into unexpected results or nuances of program impacts. For example, a quantitative analysis (e.g., a limited number of multiple choice categories) might show that program participants are not learning. Qualitative information (e.g., open-ended survey or interview questions) collected at the same time might provide an explanation for that lack of learning, such as miscommunication or misunderstanding. Collection of both qualitative and quantitative data can yield complementary information that both explains and expands on the understanding of evaluation results.

Consistency in Data Collection

Data collection requires attention to detail and consistency. Data collection tools or questions are created ahead of time and relate directly to the evaluation questions in the evaluation plan (see Chapter 3). Surveys or interviews follow a formal set of questions, while focus discussions or observations may follow a "script" of issues for discussion.

In any case, the data must be consistently recorded for each participant, group, or class involved in the evaluation. If the collection of information is inconsistent, wide variability will lead to information that isn't useful. The more consistent the data, the more specific the results of the analysis will be.

Program Outputs, Outcomes, and Impacts

As you might guess, summative evaluation is on the minds of many administrators and budget officers. Administrators often want to know about the outputs or outcomes of a program, and they often want numbers (quantitative data). There is an important distinction to make, however, in the difference between measuring outputs vs. measuring outcomes or impacts.

The problem with measuring only program outputs is that you will not know if the program is successful in meeting its objectives related to changes in participant knowledge or behavior. Outputs – such as the number of programs or the number of participants – do not give any indication of participant attitudes, knowledge, or skills. They give no indication of the potential for participants to develop life-long behavioral changes in support of aquatic resources and systems. Nor do outputs reveal whether a program is resulting in biological changes, such as improved water quality resulting from reduced sources of non-point source pollution in the home landscape.

Only measurement of outcomes or impacts will give you an indication of program success in meeting participant and environmental objectives. For example, changes in participant knowledge and attitudes, participant behaviors, and related biological parameters are important program outcomes or impacts. The best evaluation reports include information about program outcomes and impacts, in addition to program outputs.

No single evaluation tool or type of data is a silver bullet. The approach to gathering information on program outputs, outcomes, and impacts should reflect the evaluation purpose, the program needs, and the evaluation questions to be answered. Each of the different types of evaluation (planning, formative, summative) and data collection

(quantitative, qualitative) have their place in this evaluation process. As you enter evaluation planning in Chapter 3, the evaluation questions will direct you toward measuring program outputs, outcomes, or impacts.

Evaluation Definitions

- **Outputs:** The quantity of products and services delivered by an agency or program to the intended users, such as number of programs, number of participants, geographic area covered, memberships acquired, money earned, etc.
- **Outcomes:** Measurable results or consequences – both expected and unexpected – of an activity or program in meeting its stated goals and objectives, such as the percentage of participants who gain some knowledge or skill as a result of the program.
- **Impacts:** The fundamental intended or unintended change occurring in organizations, communities, or systems as a result of program activities.

Chapter 2

Create a Climate for Evaluation

Basic Ingredients for Institutionalizing Evaluation

The institutionalization of evaluation is the act of making program feedback a part of the organization's standard planning and management practices. It is a fact that evaluation leads to organizational learning, and organizational learning is linked with increased productivity and sustainability for organizations. Aquatic educators need strategies to promote an "evaluation culture" in their organizations so that learning and program improvement can continue.

For many reasons, evaluation is not a standard practice in organizations. Lack of resources is often perceived as a barrier to evaluation in organizations. Evaluation may also be intimidating. What if it turns out that existing programs aren't working? Will the organization or program suffer reduced trust or support? Will jobs or funds be cut? These fears are based on misperceptions, yet these perceptions often create a climate that discourages evaluation.

This chapter focuses on what you can do at your organization to replace a culture of fear with a culture of learning through evaluation. A number of critical factors have been identified for the successful development of an evaluation culture within an organization, including motivation and unified purpose, leadership and teamwork, and capacity and resources. In addition, at the end of this chapter you will find a table of ways to overcome barriers to evaluation in your organization and a list of benefits to organizational culture.

> ***"Evaluation will allow you to know that what you do has value. Engage in finding this out with enthusiasm."***
>
> – Elaine Andrews, University of Wisconsin Extension Service

Motivation and Unified Purpose

Build Team Motivation

The status quo is a powerful force under most circumstances. If your organization is not accustomed to evaluating programs, then why start now? A spark is needed to get evaluation started. Perhaps someone attends a conference or workshop on evaluation. Perhaps your organization hires a new employee who has done evaluation elsewhere. Even more likely, a mandate for increased accountability and evaluation may come from a funding organization or the government.

Research shows that the level of motivation is a key determinant of whether evaluation is successfully integrated into an organization. Motivation is even more important to evaluation success than are other factors, such as staff experience, program size, or infrastructure. Evaluation at your organization starts with you and your staff being sufficiently motivated to commit the time and resources necessary to perform the evaluation. This motivation includes the desire to improve the effectiveness of programs, meet program objectives, strengthen perception of program effectiveness within your organization's mission, and contribute information to support agency decisions or policies.

Ways to build team motivation:

- Define the potential benefits of evaluation and make them widely known to the managers and scientists in your organization who stand to benefit from information provided by the evaluation results. Evaluation is an opportunity for responsible program management. You will not only discover what is and is not working, but you will have the opportunity to improve programs and explore new possibilities for future programs.
- Define the potential benefits to staff members. In many organizations, staff members are better rewarded (e.g., through promotions or travel scholarships) for their contributions to the organization mission rather than for their ability to publish articles or deliver programs. Staff members

will benefit if evaluation demonstrates that the program is contributing to the organization's mission.

- Use enthusiasm and a positive attitude when making an effective case for evaluation in your organization.

Demonstrate Success by Starting Small

If evaluation is not currently established or accepted in your organization, start with one program or program component. Choose a program that has a simple structure and clear objectives, even if the objectives are unwritten. Design a basic evaluation approach and select realistic tools such as observations, informal interviews, or brief surveys. A straightforward start to your evaluation process will give you an introduction to methods, while yielding tangible results to share with administrators to gain support for future evaluation efforts. Your first evaluation will be a success and will help establish the conditions needed to make evaluation a regular part of your work plan.

Example: You may may already be having conversations with program stakeholders. If you consciously design a brief series of questions that you would like to ask these stakeholders, and then hold guided interviews with 10 or 20 stakeholders, you will have some powerful evaluation results to report. In many cases, aquatic educators can simply systematize existing conversations or observations to transform anecdotes into reportable evaluation results.

Establish and Communicate a Unified Evaluation Purpose

The purpose of evaluation can include:

- program development and improvement;
- Accountability;
- validation of how the program supports the organizational mission;
- informing management and policy decisions; and
- organizational and staff learning.

There is overlap among these different evaluation purposes. Some of them concern how you present your results to external audiences, while others help internal audiences learn, collaborate and improve programs.

These different purposes require different approaches and should be considered separately as you discuss your evaluation plan. Evaluation strategies that attempt to serve all purposes at once will run the risk of falling short. Design your evaluation to separately serve each of the important purposes determined for your organization and your programs. Time spent in one or two planning or brainstorming sessions with staff members, managers, scientists, and administrators will allow you to define these broader purposes and also to build a rapport among organizational managers and staff members.

Program Development and Improvement: This is the primary purpose of most evaluations. Program improvement involves the direct application of evaluation results and recommendations to your programs.

Accountability: Take notice of any accountability or reporting required by your organization or of funding partners, and imagine how these requirements can be addressed through your evaluation efforts. In addition to reporting program outputs (e.g., number of programs, number of people served), a well-developed qualitative evaluation (based on informal interviews or observations) is often sufficient to meet reporting requirements.

Validation of How the Program Supports the Organizational Mission: Before you discuss the evaluation, do some background research. Remind yourself of your organization's mission and vision statements, and of any goals and objectives for your section or department. Consider how the proposed evaluation activity can contribute to these values. You may find some ways to adjust your evaluation plans to better contribute to the organization's overall mission or to meet some specific organizational objectives.

"Good evaluation can help with the bottom line – it can help your customers or clients to see that you are effective and doing a good job."

– Brad St. Couer, Harbor Towne Marina, Dania Beach, Florida

Management and Policy Decisions: Although the evaluation may not immediately result in policy changes, educators often find that the results are used to inform later decisions in some very important ways. Administrators need information to make decisions, and evaluation generates information that can be useful in decision-making. Determine what kinds of information would be most useful to the decision makers in your organization. Whatever the level within the organization, staff members who develop leadership skills will earn rewards by providing information that is valuable to the organization. Organizations with active and knowledgeable staff members are recognized for their flexibility, responsiveness and effectiveness.

Organizational and Staff Learning: Organizations need an ongoing flow of information to help staff members adapt to changing situations. If learning is a priority, the emphasis should be on staff participation, stakeholder involvement, and administrator "buy-in." A participatory process creates the space and time for all of the evaluation evidence to be properly discussed and digested. As a result, the organization "learns" and adapts. Evaluation also shows how programs can be adjusted to improve "customer" satisfaction, how you can better empower program staff members to make decisions and solve problems, and how you can design innovative new programs. The evaluation can contribute to an overall shift in your organization toward an "evaluation culture," where staff and administrators become more excited about the potential uses of evaluation results.

> ***"…Senior management needs to champion lesson learning and recognize that [this approach] may mean working in different ways in an organization, including at a senior management level…this approach is well accepted in the corporate sector."***
>
> – Organization for Economic Cooperation and Development (OECD)

Once you make sense of the purposes for the evaluation, you must communicate the evaluation process and purposes to others within your organization. In this case, the Logic Model – introduced in Chapter 3 – can be a guide in communicating your evaluation plans with others in your organization. Managers and scientists in your organization will immediately understand the Logic Model's coherent flow and will recognize the importance of measuring program outcomes and impacts.

Features of an Evaluation Culture

Here are several of the key principles of creating an evaluation culture within an organization:

- Programs that get measured get attention and action.
- Program successes must be measured to learn from them.
- Program problems must be measured to correct them.
- Program results must be demonstrated to win public support.
- Without measures of program results, you can't improve the efficiency, effectiveness, capacity, or quality of programs or organizations.
- Program results must be measured to distinguish success from failure. Otherwise you might unknowingly be continuing to reward failed programs.

Leadership and Teamwork

Build Administrative Support and Leadership

Leadership has been identified as a critical success factor in the institutionalization of evaluation. The aquatic educator will probably be the "team leader" for the process, but evaluation also requires the support of supervisors or administrators. Identify a leader in your organization who understands your evaluation plan and can communicate this vision to others. This leader will encourage you to take risks, will provide rewards for work well done, and will encourage staff members to become empowered through the process.

Management support makes it easier for you to obtain the resources you need to implement your evaluation. This support also helps to persuade members of your organization to use evaluation feedback. It can be a difficult process to convince others that your ideas are worth supporting, but it is necessary if you are to be successful with a new evaluation scheme.

If organization administrators are likely to be skeptical or unsupportive of the evaluation effort, you may need to generate buy-in for your evaluation ideas. The ability to influence people's thoughts and feelings goes a long way toward generating the commitment that you need. If you feel that you are lacking in persuasive skills or competency in generating administrator buy-in, consider attending a professional development seminar or selecting a persuasive member of your staff or department to accompany you in discussions with administrators.

The importance of involving administrators in evaluation is that the level of attention to the evaluation gradually shifts from the program improvement level to the organizational level. The results of the evaluation will certainly lend to program improvements, but they will also contribute valuable information to organizational management and policy, forging the link between program success and organizational success.

Evaluation provides an opportunity to demonstrate the value of your program in the language that your non-education colleagues speak. If you work in a situation where budgets are tight and there is constant pressure to cut education programs in lieu of other organizational programs, there may be important benefits to introducing evaluation. If you can produce or promise a concise report that shows facts, such as *"65 percent of participants in a boater safety program slowed down to within 5 mph of the posted speed limit,"* you may be able to generate additional funding or attention for your program.

Evaluation Tip

To elevate the profile of your evaluation efforts, consider aligning your program with a national program related to sport fishing, boating, fish habitat conservation, aquatic resources, or recreation. This association may bring a more positive light to your program.

Case Study: Budget Pressures at Texas Parks and Wildlife Dept

An example of how evaluation can bring additional support to a program

Facing budget pressures shortly after the millennium, the Texas Parks and Wildlife Department made a decision to evaluate the effectiveness of its educational programs. At a time when keeping quiet might have seemed like prudent advice to avoid the budget axe, evaluation was instead viewed as an opportunity to define and demonstrate education's important role in the agency.

The well-stated objectives and solid evaluation offered a mechanism to "speak the same language" as scientists and administrators who regularly measure effectiveness of resource management techniques. Now, the Texas education programs are viewed as valuable components of agency success, rather than extras that need to be jettisoned during tight budget times. In addition, education staff members are getting the feedback they need to improve their own programs and provide even better service to the citizens of the state. The organization now perceives evaluation as a proactive tool that links education with the conservation efforts of the rest of the agency. The evaluation has helped hold the budget for educational programming at acceptable levels.

Source: Nancy Herron, Texas Parks and Wildlife Department

Involve Administrators, Staff, and Stakeholders

Instituting evaluation is just like adopting any other new technology. If you are a team leader, make sure to get the buy-in of your staff before beginning the evaluation process. In one case where an agency was reluctant to adopt evaluation, an agency administrator said it would have been worthwhile to use a focus group of agency directors to help design the evaluation approach. Organizations that are most successful in developing their evaluation programs are ones that involve the broadest participation of staff members while also keeping mid-level and top-level managers active throughout the process.

Evaluation can be introduced to an organization so that it is seen as a natural extension of and support for what you currently do. The excellent work already being done by you and your staff is enhanced by the opportunity for program improvement stemming from the evaluation process.

Administrators, stakeholders, and staff members involved in the process will feel a sense of ownership

toward the evaluation results and will be more likely to use those results for positive change.

To ensure participation, you will need to devote adequate time and persistence to get staff members on board with your evaluation ideas. Patience is probably the most important virtue for an evaluation team leader! If evaluation is a good idea for the long-term future of your programs and your organization, it will be worth the effort to involve staff members from the start.

> ***"To get started with evaluation, get together with your staff and pose the questions:'What do we need to know to be more effective? How does this relate to the mission of the agency?'"***
>
> – Nancy Herron, Texas Parks and Wildlife Department

Capacity and Resources

Build Organizational and Staff Capacity

It may well be that a major barrier to implementing evaluation at your organization is a lack of evaluation expertise. An early first step is to identify someone as your internal evaluation "expert." This may be you, a member of your staff, or someone else in your organization who can be a close adviser to the process. This person is someone who understands the evaluation process, as opposed to a statistician or analyst. Even if you use an evaluation contractor, you need a staff member to act as the evaluation point person.

To generate an internal expert, start by reading this Evaluation Guide and then attend or send staff members to an evaluation training program. Investing in staff evaluation skills is one of the best long-term investments you can make in program improvement. The National Conservation Training Center offers an evaluation training course (http://training.fws.gov). The course also is available online through the University of Wisconsin-Stevens Points (http://www.uwsp.edu/natres/rwilke/eetap). Evaluation training is often also available through universities and groups that support non-profit organizations.

Guidelines for Involving Administrators and Staff

- Ensure that administrators are involved, adding credibility and continuity to the process. To maintain momentum, administrators must be regularly asked for input, briefed on progress, and provided with interim results.
- Involve program stakeholders in guiding the evaluation process, if appropriate. What would teachers like to know about the impacts of the training program? What do boaters think about your evaluation plan? What can minorities tell you about avoiding bias in programs and in evaluation? Outside stakeholders are best involved in the process as advisers to evaluation design.
- Involve a broad range of staff members from throughout the organization. Involve representatives from education, communications, fiscal, and leadership ranks. Recruit specialists to fill gaps as needed, for example, in scientific or statistical arenas.
- Involve staff members in planning and making decisions about the evaluation. Begin with initial meetings to determine what information staff members need to become more effective. Encourage staff members to "test" and improve evaluation ideas.
- Give staff members plenty of opportunity to voice concerns and have open discussions before the evaluation is launched.
- Offer training and recognition opportunities for staff members involved in the process.
- Maintain workplace morale and trust by emphasizing that the evaluation is focused on program and organizational improvement, not staffing decisions.
- Remain positive and affirming about the potential benefits of evaluation in all discussions and meetings, while also being flexible about the actual scheme for carrying out the work.

A benefit of developing internal expertise is that when you seek outside assistance, you will have someone on staff who speaks "the language of evaluation." It pays to at least become familiar with some of the basic evaluation terminology presented throughout this guide and in the glossary. In addition, you can always ask an evaluation expert to explain terms or results in plain language to assist you in effectively reporting results to your supervisors and participants.

As you become more experienced with evaluation, you will develop a regular working relationship with consultants or university experts. You will almost certainly find these relationships beneficial. Many of your questions will be answered as a courtesy by virtue of the relationship that you have fostered. Bear in mind, of course, that evaluation experts are often busy and overscheduled, so it is always helpful to develop some evaluation expertise among your staff.

As your evaluation capacity increases, you may want to develop a strategy for becoming an organization that regularly refers to evaluation results at all stages of program development and implementation. To develop the vision for a culture of evaluation at your organization, you might pose the following question to yourself: *What would my organization look like if it adopted an "evaluation culture?"* Your strategy should be detailed enough to provide information on what is needed to attain the vision. While this strategy will likely remain informal, it will be a guide for your thinking about what is needed to make evaluation a permanent feature of the organization.

"Our evaluation culture will embrace an action-oriented perspective that actively seeks solutions to problems, trying out tentative ones, weighing the results and consequences of actions … to … encourage innovative approaches at all levels…. In an evaluation culture, we won't act for actions sake – we'll always attempt to assess the effects of our actions."

– William M.K. Trochim, The Center for Social Research Methods, http://www.socialresearchmethods.net

Invest Sufficient Time, Resources, and Budget

It may take more time and resources than you expect to develop a complete program evaluation system. Be sure to carefully estimate how much time and resources the evaluation will require, and add a buffer to cover unexpected events.

Include time estimates for everyone who will be involved in the evaluation. Make sure that staff members are free to devote sufficient time. Staff members need to be provided with the training, time, and incentives to incorporate evaluation into their regular workload. Organizations that find creative ways to integrate evaluation into existing work will make the most progress, whereas organizations in which employees feel overwhelmed will make slower progress.

Think about the resources that you need to carry out the evaluation, such as equipment, computers, printing, or construction of educational offerings like signs or websites. Staff members need to be provided with any specialized equipment that will help them better perform the evaluation task, or that will relieve them of other tasks and free them up to perform evaluation. For example, in one organization where increased workload was a concern, the purchase of automated data scoring and analysis equipment was a good investment because the relatively low up-front cost would pay itself back in years to come through improved staff relations, staff efficiency, and programs, and constant organization improvement.

Get into the habit of building evaluation funding into program budgets. A general rule of thumb is to devote 10-15 percent of the program budget to evaluation efforts. If program funding is sought from an outside organization, then the grant process will likely require evaluation. If it doesn't, put it into the program budget anyway – it will strengthen your proposal. If your funds are internal, it is your job to convince your supervisors that this is a worthwhile expenditure. And by this time, you have already primed your administrators with information about the potential benefits of the program evaluation! The bottom line is that evaluation should be recognized in your annual budget. With the support you have generated among administrators, you will be able to build a strong rationale for the budget request. Be sure to allow adequate time for evaluation to be recognized, funded, and completed – it may take up to a year for the budget cycle to come around.

Link Evaluation to Organizational Planning and Performance Review

Think about how to incorporate evaluation into organizational planning and performance review processes. Every situation is different – you are the expert on the budgeting and planning process in your organization.

During discussions with administrators, you can ask what information is needed to contribute to more effective management decisions. Use this feedback to understand how your evaluation results can inform organizational policy. Discuss this from the earliest stages of evaluation planning to ensure that you collect the right kind of information for the decision-making needs of your organization.

You may also insert the evaluation effort into organizational or statewide planning processes. If your organization has annual retreats or engages in strategic planning exercises, discuss the evaluation scheme. In some state agencies, education programs can be incorporated in future revisions of Comprehensive Wildlife Conservation Strategies (State Wildlife Action Plans), which are designed to include conservation education strategies. Although education is not eligible for funding under the associated grant programs, your program evaluation can gain credibility by being included under the "conservation actions" portion of the plan.

Finally, program evaluation can be linked to employee performance review. This linkage will be an added incentive for administrators and supervisors to lend their support to the evaluation scheme. Note that it is the *successful performance of the evaluation work, not evaluation outcome,* which is attached to the performance review and annual work plan for the employee. In other words, no employee is penalized for the evaluation results and employees are rewarded for successfully carrying out the evaluation work. As long as everyone – both administrators and evaluation team – agrees what work will be done, then it can be incorporated into the annual performance review.

The idea of performance review means that administrators have a reassurance that the work will be done, and the employee has an opportunity for reward. When employee workloads are being negotiated, employee job descriptions can be updated and adjusted to ensure that employees are not overburdened by evaluation tasks. The notion of incorporating evaluation into the annual review process recognizes that the administration will provide funding and support for the evaluation, while the team will have the responsibility of performing the evaluation work. Be sure to include yourself in this performance review scheme, so that everyone involved in the evaluation work is united in their desire for improved performance.

Case Study: The Clean Marina Partnership Evaluation

An example of a fully institutionalized and integrated program evaluation process

The Florida Department of Environmental Protection (DEP) is charged with enforcing clean water laws. Part of that responsibility is to issue and monitor permits for marina construction and operation. In partnership with Florida Sea Grant (a federal/state program), the Clean Marina Partnership was created. It has evolved into an effective cooperative program to protect Florida's inshore and inland waterways from chemicals and other forms of pollution through the voluntary implementation of best management practices (BMPs) by marina owners throughout the state.

Evaluation is a fully integrated and institutionalized part of the program. Marinas are recruited to join the program. Through training, new operators become familiar with the BMPs and the tracking process. A leader from an already designated Clean Marina acts as a mentor to the trainee during a trial period when a series of scheduled inspections are carried out by the DEP. When the trial period is satisfactorily completed, the new recruit is awarded the Clean Marina designation and they voluntarily comply with regulations and reporting requirements. The Clean Marina designation is recertified each year to encourage marina operators to "stay with the program." Marinas not in the program, by contrast, are subject to surprise visits by regulators in the more traditional, adversarial relationship between regulators and the regulated.

Clean Marina operators receive market benefits, including discounts on insurance premiums and submerged-lands lease permits from the state. By using inspections and voluntary compliance, and by building evaluation into the program, the partners feel good about their roles and clean water becomes everyone's interest.

Source: Clean Marina Program: http://www.dep.state.fl.us/cleanmarina/

Overcoming Barriers to Evaluation

Table 2.1 includes several ways to overcome barriers to evaluation in your organization.

Table 2.1 Overcoming Barriers to Evaluation

Problem	Potential Solution
My organization is reluctant to accept evaluation.	When institutionalizing evaluation, make sure that everyone is involved in every step of the process. Establishing an evaluation culture will involve many discussions with people in your organization who have a stake in outcomes or a role in crafting budgets. Keep the focus of evaluation on program and organizational improvements.
Evaluation is not accepted at all, despite multiple attempts.	When your attempts at evaluation have not been accepted by your organization, you can still conduct your own evaluation during day-to-day program planning activities. For example, a program brainstorming session can be considered a component of evaluation, or if you are already talking to key stakeholders you can create an instant evaluation by asking each of them similar questions and documenting the discussions. *"If you find yourself in a situation where evaluation … is not valued in your organization, conduct your own evaluation and share the results with your colleagues and supervisor. Demonstrating how evaluation information can be used to improve programs is a great way to encourage 'buy-in' and begin to make evaluation a part of program planning."* – Jan Henderson, Heifer International
Evaluation is feared as a threat to programs.	Emphasize that the focus of evaluation is on program and organizational improvement, rather than rating of "bad" or "good." Use qualitative tools to perform an evaluation without a numerical score, providing an insightful analysis of what is good and what needs improvement in the program. Emphasize the opportunity to improve programs with concrete information, and then make sure that you use the evaluation results to do just that!
Evaluation is feared as a threat to staff members.	Always keep the focus of evaluation on program and organizational improvements. Do not use evaluation as a mechanism for identifying unproductive employees. Contrary to posing a threat, evaluation can offer rewards to staff members who agree to include it in their annual work plan. Use arguments presented earlier in this chapter.
Administrator buy-in is slow.	Slow buy-in often happens when organization leaders are not fully involved in the process. Actively involve administrators in evaluation planning, discussions about evaluation approaches, and decisions about specific outcomes to be measured. Collaboration will avoid the problem of having to later revise evaluation plans.
Evaluation momentum is slow.	Invest intensive time and energy at the beginning of the process to build momentum. Use training and regular meetings to make sure everyone is on the same page. Allow extra hours to work through evaluation design. Keep administrators involved and briefed on progress. If departure of key staff members slows the flow, find someone else to take the open responsibilities and consider training more than one person.

Problem	Potential Solution
Problems occur with the evaluation process.	Follow the guidelines in Chapters 3 and 4 for planning and managing the evaluation. Use the Logic Model to focus on program outcomes and impacts. Use arguments presented earlier in this chapter to convince administrators of the value of measuring outcomes.
Problems occur in the sharing of evaluation results.	This is a major source of frustration for administrators seeking information to guide their decisions. If administrators can't get the necessary feedback, they will be much less likely to support existing or future evaluation efforts. Provide frequent updates to administrators about the evaluation process and results. For example, create a template of the most important outcomes for the administrator that can be updated on a specific schedule. Refer to Chapter 5 to communicate more useful evaluation results.
Evaluation is seen as another "thankless task" by overworked employees.	Planning sessions must include discussions of workloads and infrastructure needed to implement the evaluation. The plan should address whether new staff or equipment will be needed and whether volunteer or student labor could be used for repetitive tasks. These issues and their potential solutions can be openly discussed so that all concerns about evaluation are addressed.
There is insufficient equipment for processing evaluation data.	This can be a serious problem in large evaluations or in small organizations. If the evaluation team is handling the data itself, you may need additional staff or equipment to handle the flow of data. If raw program data are handed over to a separate office in the organization, there has to be a guarantee of when the analyzed data will be returned to the evaluation team. Consider including a line item for outside data analysis services in your evaluation budget.
People in my organization express negative attitudes about evaluation.	"Approach evaluation with a positive attitude. Create and rehearse inspiring statements about the potential for evaluation to provide useful feedback for your program and organization. Highlight that evaluation is a tool for decision making because it provides information that will enable staff to be more effective in accomplishing your organization's mission and goals. Be supportive and patient with reluctant staff members and administrators. Always be positive – 'What's going right?' Avoid the negative – 'What's going wrong?' Use evaluation to explore how you can continually improve." – Mike Spranger, University of Florida Sea Grant

Unexpected Benefits to Organizational Culture

Institutionalizing evaluation within an organization can lead to unexpected benefits, such as increased partnerships, expertise, level of knowledge, and the ability to better adopt new technologies and practices. You may want to mention these potential benefits as you argue for the value of evaluation.

Potential benefits to organizational culture:

- **Staff members at all levels will develop a greater rapport** with each other and with organization administrators. Teamwork relationships will last beyond the evaluation and will benefit everyone at the organization. For example, aquatic educators and biologists within an organization can gain a greater respect and trust for each other after collaborating on the evaluation of an environmental stewardship program.
- **Staff members involved in evaluation will be recognized as experts** by others in the organization, thus providing value to the organization. For example, in one organization, staff members experienced in leading focus groups were then called upon to facilitate staff discussions during a strategic planning retreat.
- **Staff members will develop a higher level of "inquiry mindedness"** that serves them well in other areas of their work. Staff members change their ways of working and look for information, feedback, and "lessons learned" to inform day-to-day decisions.

- **Organizational learning becomes smoother** as members of the evaluation team notice what helps with the adoption of the new practice. Future adoption of new practices or technologies becomes easier. The organization gradually "learns how to learn."
- **The organization evolves an evaluation culture.** Performance measures are more often used throughout the organization to inform decisions and plans. Evaluation recommendations are regularly revisited to guide new ideas or policies. Evaluation results become springboards for discussions about organizational values and mission and to enlighten organizational strategic planning.

Chapter 2 Summary of Best Practices

- Define and communicate the evaluation vision, purpose, and potential benefits to foster understanding, support, and momentum for your efforts.
- Involve staff members, administrators, and program stakeholders throughout the evaluation process.
- Maintain the evaluation focus on program improvement so there is no threat to administrators or individual team members.
- Build team evaluation capacity over time.
- Allocate 10-15 percent of program budgets for evaluation activities.
- Link evaluation to the organization's annual budget requests, planning, and employee review process.
- Emphasize organizational benefits of evaluation, such as improved knowledge and expertise, increased rapport, organizational learning, and better adoption of new technologies.

Chapter 3

Plan the Evaluation

Take Time to Plan

Great programs take time to get established. Most aquatic educators will say that it took a good bit of fine-tuning to create a program that provided the outcomes they wanted. Evaluation can give you the information you need to go from a program idea to exceptional outcomes and impacts. Many studies have shown that evaluation is one of the keys to program success. Thus, evaluation should not be an afterthought, and is best used as an active part of the program from the beginning.

Evaluation planning does more than just give the program manager a framework for measuring program success – it creates a roadmap to guide overall program success.

Evaluation planning provides you with the answers to key questions such as:

- Why are we evaluating?
- Who will be responsible for making decisions based on the results?
- Who will benefit from this evaluation, and in what ways?
- What do we need to know?
- When will the evaluation begin and end?
- Do we have the time, money and staff to complete this evaluation?
- What does program success look like?
- What evaluation questions need to be asked?

Evaluation Tip

Evaluation planning goes hand in hand with program planning, development and implementation. If you are in the planning stage of a new program, you can use these steps to develop the program and the evaluation plan at the same time.

Whatever your evaluation needs, as a program manager you should take time to plan a simple and streamlined process. Effective program evaluation does more than just collect and synthesize data; it allows you to continually learn about and improve programs.

Evaluation Planning Steps

There are a series of steps for planning a program evaluation:

- Step 1: Define the Program
- Step 2: Determine the Evaluation purpose
- Step 3: Understand Similar Programs
- Step 4: Assemble the Evaluation Team
- Step 5: Establish Resource Inputs and Constraints
- Step 6: Create Questions and Select Indicators
- Step 7: Develop the Evaluation Approach

Step 1: Define the Program

The first step in planning your evaluation is to define and understand the program you are focusing on. This can be a simple exercise in writing down basic program information that you already have on record. On the other hand, you may realize that certain elements of your program need to be clarified prior to evaluation.

There are four basic elements to understand about the program before you evaluate it:

Program name and description: Write down a brief description of what the program is all about.

Organizational mission: Review the mission of your organization, and also the mission of any organization that is providing funds to the program. You may have more than one mission to satisfy.

Program goals and objectives: Why was the program created? Write down a sentence about why your program was created, for example: "The Streamside Program was created to increase student awareness of and involvement in stream water quality." One of the most important questions that

aquatic educators ask is: What do I expect to accomplish with this program? This can be rephrased as: What are the program goals and objectives? The goals and objectives will be used to guide program development and to measure program success through evaluation. Most existing programs have some goals and objectives that define what the creators hoped to accomplish. If the goals and objectives have not been written, or if they are too vague, use the information in the Creating Goals and Objectives tip box for assistance in writing updated goals and objectives.

Program target audience: Think about who benefits from this program. The target audience may already be established for existing programs, or you may not yet have considered the audience if the program is in the planning stage.

Evaluation Tip

When writing objectives, use action verbs that describe a desired behavior that you will be able to measure. Here are some measurable action verbs:

Analyze	Demonstrate	Name
Explain	Describe	Define
Investigate	Present	Use
Operate	Perform	Practice
List	Classify	Combine
Evaluate	Match	Construct
Categorize	Teach	Debate
Identify	Sketch	Compare
Discuss	Locate	Estimate
Summarize	Imagine	Produce
Write	Recognize	Solve
Predict	Diagram	Select
Report	Compute	Justify
Apply	Illustrate	Plan

Creating Goals and Objectives

Goals are what you hope to accomplish with the program, or broad statements of what participants will do after the program. Goals also reflect your organization's mission. Many programs have only one goal statement. The goal is then used to trigger potential program ideas.

Example Goal: Protect the water quality of the area by increasing the amount of used oil that is recycled through county waste systems.

Program Ideas:

- Include oil recycling information in high school driver education courses.
- Provide oil recycling information to all people getting driver's licenses in the county.
- Provide oil recycling guides to area mechanics and service stations.

Objectives are specific measures of whether or not the program is successful in addressing the goal. An objective defines precisely what a participant will do, think, or feel as a result of the program. Useful objectives are clear, specific, and measurable. Write objectives that tell how participants will be affected by the program (outcomes, impacts), rather than what will be done (outputs). Program objectives are developed from the program ideas. Most programs have multiple objectives.

The ABCs of Objectives: Writing Good Objectives

A Audience: The objective defines the program audience. Example: Program participants…

B Behavior: The objective describes participants' desired actions. Example: Participants will describe…; Anglers will use…; Boaters will demonstrate…; Teachers will explain…

C Conditions: The objective states when the participant will perform the desired behavior. Example: When fishing…; After the workshop…; During the skills demonstration…

D Degree: The objective describes the degree or criteria for the desired action. Example: Over half of participants…; 20 students…; 80 percent of community members…

Example Objective: After driver education, 75 percent of participating students will be able to explain how and where to recycle their used motor oil.

Audience (Who is the program for?): Students

Behavior (What will the participants do?): Explain how and where to recycle used oil

Condition (When will this be done?): After the driver education course

Degree (To what extent or how many?): 75 percent of students

Source for ABCs: Ricker, et. al. 1998. *Water Quality Project Evaluation: A Handbook for Objectives-Based Evaluation of Water Quality Projects.* Ohio State University Extension.

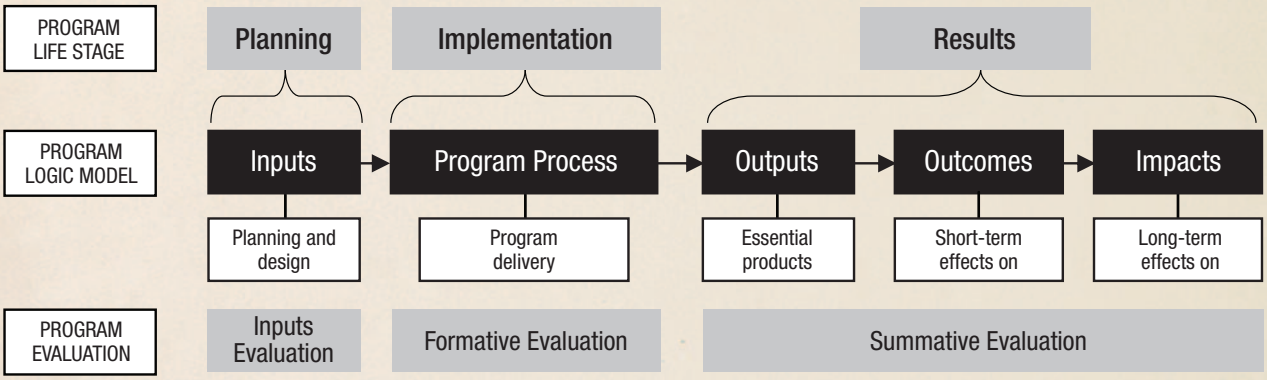
The Logic Model for Program Planning and Evaluation

The Logic Model is a tool to guide program planning and evaluation. Although the name sounds very technical, it is simply a step-by-step guide to how the pieces of your program fit together. As the name implies, it is logical!

The Logic Model encourages aquatic educators to look at the planned work that goes into a program and how it relates to the intended results. The Logic Model can guide the development of your evaluation plan, showing what you put into the program (inputs) and what you intend to get from it (outputs, outcomes, impacts). The Logic Model also clarifies how and when to evaluate the program.

The basic Logic Model components are shown in the figure below, along with indications of the program life stage and types of program evaluation. The program life stages and types of evaluation are further discussed in Chapter 4.

The Logic Model with Program Life Stages and Evaluation Types



Planning Stage: What problems or needs will the program address? What resources will go into planning the program? What form will the program take?

- **Inputs** are resources that go into creating the program such as staff, time, money, materials, and equipment.

Implementation Stage: How will the program be delivered? What media and messages are best? What resources are required for program delivery?

- **Program process** includes the tools, materials, events, actions, technology, and people that are used to produce the program. For a teacher training program, the process might include curriculum design and production, training design and delivery, instructors, volunteers, and teachers who attend the course.

Results Stage: What are the intended results of the program? What are the potential short- and long-term effects? Is the program meeting its objectives?

- **Outputs** are the essential products of a program and include things such as number of sessions, number of participants, program costs, and participant feedback.
- **Outcomes** are the changes that occur in participants as a result of the program. These include changes in knowledge, attitudes, skills, motivations, decisions, and stewardship behavior.
- **Impacts** are the long-term changes in environmental, social, economic, community, or organizational conditions that occur as a result of the program.

Step 2: Determine the Evaluation Purpose

When beginning to plan an evaluation, ask yourself: *Why am I doing this evaluation? What do I hope to achieve with this evaluation?* These questions, and the points below, are necessary to help you think about the purpose of the evaluation:

- **Why are you evaluating?** The evaluation purpose can include: to inform program improvement and development; to determine whether the program meets objectives; to measure whether the program leads to desired participant changes (outcomes, impacts); to show that the program is successful (accountability); to validate the program's role in supporting the organizational mission; to obtain additional funding; to explore useful educational approaches; to provide information for management and policy decisions; to promote organizational and staff learning; or to justify the program to reluctant administrators.
- **What type of evaluation are you conducting?** Evaluation types include planning (conducted during program planning, such as needs assessment), formative (conducted during early program implementation for improvement or modification), or summative (conducted to summarize program outputs, outcomes, and impacts). Refer to the discussion of evaluation basics in Chapter 1.
- **Who will use the evaluation information?** Think about who will be using the information you collect from the evaluation. Audiences for the evaluation include program managers, staff educators, facility directors, fisheries biologists, organization administrators, budget planners, program funders, the public, and others.
- **How will the information be used?** Your evaluation is being conducted to address a need. Take time to think about how your evaluation results will be used. What are you going to do with the information you collect? What decisions are going to be made based on the information? What actions might be taken based on the information? How will the information help you? Evaluation is fundamentally about influence and values within an organization – it is vital to recognize this from the beginning so that you can make the most of the evaluation results. Review Chapter 2 for more information about organizational dynamics and evaluation.

Evaluation Tip

Plan early to make the evaluation results as useful as possible. Think about your key report audiences, what information they need, how you might best communicate with them, and how you can ensure that the results of your evaluation will be used. See Chapter 5 for more guidance.

Step 3: Understand Similar Programs

Existing programs and research can provide a wealth of information to support the evaluation effort. There are two major sources of information that will be useful: 1) programs that are similar to your program, and 2) existing information about your own program.

Perform a library and Internet search for information about similar programs. Find out how other people have evaluated similar programs. If the evaluation results were published or reported, you can ask for a copy of the report or download a copy if it is available on the Internet. You might also review natural resources, aquatic, or environmental education journals at a university library for evaluations of similar programs. Your colleagues around the country may also have information that is helpful to you. This search will turn up ideas that are directly useful in your own program evaluation. You can also cite this information in your evaluation presentations and report to compare the results from similar programs to your own evaluation results.

At the same time, ask yourself what information already exists about your own program. Has the program been running for many years? If so, there may well be some archived information that will be useful. For example, do you have records of programs delivered and participants served for the last 10 years? Do you have samples of student work from a past teacher-training effort? Is there a scrapbook of photographs from a series of fishing training days? All of these sources of information can be directly useful in the present evaluation effort.

Case Study: Improving Bull Trout Conservation Program Success

An example of using similar research for feedback on program design

In 2001, Montana's Department of Fish, Wildlife, and Parks, in partnership with several other groups and agencies, looked for ways to improve an eight-year effort to educate the public about the need to protect bull trout. The species had rapidly declined since receiving the wrongful reputation of being a predator of valuable sport fish. Evaluation showed that the existing program was not having as much of an impact as hoped. Program leaders looked to a successful grizzly bear conservation program for ideas for tools and approaches to public education and evaluation. Feedback from the grizzly bear program was used to revamp the trout conservation program, which took a different form, including self-guided web-based instruction with pre-program and post-program tests. Participants could print out a certificate saying they had passed the bull trout conservation test. Within six months, more than 5,000 people had completed one of the online modules, each with sufficient improvement in knowledge to have earned a certificate.

Source: Janet Ady, U.S. Fish and Wildlife Service

Step 4: Assemble the Evaluation Team

Evaluation team members are those who have a stake in the outcome of the evaluation, and they are also an important audience for the evaluation results. They are the people you will work with to formulate the key questions you want to answer with the evaluation. Consider who will have influence in the definition of your evaluation questions and process. You may be the only person who needs to be involved, or you may require the input of other key stakeholders. Team involvement will ensure that you get different perspectives on evaluation needs and make sure core needs are met.

The team can include: program designers, site managers, education coordinators, biologists, program funders, grant managers, community members, and resource managers or owners. These people can be inside or outside of your organization.

Be sure to clarify the roles and responsibilities of everyone involved in the evaluation process. For example, some team members may be involved in technical or statistical aspects, while other team members may be performing field observations. Some will play a hands-on role in the evaluation, some might act as advisers, and some will make decisions based on the evaluation results. Describe your expectations for what each person will accomplish and circulate them for comments.

Ask team members what kinds of information they want and need, and discuss how the evaluation will fit within the broader purposes of the organization. Continue to involve administrators and staff members throughout the evaluation process. Set a regular schedule of evaluation meetings to take place throughout the process. You might have different team members attending different meetings, so make this clear from the beginning. Involve

administrative advisers in only a few key meetings to incorporate their input and address their concerns.

Involving team members will help you:

- collect information that is important to key people involved in the program;
- ensure that you do not miss collecting critical pieces of information;
- increase program support as team members develop a better working knowledge of the program; and
- increase support for and use of the results of the evaluation.

By involving the team when you begin to plan your evaluation, you will avoid going through the evaluation process only to hear, "You've collected interesting information, but what about…." Involving key stakeholders will improve your program's credibility and will lead to a joint understanding of the process and its outcomes. The understanding that is fostered by the team dialogue will help you gain support for your work and increase opportunities for evaluation results to be taken into account when future decisions are made.

Shape a Useful Evaluation

One of the most important uses of evaluation information is to inform decision-making processes, especially decisions about program continuation, enhancement, and funding. The list below contains questions that the evaluation team members can ask themselves or their intended audiences to shape the evaluation:

- What decisions, if any, are the evaluation findings expected to influence? Are these decisions primarily summative (about program funding, continuation, or expansion) or formative (about program improvement and development)?
- What data and findings are needed to support the decision-making process?
- By whom will these decisions be made?
- When will these decisions be made? When must the evaluation findings be presented to be timely and influential?
- Who are we trying to influence? What needs to be done to achieve that level of influence? Which audiences will have the expected degree of influence?
- How much influence do we realistically expect the evaluation results to have? What are people willing to change about the program?
- How will we know afterward if the evaluation was used as intended?
- What other factors (values, politics, issues, personalities, promises made) might affect the decisions or make the evaluation information irrelevant? To what extent has the outcome of the decision already been determined?
- If the evaluation is not going to inform any immediate decisions, how do you expect the evaluation information to be used in the future?

Step 5: Establish Resource Inputs and Constraints

Take some time to think about the needed resource inputs and constraints for the evaluation you wish to conduct. Given time and funding constraints, you may not be able to evaluate all aspects of your program, but perhaps evaluating a key part of it will produce tangible benefits. Consider which program objectives can be best addressed by the evaluation, and what kind of information your various evaluation audiences need.

In light of the objective(s) that you are going to evaluate, review the resource inputs that you currently have. Establish what additional resource inputs you may need and how you might meet those needs. Much of this process will depend on the information you want to collect (Step 6) and the methods and tools you use (Step 7). (You can come back to this step as you decide what tools will be used in the evaluation.) Be realistic about what you can do in relation to what you want to do. You need to balance practicality with scientific rigor. Your evaluation will go much more smoothly if you have the appropriate levels of time, money, and staff budgeted toward the evaluation effort.

Time: Effective evaluation is worth the hours you put into it. Include a time budget in your evaluation plan. How much time will it take to design and implement your desired evaluation? There are no guidelines for how long an evaluation will take, but the better your objectives and your evaluation plan, the more streamlined your evaluation can be. A simple evaluation of a single event could be accomplished in a matter of days or weeks, while an evaluation of a large program might take a year to complete, with evaluation work coming in short bursts throughout the timeline.

Do you or your staff have this time available? Include in your planning the time needed to analyze and report on the data you collect, which will take at least as long as planning and collecting the data. Remember to include time for follow up and application of evaluation results – the most thorough evaluation will not help your program if the results remain in a folder on a shelf!

Staff: Who will conduct this evaluation? Have they done this before? If this is the first time you or your colleagues have conducted an evaluation, build in time for the learning curve. If your staff can take the time to learn about evaluation techniques through workshops or training sessions, your programs will benefit in the long run.

Evaluation Tip

If no one on your staff has the time, interest, or ability to implement the evaluation, consider outside assistance or evaluation partners. Trained evaluators can perform the evaluation for you or act as advisers to your program evaluation. They can also design tools, collect and analyze data that best fit your needs, and help interpret results into meaningful and useful recommendations for program improvement. Their expertise lends itself to efficient evaluations that best meet the needs of your program. See Chapter 4 for more information on outside assistance.

Materials and Equipment: Do you really need equipment for an evaluation? You may need more than pencils and paper, and at the very least you likely will need access to a computer to enter the information you collect and develop your evaluation report. Depending on the evaluation tools you use (see Chapter 6 for details on different tools), you may require materials such as statistical analysis software, audio or video recorders to capture discussions during focus groups, access to the Internet to create and implement a web-based survey, or access to a library or archive of documents. Consider what you will need to collect information as well as what you will need to analyze it.

Funding: Make sure you are aware of the funding required to conduct your evaluation. Review your budget if one is already in place. Consider getting a grant to support your evaluation plan. If you are in the planning stages of a program, evaluation will be included in the preliminary budget request. Many organizations and granting agencies consider it appropriate to assign 10-15 percent of the total program budget to evaluation.

Use Existing Data

If you have a limited budget, a cost-effective source of planning information is through existing data sources. Data may come from attendance sheets, course registrations, website users, the U.S. Census, boater registrations, fishing licenses, or other sources. Existing data may reveal information about your target audience, such as who attends public programs provided by your agency, which part of the state is growing most rapidly, where different ethnic groups reside, where the majority of boats are registered, which public boat ramp receives the most use, etc. This information can create a backdrop for evaluation and may help you determine some target audience needs. Drawbacks to using existing data are that it might be inconsistent across datasets or difficult to access. Although some records may be publicly available, there may be barriers to obtaining information protected by privacy laws.

Evaluation Inputs Planning Worksheet

Resources	What you have	What you need	How to get what you need or work within resource limitations
Time			
Staff			
Materials and equipment			
Funding			

Step 6: Create Questions and Select Indicators

Determining your critical information needs will lay the groundwork for your evaluation. What is the purpose of your evaluation? What do you need to know for program improvement? What information is needed by the key evaluation audiences? These three facets will help you to determine what the central questions of your evaluation should be.

These "evaluation questions" are the general issues that you want the evaluation to address (not the specific questions that would be used for a survey or interview). These questions represent the "big picture" information that is needed from the evaluation. Refer to the Logic Model to make sure that your questions address all aspects of the program (e.g., inputs, process, and outputs).

The most important evaluation questions will emerge based on a discussion of the evaluation purpose and the program goals and objectives. Work with the evaluation team to brainstorm a list of critical questions that need to be answered. Consolidate the questions, and then determine their importance by asking "What decisions or actions can I take based on these data?" Be careful not to waste time collecting superfluous data. If you try to address too many questions, the evaluation will end up being scattered and the important results will be buried in mountains of information.

Less is Best

Fewer questions are better for a focused evaluation. Limit your evaluation questions to address only a few program objectives, features, or information needs.

Real-World Evaluation Questions

- Does the program fit with the organization's mission and values?
- Is the program needed? Does it address an important problem?
- Does the program reach the appropriate target audience(s)?
- What information or instruction is needed by the target audience(s)?
- Are program delivery methods working?
- How do program participants relate to our subject?
- Which messages and media connect with program participants?
- Does our program meet its goals and objectives?
- Is our program effective in achieving the desired outcomes?
- Is the program addressing the problem it was designed to address?
- How much does the program cost relative to its effectiveness?
- Does the program continue to be feasible in the face of changing resources or policies?
- What lessons can be learned from the evaluation of this program?
- What program feedback would be useful to organizational decision makers?

When you are creating evaluation questions:

- **Decide which information is most important to collect.** The evaluation questions will address program objectives, the evaluation purpose, or audience information needs. Consider what information would be most important, meaningful, and practical to evaluate. If you have unclear program objectives, take a few minutes to clarify them before embarking on determining the evaluation questions. In particular, determine what information will matter most to your program stakeholders and make the evaluation results most useful for your program and your organization. Here are a few potential sources to get ideas of program information needs:
  - Ask program stakeholders and past participants what information is needed.
  - Ask program staff members where the program differs from the program vision or where the program goals and objectives are not aligned with those of the organization.
  - Look at past program evaluations for information gaps.

- Review program objectives for measurable benchmarks.
- Review the Logic Model to see if you need information on inputs, outputs, or outcomes.
- Look to address organizational accountability requirements.
- Ask administrators what information is needed for decision making.

- **Ask evaluation questions.** Write down the potential evaluation questions that relate to the important information needed. In the beginning, write down all of the evaluation questions you and your team have thought about. This may be a lengthy list, but keep all of the ideas for now.
- **Select potential indicators.** How will you know when you answer these questions? Indicators are the pieces of information that let you know when your evaluation questions have been answered. Indicators are often the outputs and outcomes of the program. For each indicator, you will also have a source of information. For example, an indicator of educational program success might be improved participant knowledge. The source of information would be a program survey or interview. Indicators can also be long-term impacts that might be expected from program success. For example, an indicator of improved stream conservation behavior might be bank erosion. The source of information might be photographs taken at established points every month for a year after the stream conservation program. Add a list of potential indicators to your list of important evaluation questions.
- **Identify potential sources of information.** After defining evaluation questions and indicators, you need to establish where your information will come from. Who or what you are going to evaluate will depend largely on the type of evaluation you are conducting. Some evaluations involve feedback from a variety of audiences such as program participants, group leaders, teachers, recreationists at an outdoor site, program staff, and/or site supervisors. Other evaluations involve observation or measurement of changes in people or natural resources, or internal processes such as content analysis or brainstorming.
- **Narrow the list to the most important evaluation questions.** Examine each question and the associated indicators to see how it meets your evaluation needs. Each question should measure program objectives, provide information for program improvement, or address evaluation audience information needs. Look closely at resources and constraints and determine which of the indicators can reasonably be measured. Meet with the team to discuss the feasibility of what you intend to evaluate and to establish the importance of the various evaluation questions.
- **Make sure the indicators can be measured.** A final step is to make sure that it is possible and feasible to measure the potential indicators. If you are unfamiliar with evaluation tools and data collection, refer to Chapter 5 to learn more about data collection and analysis. With the proper questions and indicators, you will be able to collect the right data for your particular evaluation and audience information needs.
- **Make sure your objectives are realistic.** Let your experience or the results of similar programs be a guide for expected improvements as a result of your program.

Evaluation Questions Planning Worksheet

Planning Step	Example	Your Program
Program goal	To increase public awareness of watershed conservation.	
Program objective(s)	• *Following the program, 50 percent of participants will be able to name three watershed conservation practices.* • *Following the program, 75 percent of participants will indicate a willingness to engage in one watershed protective behavior.*	
Evaluation purpose(s)		
• To demonstrate program success.	• To provide information for organization's new watershed initiative.	
Evaluation audience information needs	• Program staff members need feedback to improve program delivery. • Administrators need information to direct future watershed conservation efforts in both biological and education realms.	
Evaluation questions	• Are program participants gaining awareness of watershed conservation needs? • What feedback do participants have for program delivery improvements? • What level of interest do participants have in watershed conservation?	
Potential indicators		
• Increased participant knowledge and awareness following program.	• Participant feedback following program. • Levels of participation in key watersheds.	
Source of information	• Pre-program and post-program participant interviews. • Post-program feedback cards. • Program registrations.	
Potential use(s) of evaluation results		
• Demonstrate program success.	• Inform improvement of program content and delivery. • Guide organizational conservation efforts in key watersheds.	

Case Study: OMC Foundation Grant Program Guidelines

An example of the kind of program planning that is required to get a grant

Guidelines for the Take 'Em Boating Grant Program.

Please complete a Program Description by addressing the following items:

1. Program title.
2. Who are the instructors and what are their credentials?
3. List the type of boats to be used in the program.
4. Provide the name(s) of the body of water where the program is conducted.
5. List the goals and objectives of the program.
6. How does the program promote the future of the marine industry through environmental education and safety awareness?
7. What age group is the program targeting?
8. Describe how your program is implemented or delivered to the student.
9. How is your program advertised or promoted (if applicable)?
10. Describe how your program can be adapted by other groups.
11. What results have you documented that indicate the program is successful?
12. In general, describe the areas of your program that would be enhanced with this grant. Feel free to include course outlines, lesson plans, schedules, budgets, equipment needs or any other supporting information that will help explain your program.

Source: National Safe Boating Council, http://www.safeboatingcouncil.org/awards/OmcGrantProgram.htm

Step 7: Develop the Evaluation Approach

The final evaluation planning step is to select one or more evaluation approaches and to develop an evaluation plan that will help you gather the information needed to answer the evaluation questions.

Evaluation Time Frame

Evaluation should be infused throughout the program process, not supplemental to it. Correlate the evaluation timeline with the stages of your education or outreach program so evaluation synchronizes with planning, implementation, and results. Important opportunities can be lost if evaluation is not coordinated with all program elements, such as assessing needs during program planning, collecting baseline information before program implementation, adjusting an ongoing program, or evaluating outcomes and impacts. This also highlights the importance of beginning evaluation planning as early in the process as possible.

Appropriate Level of Effort

One of the most difficult tasks will be to establish the appropriate level of effort for your evaluation. You must balance your need for information with your abilities and resources to perform the evaluation. While some aquatic education programmers want to do evaluation research for scientific publication, most others want a more straightforward way to measure program outcomes and successes. Most educators do not have the time or resources to devote to a full-blown experimental evaluation design, with a control group and pre-program and post-program data collection. For most educators, it is best to keep the evaluation design simple, evaluate just one or a few key aspects of the program, choose the tools with which you are most comfortable, and allow yourself plenty of time to complete the evaluation process.

Evaluation Tip

Evaluation should be infused throughout a program's life, to take advantage of multiple opportunities to measure needs, progress, and results.

Evaluation Methods and Tools

What tools will you use to gather the information you need? There are two main types of data that can be collected: qualitative and quantitative. Qualitative information is narrative and descriptive in nature. Focus groups, informal interviews, and case studies are examples of tools that can be used to collect qualitative information. Quantitative information can be described by numerical data. Surveys, telephone interviews, and skills assessments are tools that can be used to collect quantitative data. The tools you will use depend largely upon the evaluation questions you ask, the information you need to collect, the size of your evaluation task, and the resources available to you. In selecting tools, you should refer to the Logic Model outline for your evaluation effort. Methods and tools are covered in depth in Chapters 4 and 6.

Evaluation Tip

Use more than one evaluation tool whenever possible to "triangulate" or confirm your results from different perspectives. See Chapter 4 for more on triangulation.

Sampling Design

Sampling is a way to obtain information about a large group by examining a smaller, randomly chosen selection ("sample") of group members. If sampling is correctly conducted, the results are representative of the larger population as a whole.

Sampling Size: With small groups and events, your sample can be as simple as evaluating the entire group of people who participated in a program – a 100 percent sample is possible. If the population is large (e.g., "1,000 participants in a popular education program" or "all of the licensed boat owners in my state"), the sample may be only a small percentage (often 5-7 percent) of the group.

For medium and small populations, one rule of thumb is to sample at least 30 individuals from the overall population, and to sample at least 30 individuals per group if you want to compare different groups to each other.

For even the largest populations, a survey sample size of 411 people is generally enough to represent them. Samples of around 200 usually suffice for most statewide surveys with a 5 percent margin of error. Smaller samples will serve for smaller populations. Larger samples may be needed for evaluations that seek to make comparisons between groups, such as gender or racial groups, or groups with different learning outcomes.

An evaluation expert can help you determine the necessary sample size for your population, depending on the desired margin of error. A 4-5 percent margin of error is considered to be acceptable for most evaluation efforts.

Random Sample: A random sample is defined as a sample of a population where each member of the population has an equal chance of being in the sample. Evaluators use a random number generator or random selection software to derive a random selection from a list of names or telephone numbers. If you are working with an evaluation contractor, they can perform a random selection for you.

In some cases, the "members" of your population might actually be "activities" or "events" or any other item you could evaluate. For example, imagine that your organization has delivered 100 teacher training workshops around the state. For the evaluation, you could take a random sample of those workshops (say 10 workshops) and use that sample for your evaluation, instead of studying each and every workshop that was presented.

Sampling Explained

How can a sample of people represent everyone? If done properly, it is an established fact that a random sample can represent a larger whole. The U.S. Census is one example of how a sample can be used to represent the population of an entire nation. Taking a small random sample is much more efficient and consistent than trying to interview everybody in the population! Think about it: a doctor takes only a small sample of blood for a blood test – she doesn't need to take it all.

Systematic Sample: There are also easier sampling schemes that can mimic a random sample. For example, a systematic sample is considered to be equivalent to a random sample. In a systematic

sample, you take every "nth" record from the population. For example, if you need to sample 20 people from a population of 100, you would begin at a random point and take every 5th record from an alphabetical list of population members. If you needed to take a sample from the telephone book for a small community, you would begin at a random point and take every 20th name to derive a 5 percent sample of the population.

Stratified Sample: Another important feature of sampling is that it can help you give equal attention to subpopulations within your overall population. Stratified sampling is a method of dividing the population into subpopulations (strata) and taking random samples from each subpopulation (stratum). For example, imagine that your organization has delivered programs in a number of large urban cities and in a number of small rural towns. The final list of participants shows that 80 percent of your participants were from the cities and only 20 percent were from the small towns. In the final evaluation, you want to compare the responses of rural and urban residents in these different areas. You would divide your list of participants by city/town and then take an equal sample of 30 people from each city/town. This will give you enough people in the final analysis to compare the results between individual towns and cities, as well as to compare overall results of rural vs. urban residents.

Stratified sampling can be used to ensure that you get a large enough sample from any minority group in your population, whether that minority is based on geography, race, gender, education, income, skills, background, or any other factor. *The only caution with a stratified sample is that the overall results are not representative of the population as a whole, because you have intentionally taken a larger sample of the smaller groups.* If you wish to report overall results as well as making comparisons, an evaluation expert can help you "weight" the data so that you can make an accurate reporting of overall results in addition to making your group comparisons.

Sample of Convenience: Many evaluators select respondents for their samples because they are readily available. This type of sample is called a sample of convenience. You should be aware that respondents who volunteer for a study or who are more readily available may have certain levels of characteristics – such as ability, motivation, or attitudes – that make them a group that is different from the general population. For example, many web surveys take samples of convenience. While the results from such a sample may be interesting, they cannot be generalized to a larger population.

Chapter 3 Summary of Best Practices

Use the series of steps for planning a program evaluation:

- Step 1: Define the Program
- Step 2: Determine the Evaluation Purpose
- Step 3: Understand Similar Programs
- Step 4: Assemble the Evaluation Team
- Step 5: Establish Resource Inputs and Constraints
- Step 6: Create Questions and Select Indicators
- Step 7: Develop the Evaluation Approach

Use the Logic Model framework to guide program and evaluation planning.

Assemble the Evaluation

Now that you have gone through all of the planning steps, take a moment to fill out the worksheet below. For assistance, refer to the Logic Model in Chapter 3.

Evaluation Questions Planning Worksheet

Program Stage	Program Features	Description of YOUR PLAN
Planning	Program Name	
	Program Goal(s)	
	Measurable Objectives	
	Inputs	
	Planning Evaluation (see Chapter 4 to select tools)	
	Program Process	
Implementation	Formative Evaluation (see Chapter 4 to select tools)	
	Outputs	
Results	Outcomes	
	Impacts	
	Summative Evaluation (see Chapter 4 to select tools)	

Chapter 4

Design and Manage the Evaluation

Consider Ethics, Bias, and Accuracy

Collecting Demographic Information

When designing your evaluation, include demographic questions to capture information about your participants. Demographic questions are especially useful to confirm that your programs and the evaluation effort are reaching the desired target audience. If your program is targeted to the general public, demographic information can be used to compare your respondents with U.S. Census data to check for bias in your evaluation. Factors that might influence your program include age, ZIP code or address, race/ethnicity, educational status, gender, and annual income of participants. Language may be an important factor in highly diverse populations.

When collecting demographic information, respect participants' time and privacy. Avoid the temptation to ask about every aspect of the participants' or respondents' lives. Include only the demographic questions that are essential to your program evaluation needs. For example, if you need to find out at what age young people are most responsive to the Take Me Fishing™ campaign, then it is appropriate to ask for the specific year of birth. If you only intend to compare age groupings (e.g., ages 1-5, ages 6-10, ages 11-15), then use a multiple-choice question with those groupings. Table 4.1 includes sample demographic questions and possible reasons for collecting that data.

If you want to follow up with participants in some way, you may request their mailing address, e-mail address, or telephone number. If you are collecting contact information from evaluation respondents, use a separate response card with a code number on it. See the later discussion about ethics and privacy for more details.

Summary of Program Evaluation Standards

Since 1975, the American National Standards Institute-accredited Joint Committee on Standards for Educational Evaluation has been concerned with issues of quality in evaluation. These standards were published in 2003.

Utility: Ensure that an evaluation will serve the information needs of intended users.

Feasibility: Ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.

Propriety: Ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.

Accuracy: Ensure that an evaluation will reveal and convey technically adequate information about the features that determine value or merit of the program.

Source: Joint Committee on Standards for Educational Evaluation
http://www.jcsee.org

Table 4.1 Demographic Questions and Reasons for Collecting This Data.

Sample Demographic Questions	Possible Reasons for Collecting This Data
What is your gender? *(Female, Male)*	To measure respondent gender, or to make gender comparisons.
What year were you born?	To measure respondent age or to make age grouping comparisons. If specific ages are not needed, use multiple choice with age ranges.
What is your ZIP code?	To measure respondent geography. ZIP code gives only a general area. Ask for the address if you need more specific information.
In which (country, state, county, city) were you born?	To measure respondent heritage. This might be important in an area with high levels of emigration or immigration of people, and to refine cultural information in a community.
What race do you consider yourself? *(American Indian or Alaska Native; Asian; Black or African American; Native Hawaiian or Other Pacific Islander; White; Other Race)*	To measure racial identity. These are the official minimum U.S. Census categories. The category Multiracial may be added in some surveys. Hispanic/Latino is not a race – Hispanic people may be of any race.
What is your cultural background? *(Hispanic or Latino; Not Hispanic or Latino)*	To measure cultural background. These are the two official U.S. Census categories. Depending on the situation, it may be valuable to measure other cultural categories to refine racial information within a community, but the best way to do this is to ask for country of birth or family origin.
What languages do you speak at home?	To measure language use. This question is most useful in a needs assessment to determine program delivery languages.
How long have you lived at your present address?	To measure length of residency. This might be important if a program is trying to reach new or less knowledgeable residents.
How many children under the age of 18 live with you?	To measure family size. This might be important for family or youth programs.
What is the highest level of schooling you have completed?	To measure educational achievement. If specific grade level is not needed, use multiple choice with education categories.
Consider your household income from all sources before taxes. As I read a list, please stop me when I get to the income level that best describes your household income in *(insert year)*.	To measure respondent income level to understand or compare income groupings. This question is worded as it would be presented during a telephone interview. Income categories are almost always used to avoid privacy violations.

Performing Culturally Sensitive Evaluations

When working with groups from different cultures, learn as much about the culture prior to designing and implementing the evaluation. If possible, have members of the various cultural groups targeted by your programs as advisers to ensure that you are creating a culturally sensitive evaluation that will capture the information you need.

Factors to consider when designing evaluations for different cultural or ethnic groups:

- Assess your own attitudes, beliefs, and values. Be aware of your own perceptions, as well as any preconceptions, stereotypes, and other potential biases.
- Understand that cultural norms may be different from your own and that you may need to adjust your approach.
- Be flexible in your selection of data collection methods. Some cultures may respond better to personal interviews, for example.
- Realize that your evaluation may take additional time if you need to build rapport or trust with the target audience.
- Describe what you are trying to accomplish with your evaluation so that your target group is aware of your intentions.

The consideration of ethnic and cultural differences is especially important in large cities and other culturally diverse areas. For example, surveys of residents in Miami, Florida, about water quality issues in the Everglades are performed in English (for North American and Bahamian subpopulations), Spanish (for Cuban and Mexican subpopulations), Creole (for Haitian subpopulations), and Portuguese (for Brazilian subpopulations).

Regardless of the language being used, it is important to avoid the use of jargon for all evaluation audiences. The use of "common language" is one way to guarantee that everyone has the same understanding of the ideas being communicated. In theory, common language involves simple and straightforward communication that is understood by the common person without any technical expertise or environmental knowledge. Common language can be easily translated and understood across neighborhoods, regions, races, cultures, and countries. Table 4.2 provides some examples of common language alternatives to jargon.

Table 4.2 Examples of Common Language Alternatives to Jargon

Jargon	Common Language Alternative
The Fish and Wildlife Commission in your state is quantitatively evaluating the effectiveness of our motorized vessel operation education program in decreasing benthic impacts to the Sea Cove region of Barrister Bay.	The Fish and Wildlife Commission wants to know if the Safe Boater program has reduced damage to the ocean floor in the Sea Cove region of Barrister Bay.
The Watershed Council is working to eliminate bacterial, phosphate, and petroleum pollution in the Sandy River and Sandy Reservoir.	The Watershed Council is working to keep oil, fertilizer, and manure out of our drinking water.

Evaluating Children's Programs

Conducting evaluation with children can be a fun yet challenging experience. Some common methods used to evaluate children's programs are observation, skills (performance) assessment, and content analysis of portfolios. Remember that there are limitations on evaluating people under the age of 18. Evaluators must always seek and obtain written permission (via parental consent form) from parents of the children participating in the evaluation. State and federal programs may have further limitations on evaluating children. Ethical considerations must always be taken into account when evaluating children.

Observation: Observation can be used to see how children interact with an exhibit at a learning center (how long do they stay at the exhibit, do they read the text, interact, or perform whatever activity they are supposed to do at that site?). Observation can also be used to see how well young people "get" the message that is being delivered. For example, if a program goal is to increase the safety of children in boats, observations could be performed at boat ramps to see if children put on a life vest prior to launch.

Skills Assessment: Skills Assessment can be used when it is possible to have children demonstrate a specific skill or practice. A simple skills assessment can be performed during or after a skills workshop. For a more in-depth evaluation of learning and critical thinking skills, you can use a pre-program and post-program worksheet or skills test to confirm that the workshop itself led to the improvement in skills.

Portfolio: Another comprehensive way to assess children in a learning situation is through a portfolio of work created in a multi-session program. The portfolio includes a variety of activities (e.g., art, writing, worksheets, demonstrations) that ask students to show what they have learned.

The portfolio collection tracks students' learning throughout a program and can be supplemental to or in place of a final test of knowledge or skills at the end of a program.

Other Methods: Other methods of evaluation for children's programs include quizzes, content analysis of learning journals or logs, and analysis of responses to broad open-ended questions – a technique to which children readily respond. For example, you could ask the group What were the most memorable things that happened today? Or you could ask them to Write three things that you learned in this program. If responses are collected from all participants, this can be a fruitful and revealing evaluation approach.

Evaluation Tip

When evaluating children, take into consideration that they may not be familiar with the methods you are using. For example, young students may not know how to participate in a focus group or how to answer survey questions. Relying on standard evaluation methods may not always work; see the Florida Aquarium Case Study for an example.

Case Study: Florida Aquarium Homeschool Program Evaluation

An example of an unforeseen problem in evaluating a children's program

The Florida Aquarium conducted an evaluation of its From Source to Sea homeschool program. The evaluation included pre-program and post-program tests to assess student knowledge and attitudes. Responses to statements about local environmental issues prior to the program seemed to indicate that the students were generally neutral on most of the topics. Responses were gathered using a Likert-type scale of 1 = strongly disagree to 5 = strongly agree. Following the program, students were again asked their feelings on the same environmental statements. Their answers changed very little and in some cases, post-program means showed negative shifts in attitude. This did not reflect the results obtained from other parts of the evaluation, including open-ended questions where the same students showed positive knowledge and attitude shifts. When program managers spoke with the students' parents about the evaluation, they learned that since many of the homeschool students are not tested in the same manner as their public and private school counterparts, the students were unfamiliar with how to answer Likert-type scaled questions. Thus, their answers were a reflection of their confusion about the testing mechanism, not their opinions about environmental topics.

Source: Staci Shaut, Florida Aquarium, Tampa, Florida

Address Ethics and Privacy Concerns

Whether children or adults, when you collect evaluation information from program participants or the public, you are asking them to provide you with their time, energy, and some personal information. Ethical considerations are designed to help evaluators avoid causing any physical, psychological, or emotional harm; to avoid deceiving evaluation participants; and to ensure the anonymity and privacy of sensitive information collected during the evaluation.

Basic issues of evaluation ethics and privacy:

- Always get permission from the participants when performing an evaluation. This could be a consent form or a verbal request to participate in the evaluation (as in an interview).
- Evaluators must always seek and obtain written permission (via parental consent form) from parents of minors participating in an evaluation.
- Do not collect information unless you will be using it. Collecting excess information takes more time for evaluation participants. It will also require extra time for you to sort through all of the superfluous data to find the information you really need.
- To protect and respect the privacy of participants and respondents, you must be scrupulous in managing the evaluation data. Always keep sensitive demographic information strictly confidential. In all cases, personal identification information (e.g., name, date of birth, address, telephone number, Social Security number) should be kept separate from evaluation data. You can use code numbers to connect the two data sets. Always store sensitive private information in a secure location. This means that computer databases are locked by passwords and survey or interview documents with personal information are kept in a locked facility or cabinet. Set a date to destroy personal information after the data have been completely entered and analyzed and the evaluation has been completed and reported. You do not want to maintain sensitive information in your files over the long term. If you are collecting contact information from evaluation respondents, use a separate response card with a code number on it. With identity theft becoming more common, you must maintain audience trust by carefully managing any sensitive data that you collect.
- Programs undertaken at universities or with federal funds are usually required to obtain institutional approval for any evaluation involving people ("human subjects") prior to beginning the work. University Institutional Review Board approval can take from several weeks or months to acquire, while federal approval through the Office of Management and Budget (OMB) can take up to one year (for federal guidelines, see the U.S. Department of Health and Human Services 45 CFR Part 46). The approval process ensures that you are not conducting an evaluation that causes harm to your participants in any way. There may be some exceptions to this requirement if you are simply performing observations or conducting focus discussions, but check with your organization or institution to make sure you meet requirements. Note that there are some pre-approved survey questions available from the OMB for surveys performed under federal funding. Many of the pre-approved questions focus on recreation and aquatic issues.

Avoid Evaluation Bias

Bias is any influence that prevents you from giving fair consideration to some element of your program. Bias can creep into your program as "leading" questions, as a slanted assessment of a program, or as a cultural prejudice. Bias can never be totally avoided, but it helps to have an evaluation team provide input into the process. If you are the program manager as well as the evaluator, you may be too close to the program to see what questions could be asked to best serve the evaluation objectives. If at all possible, at least one member of the evaluation team should be from outside the organization, to bring fresh eyes to the evaluation questions. If you hire an outside evaluator, ask them to work with your evaluation team and learn more about the program prior to the evaluation, so that they also have the benefit of the team input and program knowledge in their evaluation design. Table 4.3 describes how to avoid certain kinds of bias as much as possible.

Table 4.3 Sources of Evaluation Bias and Ways to Avoid Them

Sources of Bias	How to Avoid
Designer/Evaluator Bias can occur when the evaluator is also the program manager. The evaluator naturally wants to shed a positive light on the program, or may also have some preconceived notions about the program and its outcomes.	The pairing of internal and external evaluators combines the best of program familiarity and a more detached view. Make sure the evaluation team has at least one external member. Use a stakeholder advisory team throughout the evaluation process. Work with an external consultant for part or all of the process. Perform a careful pilot test of evaluation tools to gather audience feedback so that important information is not missed. Use qualitative (open-ended) questions to capture unexpected results.
Scoring or Grading Bias happens when an evaluator gives a low value or "grade" to something that is repugnant (or a high value to something that is agreeable) to their personal views.	Use evaluation tools that require little or no interpretation. Establish unbiased measures, such as ranking schemes or rubrics, for evaluating items or for grading student skills assessments or observing behaviors. Hide or remove participant names on evaluations to prevent bias of familiarity. Provide consistent training for all evaluators who will be scoring. Perform a pilot test of the grading scheme, with several people acting as graders, to check for an unbiased tool. Use a paper checklist or grading sheet to record data while performing informal interviews or observations – this avoids having to use memory to recall information.
Racial, Cultural, or Gender Bias crops up where there is a higher or lower expectation of some group or individual based on their race or gender. Because of cultural training, this is a very difficult bias to avoid.	Instead of comparing individuals to some pre-defined or abstract criteria, evaluate the skills or performance of an individual against the skills they had when they started the program. Select people representative of the intended racial, cultural, or gender group to review the evaluation plans or tools. Use multiple and diverse evaluators or observers in situations where cultural bias may interfere.
Non-response Bias occurs when using a random sample and the people who respond to the evaluation are different from those who don't. If these differences are major, this creates a bias in evaluation results.	Demonstrate that non-respondents are similar to or equivalent to respondents. Compare respondent and non-respondent characteristics to U.S. Census characteristics for the population. Contact and interview a set of non-respondents to see if their characteristics and views differ. For a large randomized survey, ask the contractor to perform a non-response bias analysis.

When writing focus group, survey, or interview questions, use these approaches to avoid bias:

Make questions clear: When designing evaluation questions, the wording can have an impact on the data you obtain. Your in-depth understanding of a program can cause you to accidentally design questions that lead participants astray. If your questions are confusing, such as, "What did you think of the aquatic education program you attended last year?" you may receive varying answers such as, "Which program?" or "It was good." Vague questions do not provide specific or useful information. A better questions might read, "Please rate the Streamside program that you attended in March 2006," with an accompanying rating scale.

Avoid double-barrel questions: Do not request a single answer to a combination of questions. For example, "Do you feel that the Evansville Fishing Pier is maintained properly and should be kept as part of Evansville Town Park?" If the person answers "yes," you will not know which part of the question they are responding to. The better wording for these two questions is: "Do you feel that the Evansville Fishing Pier is maintained properly?" and "Should the Evansville Fishing Pier be kept as part of the Evansville Town Park?"

Keep it short: Make items short and to the point so that participants can read or listen to them without losing focus.

Avoid negative items: Negatively worded items increase the possibility for confusion. In all cases, avoid the use of double negatives. Whenever possible, especially with children, avoid using questions that include the word "not," which may be missed by respondents.

Avoid leading questions: This is often a challenge because what seems unbiased to you may actually influence respondents' answers to your questions. Identifying your question with a person or organization can influence responses; for example, "Do you support the group, Citizens for Better Fishing Opportunities, in their campaign to keep the Evansville Fishing Pier open to the public?" If the public has a negative view of the organization, they may decide that they don't support having the fishing pier open because they feel the pier is linked with the group. Questions can also be leading in the way they are phrased, such as: "Do you support keeping the Evansville Fishing Pier open to allow families to have the opportunity to fish?" If the respondent says "no," it gives the appearance that they do not support families fishing.

Maintain equal-appearing intervals: Maintain an equal visual spacing between items on printed surveys, keep equal numerical spacing between points on a scale, and keep balanced grouping for multiple-choice questions (e.g., age groupings, income groupings). For Likert-type scales (e.g., five-point scales), use a symmetrical psychological spacing in the categories. For example: Strongly disagree, disagree, neither/neutral, agree, strongly agree.

Case Study: Place-based Education Evaluation Collective

An example of combined internal and external evaluators

In early 2002, several New England organizations came together to form the Place-based Education Evaluation Collaborative (PEEC) with the intention of evaluating their individual programs and laying the groundwork for broader research into the effectiveness of place-based education. Place-based education builds partnerships between schools and communities, bringing the energy and skills of students to bear on local environmental and social issues, creating exciting and relevant learning opportunities.

Through PEEC, the organizations have jointly contracted with a team of professional educational evaluators to individually and collectively evaluate several members' programs. Core members of PEEC include universities, parks, forests, and non-profit organizations, family foundations, farms, and other institutes. Members benefits from the combined force of external and internal evaluations, receiving feedback on their individual programs and building a large central repository of research on place-based education.

Source: PEEC, http://www.peecworks.org

Increase Evaluation Accuracy

A number of factors, including different external influences and various forms of bias, can reduce the validity and reliability (or confirmability and dependability) of any evaluation. When validity and reliability are low, the rate of measurement error is high. To avoid reduced validity, keep the evaluation consistent and controlled. An evaluation expert or contractor will help you to assess these factors throughout your evaluation approach.

Quantitative Evaluation Accuracy: The accuracy of a quantitative evaluation is referred to as validity and reliability. Validity is the extent to which a measurement instrument or test accurately measures what it is supposed to measure. Reliability is the extent to which an evaluation yields consistent and uniform results if repeated under the same conditions each time. For example, a scale is reliable if it weighs a fish three times in three minutes and gets the same weight each time. A survey is reliable if it gets a similar range of responses from a sample of the population each time it is administered.

Qualitative Evaluation Accuracy: In qualitative evaluation, the parallel concepts are called confirmability and dependability. Confirmability is the extent to which the evaluation results could be confirmed or corroborated by others. For example, a confirmable observation could be repeated if another evaluator visited the same classroom to perform the observation. Dependability is a bit more complex. Like reliability, it is concerned with obtaining the same results if the same thing is measured twice. Dependability asks if something said by someone is consistent with what that person has said at another point in time. The idea of dependability recognizes, however, that it is impossible to measure the same thing twice. Thus, the evaluator must account for the changing context within which the evaluation occurs by describing the changes that occur in the setting and how these changes affected the evaluation approach and results.

Triangulate the Evaluation

Another way to maximize the accuracy and consistency of your evaluation is to use a variety of tools instead of just one. Because no one approach to data collection can completely cover all evaluation needs, triangulation ensures that you get information in more than one way so you can more clearly see the whole story.

There are several ways that you can triangulate your evaluation:

Use several different tools to collect evaluation data, especially employing both qualitative and quantitative tools.

Collect information from several different audiences that are involved in your program (e.g., students, teachers, and parents).

Use people from different professional realms to provide parallel evaluations (e.g., a biologist to monitor stream improvement and a teacher to monitor student performance of stream restoration tasks).

Use several different evaluators to perform the evaluation (e.g., a research bureau to perform a telephone survey, and the internal evaluation team to perform observations).

Develop an evaluation plan that works from the unobtrusive/informal to the more formal. Start with unobtrusive evaluation techniques (i.e., observation, content analysis, case study, focus group), then use unstructured techniques (i.e., open-ended questions, writing/journaling, informal interviews, portfolio review), and finish with in-depth or formal techniques (i.e., surveys, interviews, skills assessments).

Triangulation can provide additional proof if you are asked to discuss or bolster your evaluation findings. Triangulation is especially useful to ensure that the correct recommendations can be made from an evaluation undertaken in an environment of doubt or conflict. If your recommendations to administrators are often criticized or if you are having trouble making convincing arguments based on evaluation results, triangulation can improve your evaluation's credibility.

Evaluation Tip

When collecting field data or recording notes from interviews and meetings, it is a good idea to organize, review, and/or summarize notes as soon after collection as possible. This ensures that key points or pieces of information will not be forgotten. These notes also can serve as a starting point for an individual case study or narrative to be included in the evaluation report alongside the quantitative data analysis.

Document the Evaluation Process

Documentation is the process by which the evaluator records how the evaluation was done, what was learned, and how others may benefit from the new information. This is also sometimes referred to as an audit trail. To provide a coherent summary report, the entire evaluation process must be consistently documented!

One of the major challenges in evaluation is to gather, store, and use the information that is collected. Be sure to store *both electronic and hard copy files* so they can be easily accessed by all team members. Documentation can be as simple as handwritten notes on scraps of paper, and as complicated as a computer database or spreadsheet. It can include any or all of the following: notes, data files, photographs, sketches, audiovisual media, skill assessment results, observation sheets, participant journals, focus discussion transcripts, and news articles. The point to remember is that the evaluation process must be carefully and consistently documented, even (and especially) if you are performing an informal evaluation.

Select the Tools that Best Fit Your Program

Evaluate Different Program Life Stages

When considering the tools to use in your program evaluation, begin by thinking about your program as having stages of life: planning stage, implementation stage, and results stage. This can also include a pilot testing stage and a stage for ongoing programs. Although a "results" stage is listed, the program has not necessarily come to an end. Programs are continually evolving over time – even if a program is discontinued, the lessons of that program are applied to the design of future programs.

Look at your program and see where it fits within this framework of life stage definitions:

- **Planning:** This is the first stage in program design. The program is being formulated to address the needs within a given setting or environment and under certain constraints or limitations. This stage includes all of the steps of program planning and development.
- **Pilot Testing:** This is the preliminary stage of program delivery, when program ideas are tested with a few (perhaps a dozen) target audience members. Pilot testing can offer enormously useful feedback for program improvement before full implementation.
- **Implementation/Delivery:** The program is now being delivered to the target audience. Program evaluation can determine necessary improvements or modifications.
- **Ongoing/Recurring:** An ongoing program can be evaluated at any time, to assess each separate program element as it is delivered or to examine a three-month or one-year record to find clues that the program is achieving goals and objectives.
- **Results:** The program may not be "finished," but it is time to measure and report conclusive results. This is the time to evaluate program outputs, outcomes, and impacts.

The stage of your program plays an important part in the evaluation tools you choose. In addition, your tool choices will be influenced by whether or not evaluation has previously been conducted. Consider where you stand with the program to identify the type of evaluation and tools needed.

Refer to the Logic Model in Chapter 3 to help guide your selection of evaluation tools, which are described in detail in Chapter 6.

Table 4.4 will help you begin your search for appropriate evaluation tools according to the stage of your program.

Table 4.4 Evaluation Tools for Stages in Program Life

Program Stage	Past Eval?	Type of Evaluation	Typical "Big Picture" Evaluation Questions	Potential Tools
Planning	–	Planning, Needs assessment, Feedback from past programs	Does the program fit with the organization's mission? Does the program address an important problem or need? Who is the target audience(s)? What information is needed by the target audience(s)? What program feedback would be useful to leaders?	brainstorming/nominal group technique, focus group, expert opinion/ Delphi group, interviews, citizen advisory group/ public workshop, content analysis
Pilot testing	No	Planning	Is the information being received by the target audience(s)? Are the messages, media, and delivery methods working? What can be adjusted to improve the program?	observation, skills assessment, interviews, surveys, content analysis
Pilot testing	Yes	Planning and Formative		focus groups, interviews, observation
Implementation/ delivery	No	Planning and Formative	How do participants relate to the subject? Are the messages, media, and delivery methods working? What can be adjusted to improve the program? Is the program meeting its objectives or desired outcomes? Is it addressing the problem?	observation, skills (performance) assessment, interviews, website tracking
Implementation/ delivery	Yes	Formative		interviews, surveys, observation, skills (performance) assessment,
Ongoing/ recurring	No	Formative and Summative	Is the program meeting its objectives or desired outcomes? Is the program adequately addressing the problem? Is the program cost effective? What lessons can be learned from evaluating the program? What program feedback would be useful to leaders?	surveys, skills (performance) assessment, interviews, website tracking
Ongoing/ recurring	Yes	Summative		surveys, skills (performance) assessment, case studies, interviews, website tracking
Results	No	Summative	Is the program meeting its objectives? Is the program achieving desired outcomes? Is the program adequately addressing the problem? Is the program cost effective? What lessons can be learned from evaluating the program? What program feedback would be useful to leaders?	surveys, observation, skills (performance) assessment, interviews, license sales tracking, website tracking, internal review
Results	Yes	Summative		surveys, observation, skills (performance) assessment, stewardship monitoring, case studies, interviews, license sales tracking, website tracking, public meeting, internal review, longitudinal study

Choose Specific Evaluation Strategies and Tools

Choose evaluation tools based on the type of program being evaluated, the life stage of that program, the goals of the evaluation, the program objectives being evaluated, the evaluation information needs, and the resources available to conduct the evaluation.

Use Table 4.5 to help guide your choice of evaluation tools for typical aquatic education programs. Note that these program categories are not mutually exclusive – for example, a school program may also be a science program; a community program may also be a recreation program. Remember that almost any evaluation tool can be used for almost any aquatic education or natural resources program, so the table suggests only a few tools especially suited to each program type. The evaluation tools are described in further detail in Chapter 6, where you will find a table comparing the advantages and disadvantages of each type of tool.

Table 4.5 Evaluation Tools for Typical Aquatic Education Programs

Program Types	Tools or Strategies for Small or New Programs	Tools or Strategies for Large or Ongoing Programs
Community, family, or non-formal programs such as informal courses, backyard conservation, camp or nature center, aquarium, education displays or kiosks, information resources (e.g., maps, newsletters, field guides), coloring books or games, and online or paper publications.	brainstorming focus groups pilot testing content analysis informal interviews internal review	focus groups content analysis case study surveys or interviews stewardship monitoring skills assessment longitudinal study
School-based programs such as curricula and classroom activities, learning and sharing networks (often electronic), and science competitions.	pilot testing observation learning log/journal /portfolio open-ended questioning website tracking	observation surveys or interviews worksheets/quizzes open-ended questioning skills assessment website tracking
Science and environmental education programs such as environmental research or databases for public access. These can be in schools or non-formal settings.	brainstorming expert opinion pilot testing website tracking	observation surveys or interviews worksheets/quizzes open-ended questioning skills assessment case study website tracking
Teacher training or leader development programs such as Project WILD/WET, Wonders of Wetlands, and teacher training institutes.	focus groups pilot testing observation informal interviews	focus groups observation case study surveys or interviews
Recreation and leisure outreach programs such as sportfishing campaigns, boating education and safety, fishing clinics or rodeos or tournaments, free fishing days, and fishing and boating shows or festivals.	brainstorming focus groups pilot testing observation informal interviews	focus groups observation case study skills assessment stewardship monitoring longitudinal study surveys or interviews

Program Types	Tools or Strategies for Small or New Programs	Tools or Strategies for Large or Ongoing Programs
Youth leadership development programs such as *Reel Kids.*	citizen advisory group brainstorming focus groups pilot testing observation informal interviews	citizen advisory group focus groups observation skills assessment case study surveys or interviews
Stewardship, watershed, or ethics education programs such as clean water or habitat restoration, watershed, angler and boater ethics programs, fish line recycling. citizen advisory group brainstorming focus groups	pilot testing observation informal interviews	citizen advisory group focus groups stewardship monitoring longitudinal study case study observation surveys or interviews
Outdoor or adventure learning programs such as outdoor skills for women, men, or families (e.g., Becoming an Outdoor Woman), adventure clubs, summer skills courses.	brainstorming focus groups pilot testing observation informal interviews	focus groups observation skills assessment stewardship monitoring case study longitudinal study surveys or interviews
Volunteer or partnership programs such as community service, citizen restoration, partnerships with schools and businesses.	citizen advisory group brainstorming focus groups pilot testing observation informal interviews	citizen advisory group focus groups observation case study surveys or interviews

Measure Long-Term Changes

By assessing long-term changes in a program, you can begin to detect patterns or trends in various program features, such as program participation, participant characteristics, program progress and development, participant outcomes (knowledge, behavior), and environmental impacts (changed conditions). Measurement of change over time in long-term programs may be called a longitudinal study or trend analysis.

If you have evaluated your program in the past, you may be able to assess longitudinal changes in your program. At the very least, you should be able to compare current conditions with past measurements if you use the same or similar measures and tools. You may also be able to use other data sources to conduct a trend analysis, such as changes in U.S. Census data as compared to your target audience, or similar analyses. Several types of longitudinal change analysis are outlined in Table 4.6.

Table 4.6 Types of Longitudinal Change Analysis

Longitudinal Study	Benefits	Drawbacks	Sources of Data
Trend Analysis: Looks at changes in factors over time. For example: Changes in the number of people that participate in fishing.	Shows if a trend is occurring (e.g., if fishing participation is increasing or declining). May require only a little time and effort.	Can show trends, but cannot show you why the trend occurred. Large and/or lengthy studies may be required to gain adequate statistical power to show trends.	Past evaluation, content analysis, literature search, U.S. Census data, participation rates, registrations
Cross-Sectional Studies: Like a slice of a tree, this study measures a certain characteristic in a defined population at single point in time. For example: Measuring all of the children in a class to see whether fishing experience impacts knowledge.	Useful in many settings. The "exposure" (e.g., fishing) and the "outcome" (e.g., knowledge) are determined at the time of measurement, so no additional measurement is required. The measurement can be repeated for a trend analysis.	Cannot be generalized to a larger population, but does allow comparison of subgroups within the sampled group.	Surveys, interviews, worksheets/quizzes, focus groups, observations, skills assessments
Panel Studies: The same group is measured or interviewed at different points in time. For example: A consumer panel is regularly interviewed to assess marketing plans.	Because the group members are the same, the study shows how individuals change over time. The "paired" statistical analysis for data from the same individuals is very sensitive to showing change.	This is very time consuming and can be expensive. Keeping track of people over time is difficult.	Focus groups, surveys, interviews (Marketing firms may have existing panels that you can use.)
Cohort Studies: Looks at a segment of the population (a group with similar characteristics) during different periods. For example: A study over time of Baby Boomer water conservation behaviors.	The study shows if and how things have changed within a segment of the population over time.	Results can be applied to the group you are evaluating, but cannot be generalized to the broader population.	Focus groups, surveys, interviews, observations
Pre-Program and Post-Program Studies: This approach measures participants before and after a program takes place. (Although this is not a "longitudinal" approach, it is listed here to remind you of the possibility of simply measuring before and after your program.)	This technique can show the direct outcomes of a program. If the same individuals are measured (e.g., in a small program), you gain the advantage of the more sensitive "paired" statistical analysis. If you measure a random selection of citizens (e.g., in larger surveys), you can compare averages from before and after the program.	Very labor intensive to perform on a large scale.	Surveys, interviews, worksheets/quizzes, focus groups, observations, skills assessments

Here are a few longitudinal analysis data sources:

- **Documents for content analysis:** Any long-term documentary information, such as long-term news coverage records, long-term recreational use records, and long-term biological information that correlates to changing human behaviors.
- **United States Fish and Wildlife Service National Survey of Fishing, Hunting, and Wildlife-Associated Recreation:** A national recreation survey that can help you understand current and future recreation demands for wildlife observation, fishing, boating, and other outdoor activities. The study is performed every five years, so there is now trend data (and state-by-state data) from 1991, 1996, 2001, and 2006. (http://www.census.gov/prod/www/abs/fishing.html).
- **U.S. Census and State Census data:** Mandated by the U.S. Constitution, the U.S. Census takes place every 10 years for the purposes of allocating Congressional seats, electoral votes, and government funding (U.S. data is available online at http://www.census.gov). Some states also conduct censuses.
- **License sales and boat registration tracking:** Boat registrations and license sales can give you an idea of how many people plan to participate in boating and fishing. Boat registration information can provide data on the types of boats used and the geographic distribution of boating and fishing participants. More information on license sales tracking can be found in Chapter 6.
- **State Comprehensive Outdoor Recreation Plans (SCORPs):** Many states create these plans to assess current and future recreation trends. SCORPs are specific to county and activity. When SCORP plans are not available, other recreation trend documents may be available.
- **Other Information Sources:** See the Resources section at the end of this guide for some additional sources of information. Compare your program with similar programs to gain insights into changes over time or differences across geographic borders.

Case Study: Assessing the Maine Watershed Stewards Program

An example of a longitudinal evaluation and survey

The University of Maine initiated the Watershed Stewards Program to educate people about the threats to water quality in lakes. Participants received 20 hours of training and performed 20 hours of service to their lake watershed. The evaluation sought to measure knowledge levels over time and to compare participants and non-participants living on the same lakes. The longitudinal analysis looked at the consistency of participant post-program test scores over five years and found that knowledge outcomes stayed fairly steady over the years. A survey to compare groups found that 68 percent of non-participants tried to reduce runoff, while 98 percent of trained stewards protected their lakes from polluted runoff.

Source: John Jemison et al., 2004, *Journal of Extension* at http://www.joe.org/joe/2004june/rb4.shtml

Manage the Evaluation

Successful evaluation requires management of both the process and the people involved in the evaluation. Putting time and energy into managing your evaluation gives you a better end product and actually saves time as the evaluation moves forward. Evaluation management tasks include developing a scope of work and evaluation budget and (if applicable) choosing outside contractors through the requests for proposals process.

Develop the Scope of Work

A scope of work is a detailed description of the tasks involved in the evaluation process. If you are internally conducting the evaluation, the scope of work is a guiding document for the evaluation team, along with being a resource for managers, staff members, funding organizations, and other program stakeholders. If you are partnering or contracting with an outside individual or group to conduct your evaluation, the scope of work specifically describes the tasks that need to be completed and roles that need to be fulfilled. The evaluation plan that you created in Chapter 3 will serve as a framework for the scope of work.

Key elements of a scope of work:

- purpose of the evaluation
- goals and objectives of the program
- definitions of terminology specific to the program evaluation
- description of evaluation questions to be answered, evaluation information needs, indicators or variable to be measured, preferred methods of measurement, and available data sources (if known)
- evaluation tasks to be completed
- task responsibilities (tasks assigned to internal and external evaluation team members)
- evaluation timeline
- evaluation budget

Define Evaluation Tasks

Your scope of work includes a list of tasks to be completed at each phase of the evaluation. Depending on your understanding of what is needed and the complexity of the evaluation, the task list may be comprehensive or simple. If you are hiring a contractor to assist with the evaluation, include the tasks you wish them to accomplish. If you are unsure of the tasks to be completed, describe them to the best of your ability in the scope of work or refer to Chapter 3 for more information on evaluation planning.

Create an Evaluation Timeline

The evaluation timeline includes allotted timeframes for all phases of the evaluation work. It allows for unexpected delays and accounts for any potential constraints, such as when events take place or when students are out of class during the summer. The most frequent problem encountered when creating a timeline is the urge to be too optimistic about how much time is required to complete tasks. Make sure your timeline includes ample time to collect information, analyze data, report results, and get feedback from your evaluation team. Consider adding an extra 10-20 percent to your original time estimates to allow for unexpected delays. The following is an example of an evaluation task list and timeline for a planning, formative, and summative evaluation of a one-month education program:

Hypothetical Example: Sarah is planning a fairly major aquatic education program that will last one year. She would like to design the program based on community needs. To do this, she will conduct a focus group before the program to get an understanding of the information needs and delivery methods best suited to the audience (planning evaluation). She will also perform a pilot test of the program with some audience members to get feedback about program content and delivery (formative evaluation). She also hopes to measure changes in knowledge and attitudes by conducting pre-program and post-program surveys of participants (summative evaluation). In addition, she wants to conduct follow-up interviews with a few participants to collect information on issues or unexpected results that may arise during the evaluation process (summative evaluation).

The task list that Sarah creates for her program (Table 4.7) could be condensed or expanded to fit any program, from a day-long workshop to a long-term education program. If the program already exists, you can start the timeline at any stage. You can also take a step back at any program stage and conduct planning or formative evaluations on existing programs, or add a pre- and post-program survey for participants.

Table 4.7 Example Task List and Timeline

Tasks	Timeline
Hold planning meeting with evaluation team	Month 1
Design and submit evaluation plan to evaluation team	Month 1
Receive comments on evaluation plan from evaluation team	Month 1
Create needs assessment	Month 2
Conduct needs assessment	Month 2
Summarize and report needs assessment results	Months 2-3
Design or modify program based on needs assessment	Month 3
Design evaluation tools and submit for review to evaluation team	Months 3-4
Receive comments on evaluation tool from evaluation team	Months 3-4
Conduct pre-program evaluation (pre-program survey can also be presented to participants before each program session, months 5-17.)	Month 4
Pilot test program	Month 4
Implement program	Months 5-17
Conduct post-program evaluation (post-program survey can also be presented to participants after each program session, months 5-17.)	Month 18
Conduct follow-up interviews with participants	Month 18
Analyze data collected	Month 19
Create draft final report and submit to evaluation team	Month 20
Receive feedback on draft final report from evaluation team	Month 21
Create final report	Month 21
Share evaluation findings and monitor use of evaluation results to improve programs and inform decisions	Months 22+

Develop an Evaluation Budget

Budgeting helps you anticipate the funds needed for your evaluation and allows you and your team to get a handle on staffing and resource needs. If your program receives external funding, evaluation will likely be a required element of your funding proposal. Many grants and funding groups expect at least 10-15 percent of your program budget to go towards evaluation. If your program is internally funded, make an evaluation budget to guide your expenditures and efforts. If your program is externally funded, update the budget that was a part of your original proposal. Discuss with supervisors or administrators what kinds of feedback information would be most useful and discuss the levels of program and evaluation efforts and budget with them.

Key components of an evaluation budget :

- **Staff:** The cost of personnel is usually the largest part of an evaluation budget. Determine who will

be conducting your evaluation. Use your timeline to estimate the amount of time staff will spend on tasks. Will this be senior staff members, interns, or regular staff? Account for base pay rates plus any associated overhead costs or benefits.

- **Travel:** Include travel costs, such as car rental or mileage, airplane tickets, lodging, parking, taxi fares, tolls, and other costs.
- **Communications:** If you anticipate long-distance phone calls or postage to be a significant part of your evaluation (such as in a mail or telephone survey), include these costs in your budget. If performing a mailing, determine the weight of a potential mailing and the number of people to be sampled in order to estimate postage costs.
- **Materials:** Consider the types of materials that will be used for the evaluation. You may need to make copies, or you may need a computer with specialized statistical software for analyzing the collected data.
- **Consultants:** The costs for outside contractors are separate from the costs for existing staff labor and materials. Consultants can be paid by a flat or hourly fee. You can estimate the number of days you think the evaluation will require, and divide the total consultant budget to determine the daily rate you can afford. If you are unsure about what to budget for this task, talk to other practitioners who have hired outside assistance or talk to several contractors to find out an appropriate price point. If your contractor budget is limited, remember that data entry and analysis often are the key service items to hire.
- **Other direct costs:** These include hiring a company to implement phone surveys or a marketing firm to conduct focus groups. Request a cost estimate from a few organizations or individuals that offer these services to estimate these potential costs.

Choose Outside Contractors

If your staff does not have evaluation expertise or time to conduct an evaluation, carefully select and use outside contractors who can provide the specific services, expertise, advice, or assistance that is needed by your organization or evaluation team. When trying to locate appropriate vendors, begin by asking other practitioners who they use and like. You can also contact university departments (such as social science, survey research, human dimensions, marketing, recreation and tourism, forestry, wildlife, fisheries, natural resources, education, or environmental science departments) for recommendations. Consider accounting firms if you need simple data entry and analysis tasks.

Evaluation Tip

In addition to estimating the time needed for each portion of your evaluation, look at the dependence each evaluation segment may have on the others. For example, if a needs assessment is being used to develop a program, you need to provide enough time for the needs assessment to be conducted and summarized if the results are going to influence program content. In addition, consider which sections of your evaluation may be completed simultaneously to save time.

Prior to hiring an outside contractor, consider the following questions:

- Have you defined what you want to get from your program evaluation?
- Are you able to afford an outside contractor? Remember that an outside contractor may be able to perform tasks more efficiently than an inexperienced internal team.
- Are your program objectives specific and measurable?

If you decide to go with an outside contractor, look for the following qualifications:

- education and training in evaluation
- experience with evaluation applied in your setting
- familiarity with your specific education topic and with the type of evaluation required

Consider these main sources when looking for an outside contractor:

- universities
- firms with staff that specialize in evaluation
- accounting firms that can perform accurate data entry and/or analysis
- marketing or market research firms

Each of these sources has strengths and weaknesses. Table 4.8 highlights some of what you may expect to encounter if you work with one of these contractor groups.

Table 4.8 Strengths and Weaknesses of Evaluation Contractors or Partners

Vendor	Strengths	Weaknesses
Universities	• access to specialists • credibility • large pool of resources • easy contracting if your organization is a state or federal agency • may be cost effective since faculty salaries are covered by the institution • may be cost effective if using student research or if performed as a class project • student researchers often are dedicated and hard working	• may take longer and be less responsive to specific or immediate needs • deliverables schedule may not fit with faculty, student, or class schedules • student researchers may be less experienced • faculty members may incorporate other research objectives related to their own research interests • cost depends on the complexity of the evaluation
Specialized evaluation, consulting, or accounting firms	• usually available and responsive to specific and immediate client needs • focus is on getting the client what they want • flexibility in contracting. • special expertise and training	• some firms may lack sufficient subject area knowledge or evaluation experience • large jobs may require that some tasks are outsourced • smaller pool of consulting talent to draw from • cost depends on the complexity of the evaluation
Marketing or market research firms	• quickest turnaround time • standardized methodology • ability to conduct large scale evaluations efficiently	• may not want smaller jobs or may require certain minimum payments to complete work • may not specifically tailor the evaluation to what you need, because they are set up to use standard marketing methods • may cost more than universities or small firms because they often work for large business and industry • evaluation rigor and design may not be as high quality if the firm is committed to multiple major research projects • cost depends on the complexity of the evaluation

Write a Request for Proposals

If you decide to get outside help with your evaluation, create a request for proposals (RFP; also called a request for quotation in some agencies). An RFP allows you to advertise the scope of work and to compare the proposals (evaluation approach, project experience, and costs) of different respondents. To get good proposals, make the RFP as clear as possible and get at least three bids so you can compare costs, approaches, and experience. Your organization may have sample RFPs on file, or you can find lots of samples through an Internet search. If you work in government, your agency will have a specified process for releasing RFPs. Check with your contracting or financial office for details. If you work for a non-governmental organization, you will want to release the RFP to aquatic education organizations and/or through agency and university colleagues who might be able to spread the word to the appropriate contractors.

Key Components of an RFP

The RFP will have two major sections. First will be a description of the program and evaluation needs (the scope of work) along with the technical details of when the proposals are due and how they will be reviewed. Approach the first section of the RFP by providing information on the purpose of your evaluation and what you want to accomplish from the evaluation. This information can be provided along with the other key parts of the RFP:

- brief background on the program and organization
- reasons and objectives for evaluation
- why you are evaluating
- type of evaluation you are conducting
- decisions that will be made as a result of the evaluation
- what the evaluation will help you accomplish
- scope of work and tasks that need to be completed (tasks can be very specific if you know how you want the evaluation conducted, or they can be very general if you want contractors to come up with creative approaches to the evaluation)
- potential evaluation timeline
- proposal submission deadline and timeline for proposal review and awarding a contract
- desired qualifications of the contract firm and personnel (include the kind of experience they should have, what they should know, and what they should be able to do)
- information on how you will rate or score proposals

This information will help potential contractors understand your evaluation needs, resulting in a good selection of proposals to choose from. An RFP that asks applicants to submit their suggested approaches also provides you with options for how the evaluation might be completed.

Including Budget Information in Your RFP

If you know the evaluation budget, you will have a choice of whether to reveal it in the RFP. If you do not provide a budget in the RFP, the proposals may not be meaningful or affordable. If your RFP contains a budget, you'll get a diversity of suggested approaches that are within your price range. However, the proposals will only end up telling you what services you can get for that budget, as most contractors will bid the maximum amount. An alternative is to provide an incentive for lower bids by allocating "bonus points" to applicants based on how low their bids are, so that the highest bidder gets no "bonus points" and the lowest bidder gets the maximum "bonus points" during proposal review.

What to Request from Your Applicants

The second section of your RFP will tell the respondents what information they need to submit for review. Here are a few key items that you might want potential contractors to include in their proposals:

- **Qualifications and key personnel:** Ask the contractor to briefly describe their qualifications and to provide information on key staff members who would be working on your evaluation, including their areas of expertise and experience. You might also ask the respondent to define their level of availability to perform the evaluation, so that your evaluation doesn't compete with the firm's other commitments.
- **Past evaluation experience and references:** Ask the contractor to describe at least three relevant past projects (within the past few years), including names of references relevant to the firm's evaluation work. Look for someone who understands the process of evaluation for programs

similar to yours, such as other aquatic, natural resources, or environmental education or outreach campaigns, school programs, online training, teacher training, statewide campaigns, or similar efforts.

- **Evaluation approach:** Ask for a detailed description of how the contractor plans to approach the evaluation tasks that you have listed in the RFP. The contractor should include a description of their evaluation philosophy, how they work with clients, and instruments that they propose to use for the evaluation.
- **Evaluation cost:** Ask the contractor to provide a cost budget for the evaluation, with a detailed breakdown of how the applicant proposes to spend the budget.

Proposal Review Guidelines

Many RFPs also include a section describing how the proposals will be reviewed or scored. To establish this process, make a list of the most important proposal qualities and decide how valuable each factor is for your evaluation. Develop guidelines for scoring each factor (see the section on creating a rubric in the Skills Assessment fact sheet in Chapter 6 for ideas). For example, you might ask reviewers to score each item from 0 to 5, or you might ask reviewers to grade each item on a scale of 1 to 100. A final score will be calculated for each proposal by each reviewer, and average scores can then be calculated if there are multiple reviewers. Using this process, you will be able to make an unbiased decision about which proposal best serves the evaluation needs of the organization.

Once the proposals come in, the evaluation team or a special proposal team will review the submissions. Ask additional technical experts to join the review process if you need help reviewing evaluation approach or other technical aspects. Expect to engage in a careful review of the RFP process so that no contractor can claim that it was unfair.

Scoring statements that may be of interest to aquatic program evaluation proposals:

- The proposal clearly conceives, defines, and describes the evaluation approach.
- The proposal contains appropriate strategies and timetable.
- The applicant has clearly justified the proposed evaluation approach.
- The proposed evaluation approach has technical merit.
- The proposal suggests an innovative approach.
- The proposed evaluation process incorporates appropriate stakeholder input and involvement.
- The proposed approach meets all applicable ethical and environmental standards.
- The proposed evaluation approach will yield results that benefit our organization.
- The applicant has the necessary experience to complete the work.
- The proposed budget is within the advertised budget limit.
- The proposed budget is lower than the advertised budget limit. (This score can be used as a bonus incentive to encourage streamlined budgets. To be effective, the bonus point approach must be included in the RFP.)

Chapter 4 Summary of Best Practices

- Collect only the information you need and maintain high ethical and privacy standards.
- Avoid bias and maximize the accuracy of your evaluation approach by increasing validity/reliability and using multiple evaluation tools.
- Choose evaluation tools based on the type of program being evaluated and the life stage of that program.
- Create an evaluation scope of work that is specific to the evaluation tasks that need to be completed.
- Use outside contractors or partners to provide your organization with specialized evaluation expertise, advice, and assistance.
- When hiring contractors, create a detailed request for proposals (RFP) so you can receive proposals that best meet your evaluation needs.

Chapter 5

Create Useful Results from the Data

Make Evaluation Results More Useful

To make your evaluation results more useful, they must be shared with appropriate audiences, such as managers and organization administrators, budget directors, program planners, and outside stakeholders. The conclusion and recommendations will address both program objectives and improvement needs and the needs of the evaluation audiences.

The content of the report depends on your primary audience(s). For example, ***top-level administrators*** may be most interested in knowing whether the program met its intended goals or whether the program was cost effective. Administrators may be most interested in data and details from a summative evaluation, to assist with decisions about program funding and resource allocation.

Program staff might be more interested in seeing the evaluation of program activities to improve the overall quality of the program. Staff might also want to use positive feedback and results in future marketing and promotion activities. This audience would benefit from a general report of quantitative and qualitative evaluation results, including qualitative measures from feedback and interview transcripts, along with recommended actions to improve programs.

Funders are likely to desire a report that includes an executive summary, a description of the organization and program under evaluation, an explanation of the program goals and objectives, a summary of the evaluation methods and analysis, and a listing of the conclusions and recommendations.

In addition to internal and funding audiences, most evaluations have other audiences as well. To reach these wider audiences, it may be helpful to include a descriptive program timeline: how and why the program was planned and developed; how the program was implemented; factors that aided or challenged the program's overall success; a summary of evaluation results; and recommendations for program modifications.

Typical Audiences for Evaluation Information

- program managers, staff
- top-level administrators or decision makers
- program participants or clients
- board members
- agency biologists
- industry representatives
- university researchers
- funders or investors
- conservation organizations

Analyze the Evaluation Data

Data Coding and Entry

Whether you are performing a needs assessment (planning evaluation), improving an existing program (formative evaluation), or looking at program outcomes and impacts (summative evaluation), you will come to a point when the data need to be compiled and analyzed. When you reach that point, the evaluation team will meet to discuss issues such as the coding of data, open-ended responses and qualitative results, and the overall organization of the database.

The data analysis choices are further discussed below. The fact sheets in Chapter 6 include suggestions for compiling and analyzing data for each different type of evaluation tool. Consult with a statistician or evaluation expert if help is needed with any of these steps. See Chapter 4 for information on finding contractors to help with the evaluation process.

Scales of Measurement

There are four basic levels of measurement that can be used for evaluation data. Even if these descriptions seem abstract at this point, you will later find these concepts are a useful language to connect to real-life evaluations and the lingo of evaluation contractors:

- **Categorical/Nominal:** Numerals, labels, or names are assigned to the data such as gender, race, religious affiliation, political party, college major, or birthplace. The only comparison that can be made among nominal values is whether they are the same or not, there are no "less than" or "greater than" relations. The mode (the most common answer in a data set) measures central tendency.
- **Ordinal/Rank:** The numbers assigned to objects represent the rank order (1st, 2nd, 3rd…) of the entities measured. You can determine which variable is greater or lesser or equal to other variables, but you don't know the intervals between the ranks. Examples include the results of a race (without time intervals), and most measurements in the social sciences, such as attitudes, preferences, and social class. Both the mode and median (the number that separates the upper and lower halves of the distribution of answers) can be used here to measure central tendency.
- **Interval/Discrete:** These values have all the features of integers (whole numbers), with equal values between the numbers. A mean (or average) can be calculated for these values, in addition to mode and median. Examples include dates, temperatures (Celsius or Fahrenheit), IQ scores, and scores on many social survey questions.
- **Ratio/Continuous/Scaled:** These measures have all of the features above, but they are related to a scale with a defined zero point. Examples are distance, length, temperature (Kelvin), age, number of years of residence in a given place, number of fish caught in a day, or number of events produced in a year.

Qualitative Data Analysis

Qualitative evaluation recognizes that programs take place within a framework of subjective experience, social context, and historical time.

Qualitative evaluations seek to discover how people think and feel about the circumstances of an educational program or other event. Rather than starting with a hypothesis, the qualitative process seeks to explain how and why something operates in the way that it does.

Qualitative data come in many forms. The data may consist of varied sources, such as transcripts from open-ended or in-depth interviews, recorded observations, focus groups, texts and documents, multi-media or public domain sources, policy manuals, photographs, and autobiographical accounts.

After the raw data are entered into a database, the interpretation process will help you to understand your program audience and to see commonalities, variations, and relationships in the information. Maintain a focus on generating the results that are of greatest interest to the anticipated users of the evaluation results. Then engage in further exploratory analyses that may be of broader interest.

Qualitative analysis usually involves intra-case analysis and cross-case analysis. A case may be a single individual, a focus group session, or a program site. Intra-case analysis examines a single program site or a single event. Cross-case analysis systematically compares and contrasts multiple program sites or participants, such as a comparison of two people experiencing the same program. The analysis seeks to identify patterns or commonalities, uncover the essential nature of a program or event through deep description, recognize behaviors that spring from cultural patterns, or analyze stories or interviews (narratives) for clues and insights into participant experiences.

Simple steps for qualitative data interpretation:

1. Decide on the data documentation and interpretation scheme (see Chapter 6) and enter the data in a database or other analysis program.

2. Understand the program or participants by creating a "thick description" of the phenomena being evaluated.
3. Create detailed case studies or portraits of specific aspects or relationships that appear within the data.
4. Use intra- or cross-case comparisons to look for patterns or themes that explain how and why relationships appear as they do.
5. Put the new knowledge about the program and relationships into a real-world context to create results and recommendations.

The evaluator can ask the following questions during the qualitative analysis:

- What patterns and common themes emerge in responses to specific items?
- How do these patterns (or lack thereof) help to illuminate the broader evaluation question(s)?
- Are there any deviations from these patterns? If yes, are there any factors that might explain these atypical responses?
- What interesting stories emerge from the responses?
- How can these stories help to illuminate the broader evaluation question(s)?
- Do any of these patterns or findings suggest that additional data may need to be collected?
- Do any of the evaluation questions need to be revised?
- Do the patterns that emerge corroborate the findings of any other evaluations that have been conducted?

The practice referred to as Computer Assisted Qualitative Data Analysis (CAQDAS) employs computers to identify possible themes, concepts, and contexts within a mass of qualitative data. Popular qualitative analysis software systems include ATLAS.ti, HyperRESEARCH, MAXqda2, N6, NVivo, QDA Miner, Qualrus, and Transana. These programs are designed to help evaluators sift through transcripts, case notes, survey results, articles, pictures, and other varied documents for content that can provide insight into program success.

Quantitative Data Analysis

Quantitative data are usually entered into a spreadsheet or statistical analysis program, if the dataset has not already been entered by a survey service, accounting firm, or other evaluation contractor.

Simple steps for data entry and analysis:

1. Decide on a data coding or transformation scheme (see Chapter 6).
2. Enter the data in a spreadsheet or database program.
3. Perform descriptive summaries on the data.
4. Perform inferential summaries on the data.
5. Perform group comparisons or other higher order statistical analyses on the data.

After the raw data are entered in a spreadsheet, data analysis will help you to understand your program audience and to see relationships, similarities, and differences in the evaluation data. Keep the early data analysis process focused on generating the results that are of greatest interest to the anticipated users. Then engage in further exploratory analyses that may be of interest.

Two major summaries are used to describe and glean information from evaluation data:

- **Descriptive statistics** are used to summarize or describe data. Descriptive statistics include the frequencies for all of the variables (how often questions were answered in what way: for example, *75 percent of respondents said…*) and the means (averages) and standard deviations for items measured with interval variables.
- **Inferential statistics** are used to model patterns in data or to draw inferences about the larger population (e.g., from which the sample was taken), while accounting for randomness and uncertainty in the data. Inferential tools help with hypothesis testing, predicting future observations, describing associations (correlation), or modeling relationships (regression). Other modeling techniques include Analysis of Variance, time series analysis, and data mining. Inferences may only be extended to the whole population if the sample is random and representative of that population. (See glossary for further definitions.)

If you are unsure of how to analyze the data you've collected, use information available with statistical software and consult with one or more professionals who are particularly skilled in that aspect of data analysis. Common spreadsheet programs (e.g., Microsoft Excel and similar) perform most of the simple descriptive and inferential statistics needed to summarize evaluation data. This makes it easy to perform a simple evaluation with an in-house data analysis plan.

Two popular software packages that are used for statistical analyses are SPSS (www.spss.com) and SAS (www.sas.com). There are also a number of free online statistical analysis tools that are available. For example, links to tools can be found at the website of the International Statistical Institute (ISI) (http://isi.cbs.nl/FreeTools.htm) or at "Web Pages that Perform Statistical Calculations" (http://statpages.org/javastat.html).

Statistical Analyses

It is not necessary to know statistics in great detail to perform an evaluation! Statistical experts are available to help with the nitty-gritty aspects of data analysis.

The purpose of statistical data analysis, in the broadest sense, is to summarize the similarities in a set of observations, and the differences between subgroups within that set of observations. For example, the statistical summary will tell you how many people you have with certain demographic characteristics (e.g., "the group was 27 percent African-American"), how many people are knowledgeable about a certain topic (e.g., "85 percent of respondents correctly identified the picture of a striped bass"), and how subgroups differ (e.g., "women were significantly more likely to support watershed protection measures than men"). For subgroup comparisons to be "statistically significant," they must have a statistical value of $p<0.05$ (or sometimes $p<0.01$). The "p" value is defined as the probability that the observed differences are due to chance alone; when the "p" value is low, then the differences are real and your hypothesis is supported.

Table 5.1 lists types of higher order statistical analyses and their uses.

Case Study: Missouri Stream Teams Evaluation

Example of the use of the ANalysis Of VAriance (ANOVA) statistic to compare groups

A survey was performed to assess the effectiveness of the Missouri Stream Team Program, which provides education about stream ecology and stewardship responsibility through a school-linked stream adoption program. The ANOVA was used to determine the overall effects of the program and compared students by experience level, rural vs. urban residency, and school affiliation. The results demonstrated that experienced students showed more positive overall environmental knowledge and attitudes and that the differences were statistically significant at $p<0.05$. The evaluation team concluded that the Missouri Stream Team Program has a positive effect on participant knowledge and attitudes toward the environment.

Source: Brian Roddiger and Janice Schnake Greene, Southwest Missouri State University

Table 5.1 Types of Higher Order Statistical Analyses and Their Uses

Statistic	What it Measures	Data Types	Use in Evaluation
Correlation / Pearson Coefficient / Spearman's Rank Coefficient	Correlation is a statistical measure of how much the movements of two variables are coincidentally related.	Interval or ratio data for two or more independent variables.	Correlation does not show causality, but you can report the correlation and suggest further investigation.
Chi-square (Greek chi: 2)	A statistic used to compare frequencies of two or more groups of nominal data. The chi-square is used to determine whether a value deviates from the "expected" outcome solely by chance.	Categorical data for two or more independent variables.	If two groups are measured to be supportive of a certain program, but one group is slightly more supportive than the other group, the chi-square test can tell you whether that difference is statistically significant.
Mann-Whitney U / Wilcoxon rank-sum	This test allows us to say if one of two sets of independent observations is significantly larger by comparing the medians of the data. These tests are for non-parametric (non-normal) data.	Ordinal data for two independent variables.	Used to analyze randomly collected data that does not fit a "normal" (Bell) curve, called non-parametric data.
Student's t-test (provides the "t" statistic)	A statistical significance test used to compare differences between means of two groups.	Means of interval or ratio data for two groups.	It can be used to compare independent samples (e.g., males vs. females) or on paired data (i.e., two measurements taken from the same person, perhaps before and after an event).
Analysis of Variance / Fisher's ANOVA / ANCOVA / MANOVA (provides the "F" statistic)	ANOVA tests the statistical significance of the differences among the mean scores of three or more groups on one or more variables by splitting the variance (variability) into different parts.	Means of interval or ratio data for two groups.	ANOVA tests the difference between the means of two or more groups, so is useful for comparisons of multiple groups.
Regression Analysis / General Linear Model	A method for investigating and modeling the relationship between a dependent variable and one or more independent variables. The regression equation defines a straight line that approximates the information in a group of data points, and shows any trend that exists among factors.	Continuous data for two or more independent variables. The general linear model is a "mixed" form that allows the use of both ordinal and continuous data.	Any analysis where you wish to predict a relationship between variables.

Statistic	What it Measures	Data Types	Use in Evaluation
Discriminant Analysis	A variation of regression analysis or analysis of variance where the independent variable(s) are categorical.	Categorical data for two or more independent variables.	Any analysis where you wish to predict into which of two or more groups an object is likely to fall. It is often used to analyze factors contributing to complex behaviors.
Factor Analysis / Principle Components Analysis / Cluster Analysis	A multi-variate data reduction technique that aims to summarize a large number of variables with a small number of factors. The analysis is based on a matrix of correlations between factors.	Correlations of interval or ratio data for up to 100 independent variables.	Factor analysis can be used to reduce a broad set of attitude or behavior measures into several indices for better understanding.

Reach Coherent Conclusions from the Evidence

Where's the Evidence?

As the data analysis progresses, you will begin to see what evidence has been developed by your evaluation. Evidence can be defined as the data, documents, objects, pictures, or verbal statements that prove or disprove your notions (hypotheses) about the program. Just like a lawyer in a court of law, you will use this evidence to build a strong case for the conclusions and recommendations included in your report. The manner in which conclusions should be stated is primarily dependent on the sampling strategy and sample size employed. For random sampling, you can generalize about the population from which the sample(s) were drawn, but for the other sampling types you can only refer to the group of respondents (refer to Chapter 3 for a discussion on sampling design).

At this point, the evaluation team will meet to discuss the results of the statistical analysis. The goal is to understand what the results mean in real life and what this reveals about the program. The best way to approach this is to summarize each piece of data analysis in plain language, such as "23 percent of audience members are over the age of 65"; "85 percent of watershed residents say they are willing to reduce lawn fertilizer use"; or "33 percent of program participants were able to perform the necessary skills to cast a lure."

The task for the evaluation team is to determine what the key evaluation evidence is. More than just facts, evidence is any information that helps you answer the evaluation questions that you developed in Chapter 3. Evidence can take many forms, including your own experience with the program as well as the various data collected through the evaluation tools. Evidence is informative, robust (i.e., not open to question), and updated.

Refer to your evaluation plan and Logic Model to determine which results represent program outputs, outcomes, or impacts. If the results are not understood by team members, ask a statistical expert for assistance with interpretation. They may be able to determine why you are getting mixed results. You will also want to ask program staff members (e.g., instructors) for their thoughts, as they may know some details (e.g., weather, participant differences) that might explain incongruous results or differences (or lack thereof) in the evaluation results.

Identify the Best Evidence

Evidence is the information that helps to answer your evaluation questions. To determine which information provides the critical evidence for your report, consider the following questions:

- Which information indicates how the program has performed? Which information demonstrates that the program has or has not met its goals and objectives?
- Is there information to show that the program has met the standards set for a successful program?
- What conclusions regarding program performance are justified by comparing the available information to the selected standards of success?
- What additional information is provided by a close-up or detailed observation of the program?
- What information is available on the outcomes or impacts of the program?
- How can information about the "lessons learned" from the evaluation be used to improve program success?

One important concept you need to consider is the notion of causality – whether one thing caused another thing to happen. For example, it would be wonderful to show that more people are licensed to fish as a result of your statewide fishing campaign. You will need more than a correlation between the two variables, however, to prove that the campaign caused people to get fishing licenses. The increased license sales could be due to any one of a number of other coinciding factors, such as increased leisure time, less travel to distant destinations, an aging population, or some coincidental coverage in the news.

You can see how difficult it is to show that a program caused a change in behavior! Causality usually is not proven through simple program evaluation. It is usually demonstrated through rigorous experimental research, where a "control" group demonstrates that other outside factors did not cause the observed effect.

So instead of trying to show causality, you can apply more reasonable criteria in program evaluation. The first requirement is to maintain confidence in the accuracy of your data, by maintaining the validity and reliability of your evaluation techniques. Refer to the Chapter 4 section on bias and accuracy to review these concepts.

A more reasonable criterion for program evaluation might be if you find a strong correlation between your program and some desired outcome. You can then say that the results demonstrate a potential relationship between your program and the desired outcome, and that you might reasonably expect that the program had some effect on the results, even though you cannot prove it. It is fairly safe to say, for example, that a program influenced shifts in knowledge or attitudes by using surveys or interviews that take place immediately before and after a program.

Another important tactic in analyzing evaluation evidence is to avoid preconceived notions about your program and its effects. Preconceived notions might cause you to miss something important. For example, if you are busy looking for an expected program impact, you might overlook an unexpected outcome that demonstrates an unusual or unlikely (and potentially positive) result of the program.

Develop Sound Recommendations

In the "conclusions and recommendations" section of the evaluation report, you will present findings about the strengths and weaknesses of the program and recommendations for program improvements. Based on the results of the evaluation, you will suggest specific actions to help the program better meet its goals and objectives.

Be careful to be concise in this section of the report. Focus on relevant findings that you are confident about and that support your recommendations. Choose findings that are the most conspicuous, outstanding, and representative of the program and that are needed for your reporting purposes – and for the anticipated users of the report. Keep your recommendations within the scope of documented evidence. To expand your recommendations beyond the information you have will open your evaluation to criticism and credibility issues.

Beyond evaluation findings, it may be beneficial to capture any additional insights that your team has gained from participation in and knowledge of the program. These might be insights that are not derived from the data alone. This information enters the realm of what is called "anecdotal" information and evidence. Anecdotal information often gets a

bad rap, but in some cases it may be your best resource to provide "inside" information on what works best in a program on the ground. The main flaw in anecdotal evidence is that there is no guarantee that it is not hand-picked. As long as anecdotal information is clearly labeled as such, it isn't a problem.

Making a recommendation doesn't necessarily mean that action is going to take place, but it provides program managers and administrators with a baseline from which to make decisions. The next important step is reporting your results to key decision makers.

Communicate, Use, and Monitor Evaluation Results

The final step in the evaluation process is to create a summary report of your findings. This final report will include a background of why you conducted the evaluation, a breakdown of the results, and a summary of conclusions and recommendations supported by these results. The report is the key to communicating results to your intended audience.

Reporting Questions for the Evaluation Team

When preparing the final report and other products that will communicate evaluation results, consider the following questions:

- How can the substance and format of the evaluation report be tailored to meet the needs and interests of a given audience?
- How will the evaluation report be organized? How can the findings based on qualitative and quantitative methods be integrated?
- Does the evaluation report distinguish between conclusions based on robust data and those that are more speculative?
- Where findings are reported, especially those likely to be considered sensitive, have appropriate steps been taken to make sure that promises of confidentiality are met?

Developing Conclusions and Recommendations from the Evaluation Evidence

Here are some suggestions for developing conclusions and recommendations:

- The conclusions and recommendations are supported by the results of accurate and robust data and systematic analysis of both quantitative and qualitative data.
- Include stories and quotes to enrich the report. These items are especially useful in bringing a real face to the report and in reflecting the qualitative analysis you have performed.
- The report may also include some sweeping insights that are based on involvement, impressions, and anecdotes – as long as they are separate and clearly labeled as anecdotal evidence.
- In addition to the methods section in your report, be sure to provide full documentation (in an appendix) for all findings.
- If problems were encountered that may have affected the findings, discuss possible biases and efforts used to overcome them.

Format the Evaluation Report

Flexibility should be a required quality for developing a final evaluation report. You must develop a reporting vehicle that is in the most useful format for each group of anticipated users. For example, you might want to develop a preliminary report for team review and feedback, a technical report for review by analysts and administrators, and then a series of shorter reports or press releases for various evaluation audiences.

Taking flexibility into account, many evaluation reports are organized into standard sections so that audience members have access to the information they need.

Title Page and Table of Contents: Insert a title page and table of contents at the beginning of the report. Include the names of the evaluation team and report-writing team and the date of the report. Depending on the length of your report, consider

Standard Report Format

Written evaluation reports follow a standard format that consists of:

- Title Page and Table of Contents
- Executive Summary
- Background and Evaluation Questions
- Methods
- Results
- Conclusions and Recommendations
- Appendices

including a list of all charts and graphs so that your audience can quickly locate them.

Executive Summary: Immediately after the title page, present an executive summary of your evaluation report. The executive summary is usually one or two pages long and provides an "up-front" summary of the most significant results and recommendations of the evaluation. Often formatted into a bulleted list, each concise statement might include a key result and the recommendation stemming from that result. The executive summary must be easy to read – many of your key audience members are more likely to read the executive summary than the full document. This summary can also be used as a reference when your reader has finished the report and wants to review the major points. Despite its early presentation in the report, prepare the executive summary last, when you are most familiar with the results of your evaluation.

Background and Evaluation Questions: Before you get into the details of your report, you will want to provide a general background to the program evaluation. Explain the basis for your evaluation, the factors that motivated you to conduct the evaluation, the objectives for the evaluation, and any other important information relating to the evaluation planning process. State the "big picture" evaluation questions (see Chapter 3) that you are trying to answer through the evaluation. By explaining the justifications for the evaluation, your audience will have a better understanding of the results.

Methods: This section contains a complete description of how you got your data and information. Be specific about what evaluation tools you used, how the evaluation questions or instruments were designed, how you derived your sample (if applicable), and how the evaluation and data analysis were performed. Make note of how many people participated, response rates, and the time it took to conduct the evaluation. For archival purposes, raw data may be included in the report appendix or on attached electronic media for future reference or comparison.

Results: This section of the report will be fairly detailed, explaining the major and important evaluation results, along with any unexpected or surprising outcomes. Display the results in the form of tables, charts, graphs, and pictures. Incorporate descriptive text to explain what these visuals mean, and to emphasize important points. Present the text and visual elements side by side to allow readers to easily make important connections. Refrain from making value judgments about the results in this section of the report – make it a factual summary of the results.

Conclusions and Recommendations: The conclusions and recommendations section is where you will summarize the key points and express thoughts about future actions that are supported by the data analysis. Your conclusions and recommendations are directly based on the data analysis results. Suggest courses of action in concise statements so that an action plan can be created or decisions can be made based on your results. The conclusions will often be presented in narrative form, while the recommendations can be presented as bulleted action items. Like the executive summary, this section of the report may get the most attention from readers and can affect administrators' decisions with respect to future program support.

Appendices: Appendices contain information that is important to your report, yet does not fit into the main body of the document due to length or relation to the text. As appropriate, the appendices might include data collection forms or instruments, the raw data in tabular/spreadsheet or electronic form, transcripts of interviews or discussions, testimonials, copies of observation sheets or skills assessment records, case studies, and related literature.

Reporting for Future Use

If there is a chance that the evaluation will be repeated in the future, or if the evaluation results are to be filed as part of the public record, a detailed archival evaluation report must be produced, including raw data in spreadsheet or electronic format. The report must include enough information so that future program managers or contractors can understand how and why you performed your program evaluation and can duplicate part or all of the evaluation. A detailed report and dataset are necessary if future evaluators are to be able to make comparisons over time. Evaluation reports generally are stored in central paper and electronic files, as well as being presented to the public on accessible websites in summary and/or PDF form.

Typical Grant Organization Reporting Requirements

Programs receiving government or private grant funds are often expected to establish a program evaluation plan and to submit a final report to present evaluation results and highlight program success. It is in your interest to design and implement the evaluation to suit the needs of the granting organization, as well as to gather information to improve the program. Granting organizations typically require that final reports answer some or all of the following questions:

- What measurable objectives were set for the program and what indicators were used to measure performance? To what extent did the program achieve these objectives and levels of performance? How and why did the program succeed? Where did the program fall short? Were there unanticipated or unplanned outcomes?
- Did the program encounter internal or external challenges? How were the challenges addressed? Did the program employ any new tools, strategies, or approaches? Was there something the funding organization could have done to assist in program success?
- Have there been other sources of support for the program, such as grants, business partnerships, or community alliances?
- What lessons were learned from undertaking the program?
- What impact has the program had to date? What are the plans to measure long-term impacts of the program?
- Were volunteers used in the program? In what ways were volunteers most and least effective in helping to achieve program goals?
- What are the future plans for continuation and/or modification of the program?
- How were program funds spent? Was the program cost effective? Were there any deviations from the original program budget, and if so, why?

Case Study: Future Fisherman Foundation

An example of real-life final reporting requirements

Final report requirements for the Future Fisherman Foundation National Fishing and Boating Education Initiative:

Program Background Information

Unit and Lesson Plans: Include an overall unit plan and daily lesson plans for the fishing or boating unit, including student learning objectives, materials needed, and methods of assessing student achievement.

Photos of Program Implementation: Pictures are to be labeled with student, teacher, or volunteer names and location. Include signed photographic release forms for all people included in photographs.

Financial Accounting Report: Include the original program budget and receipts for purchases, and return any money not spent or accounted for through receipts. Funds up to 10 percent of total budget may be transferred to a different category within your budget without prior approval. To transfer fund amounts over 10 percent of total budget, you need to receive approval from the Future Fisherman Foundation.

Assessment: Include the number and grade levels of students who participated in the fishing or boating unit. Include information of any further student participation in addition to the required classroom time. Include both examples of your classroom assessments and the results (e.g., "95 percent of students were able to cast into a hula hoop at a distance of 20 feet"). Include anecdotal stories about how the unit was received by students, parents, and the community.

Source: Future Fisherman Foundation, http://www.futurefisherman.org

Avoid Common Reporting Mistakes

An evaluation report is most likely to be read if it is timely and informative and if evaluation stakeholders have been involved in the process from the beginning. You may be surprised to find that the report you so carefully crafted is not widely read by your intended evaluation audience. To overcome barriers to the acceptance and use of recommendations in the evaluation report, make sure that it meets all of the needs of the report audience in advance.

Most common complaints heard in organizations when an evaluation report is released are:

- the report is late and decisions have already been made;
- the report is very thick and nobody will read it;
- the questions asked are the wrong ones;
- the information does not meet the audience needs ("You didn't tell us what we need to know!"); and
- the report is full of jargon that makes it boring and hard to understand.

The response of the evaluation team to these potential problems seems simple: make the report timely, concise (and with an executive summary), topical, relevant, and easy to understand. Regardless of the audience for which it is written, the final report must engage the reader and stimulate attention and interest with its grasp of the subject matter.

Communicate Results in the Organizational Setting

The evaluation team must disseminate, communicate, and discuss evaluation results with members of the key audiences. Frequent discussion of results is the most important factor in establishing a mutual understanding of the results and other relevant report contents and recommendations.

The evaluation team will meet to define and list the potential audience(s) for the evaluation report. As the team makes a list of report audiences, team members can also be thinking about how to best communicate with those audiences by filling in the Evaluation Planning Worksheet (see next page). Follow the example to tailor communication ideas and purposes for your particular report audiences.

As you define your report audiences, decide what forms of communication would be best for each key audience. You will certainly want to write a full evaluation report for program and organizational records, but for other audiences, whose interests may be limited to just a few of the topics covered in the full report, shorter summaries, oral briefings, conference presentations, or workshops may be more appropriate.

For example, you might want to create a separate "Executive Summary of Evaluation Results" for the organization's administrators, represent results in visually in graphics for decision makers, and develop

Tips for Writing Good Evaluation Reports

- **Start early:** A good deal of final report work can (and often does) take place before the data are collected or analyzed. Background and methods sections can be developed from the original program proposal and from material developed during evaluation planning. With parts of the evaluation report prepared ahead of time, the team will be ready to work on the results and recommendations as soon as the analysis is completed, making for a more timely report.
- **Allow enough time:** Be sure that your evaluation plan includes adequate time for the analysis, interpretation, and reporting of evaluation results. Too often, insufficient time and resources are allocated for evaluation reporting. Be sure that evaluation team members and program staff members have adequate time to review reports and prepare presentations before the information is broadly released. Allow extra time for consulting with organization public relations officers if you plan to release a Press Release about your evaluation results or program improvement plans.
- **Make the report concise and readable:** As a rule, only a fraction of the data tables and case narratives prepared for the evaluation need to be displayed and discussed in the report. One method for limiting the size of the report is to include only narrative and data that are tied together and related to the most important evaluation questions or needs for the key audience(s). The inclusion of quotes, visuals, graphs, and tables will help to break up the tedium of a technical report – as long as those visuals provide context or relate to the data being presented. Report language and terminology should be understandable and consistent.
- **Solicit feedback from report audiences:** Early in the evaluation process, the team can solicit feedback from the primary report audiences to make sure that the evaluation report will meet their information needs. More feedback can be solicited as the report is being developed, particularly from program staff and managers and the entire evaluation team through early reviews of the evaluation results.

a press release for state news outlets to reach the general public. Direct personal communication is often considered the best way to deliver results to top-level administrators and other policy makers or decision makers, so plan to deliver your summary of results in a personal meeting for best effect. Oral briefings allow the sharing of key findings and recommendations with those decision makers who lack the time to carefully review a voluminous report, and also give you an opportunity to answer any questions about the program for the decision-making audience.

The evaluation team may also wish to consider the timing of communicating evaluation results. For example, it may be valuable to share information about results with team members and program staff members first, to generate feedback and comments before the results are shared with administrators or decision makers. It is also important to share results with decision makers well in advance of any important decisions about program continuation or budgets – for example, well in advance of the annual budget planning period. In cases where you need to share information but your final report has not yet been developed, you can still give a "sneak preview" of evaluation results to key decision makers.

Evaluation Tip

When evaluation results are effectively communicated, there is a greater chance that the recommendations will be used to reinforce, improve, or modify program activities.

Another way to share information is through presentations at meetings or conferences, which often focus on sharing with colleagues the "lessons learned" and best practices that are evident from your evaluation results. In addition, conference

Report Media to Consider

A final report may be delivered in a variety of different formats – from a written manuscript to a personal conversation or an electronic presentation. However it is presented, it should be organized to show your approach, methods, results, and conclusions.

There are many ways to share your evaluation information:

- detailed written report
- executive summary of evaluation findings and key conclusions
- face-to-face oral presentation
- brochure on the principal evaluation lessons and recommendations
- annual report
- article in technical publication
- press release
- press conference or media appearance
- public meeting or workshop
- conference or meeting presentation
- electronic distribution (e.g., e-mail, website, newsletter, PowerPoint presentation)

presentations and workshops can be used to focus on special themes or to tailor messages to the interests and background of a specific audience. In some cases, it may be worthwhile publishing the evaluation report either in its entirety or in shorter versions for dissemination to a wider audience.

Seminars, workshops and discussion groups can also be organized with stakeholder or public audiences. These working sessions offer opportunities for stakeholders to hear about evaluation findings and to interpret those findings (through discussion) into a meaningful construct – in other words, these are opportunities for mutual learning among stakeholders and program managers.

Use and Monitor the Evaluation Results

Active follow up and monitoring is often necessary for program managers and administrators to implement report recommendations. At this point you are essentially managing knowledge, which involves sharing and leveraging information – getting information to the right decision makers, helping them understand it, and encouraging their action in a timely manner. Chapter 2 discusses how change happens within organizations.

Implement the Results: The evaluation results should be used both for program improvement and for organizational growth. There are many different ways in which your evaluation results can be used (see box on next page). The benefits that are derived from the evaluation may ultimately depend on how the results are distributed and used by the intended audiences.

Evaluation Reporting Worksheet

		Potential Uses of Evaluation Results			
Potential Audience	**Best Methods or Media to Communicate with Audience**	**Determine Success**	**Decision Making**	**Program Support**	**Future Practice**
EXAMPLE: Top-level administrators	Executive summary of report, oral presentation, face-to-face meeting, talking points	X	X	X	

Factors that Influence Use of Evaluation Results

- Interests: Personal interests strongly influence the policies set by decision makers.
- Ideology: Decision makers are influenced by both personal and organizational ideologies and "beliefs."
- Institutions: Decisions are made within the institution, reflecting previous decisions, organizational history, agency culture, and norms. General agency direction is set and institutional factors may constrain decisions.
- Information: An evaluation is only one of many sources of information that decision makers take into account. Advisers, colleagues, interest groups, and other sources of information may carry more legitimacy with the decision maker.

Source: Adapted from Weiss (1999)

Given the many factors that influence decision making, the prospects that your evaluation will have a strong and direct influence on a particular decision may be low. But you can increase those odds by strategically packaging and distributing the evaluation results. If you can give decision makers the right information and the right reasons to use it, your evaluation will have a strong influence on program improvement and on future decisions and policies.

Keeping in mind the different types of uses and the factors influencing those uses, the evaluation team can establish an implementation plan based on the recommendations. The plan will include a timetable and will identify who is responsible for follow-up actions and monitoring of those actions. The plan will also include the worksheet of evaluation report audiences that you filled out earlier in this chapter. The members of the evaluation team will primarily be responsible for the necessary actions in your plan, such as meeting with administrators and preparing documents or presentations about the results. Discuss how the team members might be able to influence those people to take the necessary actions.

Monitor the Implementation of Results: Over the coming weeks and months (and perhaps years), the evaluation team will monitor the status of implementation of recommended actions, and by so doing will advocate the use of evaluation results. If the evaluation report was timely, the information will be readily available to inform pending decisions. If the evaluation report was poorly timed, the team still can encourage actions by bringing up the pertinent results when related decisions are being made at a later date. Team members can have face-to-face conversations with the users of the evaluation results to promote the use of the results. Team members can continue to discuss the evaluation results in meetings or decision-making sessions when pertinent. After all of the work that went into the evaluation, the team will be motivated to advocate for the actions recommended by the analysis and results. With persistence, the evaluation team will see the fruits of their labors in the improved programs and policies of their organization.

Checklist of factors that influence the use of evaluation results in the organizational setting:

- The evaluation team is able to explain the evaluation process and build commitment among the key audiences.
- The primary intended users for the evaluation have been identified and are providing input.
- The potential contribution of evaluation results to major decisions or policies has been considered.
- The potential contribution of evaluation results to program improvement and general knowledge have been considered.
- High priority evaluation questions have been addressed.
- Potential barriers to the use of evaluation results have been identified, and the results can be used as intended.
- Primary intended users are informed of interim findings to maintain interest in the evaluation.
- Primary users are involved in helping to generate recommendations.
- Results are disseminated to intended users and the evaluation team works with intended users to apply the results in intended ways.
- The evaluation team stays in contact with the intended users to determine the extent of use of the evaluation results.
- Evaluation results are used to improve the program.
- Evaluation results are used to influence organizational decisions.

Source: Utilization-Focused Evaluation Checklist (Patton 2002)
http://www.wmich.edu/evalctr/checklists/

Uses for Evaluation Results

Direct use: A decision maker acts on the basis of the conclusions and recommendations contained in the evaluation report or presentation. This is actually a fairly rare situation.

Indirect use: A decision maker reviews the evaluation report and combines it with many other sources of information and advice to prepare a position or policy on an issue of current importance. The recommendations are selectively incorporated into a broader decision or policy. The personal advice of the evaluation team is particularly important here.

Symbolic use: The evaluation recommendations are "publicly" accepted, but are not used to inform decisions. The organization has carried out the evaluation to meet requirements or to give the appearance of being responsive, but is not actually concerned with improving performance. Fortunately, it is increasingly difficult for an organization to make symbolic use of evaluation results in a climate of public concern about performance and accountability.

Process use: This is the use of the evaluation process itself to generate participation of managers, staff members, and other stakeholders. By involvement in the process itself, the participants are changed. Although process use changes individuals' knowledge, skills, and attitudes, it also results in changed decisions and policies at the organizational level. The organizational culture begins to change from the inside out.

Case Study: Evaluation of Chesapeake Bay Foundation Programs

An example use of evaluation results for conservation education programs

Researchers conducted an evaluation of the Chesapeake Bay Foundation's (CBF) conservation education programs to determine the extent to which they promoted participants' environmentally responsible behavior and reduced teachers' perceived barriers to teaching about the bay. Outcomes of five youth-oriented and two teacher programs were assessed through pre-program, post-program, retention tests and mailed questionnaires from samples of current and past participants.

Based on improvements in characteristics that promote environmentally responsible behavior, such as perceived knowledge of issues and actions, environmental sensitivity, and intention to act, the researchers concluded that the education programs increased some youths' and many teachers' environmentally responsible behavior. They were able to show that the teacher-education programs augmented teaching about the bay. As a result of recommendations based on evaluation results, CBF implemented several changes, including focusing programs to target more specific and attainable goals, coordinating programs to provide experiences that build on one another, and conducting periodic evaluations.

Source: Michaela Zint, Anita Kraemer, Heather Northway, and Miyoun Lim, 2002, *Conservation Biology*, 16(3):641.

Chapter 5 Summary of Best Practices

- Focus your evaluation and reporting on providing useful results to the intended users.
- Use the data analysis process to help you see relationships, similarities, and differences in your data.
- Use outside contractors or partners to provide your organization with specialized data analysis and interpretation expertise, advice, and assistance.
- Use the evidence to build a strong case for the evaluation conclusions and make recommendations for specific actions of program and organizational improvement.
- Organize, report, and discuss the evaluation report in formats that are best suited for intended users and other key audiences.
- The evaluation report is more likely to be read if it is timely, easy to read, and addresses key evaluation questions.
- Active follow up is often necessary for program managers and administrators to implement report recommendations.
- Monitor the use of evaluation results and the changes that flow from the recommendations.

Chapter 6

Aquatic Education Evaluation Tools

Introduction to Tools

This chapter presents a variety of tools you can use to evaluate your programs. The descriptions of evaluation tools in the following pages are presented as "fact sheets" and focus on how to apply each evaluation tool. Each fact sheet provides an outline of the basic steps involved in using each tool, an explanation of when it's best to use the tool in your aquatic education evaluation and helpful suggestions and hints for using each tool. In addition, case studies of programs that use these tools are provided throughout the chapter.

Basic Steps for Using Evaluation Tools

Review the evaluation cycle (introduced in Chapter 1) which is followed for any type of evaluation, regardless of the tools or strategies. The fact sheets included in this chapter assume that you follow the evaluation planning process that has been explained throughout the Evaluation Guide:

- Create a Climate for Evaluation (Chapter 2)
- Develop an Evaluation Plan (Chapter 3)
- Design and Manage the Evaluation (Chapter 4)
- Analyze Data, Communicate, Use, and Monitor Results (Chapter 5)

Important Note

The evaluation planning steps are very important to the overall success of your evaluation effort. Although it is tempting to jump directly to the tools, the rest of this guide provides the framework needed to successfully use these tools to evaluate your program(s).

For each tool, you will establish a purpose, create a timeline, and keep proper documentation. After you have completed the evaluation, you will meet with the team to make any data coding or analysis decisions. The team members will then examine the evidence to reach conclusions and recommendations.

Characteristics of Aquatic Education Evaluation Tools

Table 6.1 provides a comparison of various evaluation tools, comparing their resource requirements, characteristics, and advantages/disadvantages. The table is followed by a series of fact sheets for the following evaluation tools:

- surveys
- interviews
- brainstorming/nominal group technique
- focus groups/focused conversation
- citizen advisory group/public workshop
- observation
- content analysis
- skills (performance) assessment
- case study
- stewardship monitoring
- other tools:
  - expert opinion/Delphi group
  - website evaluation and tracking
  - longitudinal study/panel study
  - internal review
  - license sales tracking
  - cost-benefit analysis
  - open house/public meeting

For additional help in selecting tools for various types of aquatic education programs, please refer to Table 4.5 in Chapter 4.

Table 6.1 Characteristics and Advantages of Evaluation Tools

KEY: ■ Best ◘ Moderate □ Worst	Resources				Capacity				Type of Evaluation				Description and Advantages
	Relatively Low Evaluation Cost	Relatively Low Time Investment	Relatively Fewer Personnel Needed	Relatively Easier to Perform	Relatively Easier to Use	Database Needed (Yes, No, Maybe)	Measures Demographic Characteristics	Measures Knowledge, Attitudes	Generates New Ideas	Provides In-Depth Information	Promotes Participation	Relatively Unbiased	
Individual Techniques													
Surveys, Mail	◘	□	■	□	■	Y	□	□	■	■	■	■	Mail surveys are confidential, low in bias, relatively inexpensive, and allow for more complex questions. They may get lower response rates than other types of surveys or interviews.
Surveys, Phone	◘	■	□	□	■	Y	□	□	■	■	■	◘	Telephone surveys offer more flexibility and speed and higher response rates in return for a higher cost and a bias against unlisted numbers.
Surveys, Web	■	■	■	■	■	Y	■	■	◘	■	■	□	Web surveys are inexpensive and have similar benefits to written surveys, but usually do not garner a representative sample.
Interviews, Informal	■	◘	■	■	■	M	■	■	◘	■	■	□	Informal interviews are inexpensive and easy. Make your existing conversations more consistent to derive useful interview evaluation data.

	Resources			Capacity			Type of Evaluation						
KEY: ■ Best ◘ Moderate □ Worst	Relatively Low Evaluation Cost	Relatively Low Time Investment	Relatively Fewer Personnel Needed	Relatively Easier to Perform	Relatively Easier to Use	Database Needed (Yes, No, Maybe)	Measures Demographic Characteristics	Measures Knowledge, Attitudes	Generates New Ideas	Provides In-Depth Information	Promotes Participation	Relatively Unbiased	Description and Advantages
Interviews, Formal	□	□	□	□	■	Y	□	□	◘	□	■	◘	Formal interviews are especially good for asking people about their attitudes and motivations, but are potentially costly and vulnerable to interviewer bias.
Group Techniques													
Brainstorming/ Nominal Group	◘	■	■	□	□	N	□	□	■	□	◘	◘	Nominal or Delphi groups are guided brainstorming sessions that can generate new ideas and promote interaction and a sense of ownership in the process.
Focus Groups	◘	■	■	□	□	M	□	□	□	□	◘	■	Focus groups are especially useful for collecting information during outreach program design stages.
Citizen Advisory Group	■	□	■	■	□	M	■	■	□	■	□	□	Citizen advisory groups often result in the same outcomes as workshops and public hearings, but result in a more supportive public due to participation in the decision process.

KEY: ■ Best ◘ Moderate □ Worst	Resources: Relatively Low Evaluation Cost	Resources: Relatively Low Time Investment	Resources: Relatively Fewer Personnel Needed	Capacity: Relatively Easier to Perform	Capacity: Relatively Easier to Use	Capacity: Database Needed (Yes, No, Maybe)	Type of Evaluation: Measures Demographic Characteristics	Type of Evaluation: Measures Knowledge, Attitudes	Type of Evaluation: Generates New Ideas	Type of Evaluation: Provides In-Depth Information	Type of Evaluation: Promotes Participation	Type of Evaluation: Relatively Unbiased	Description and Advantages
Observation and Measurement													
Observation	■	◘	◘	□	■	M	□	■	□	□	■	■	Observation yields in-depth information about events, but requires special training and may introduce observer bias.
Content Analysis	■	■	■	□	■	Y	■	■	■	◘	■	■	Content analysis describes or rates items in documents or news coverage, like the frequency or style of coverage of a specific topic.
Skills Assessment	◘	◘	◘	◘	◘	M	■	■	◘	◘	◘	◘	Skills assessment provides the opportunity to see if participants learned what you intended. Can be simple or complex depending on assessment.
Expert Opinion	◘	■	■	□	■	M	■	■	□	◘	■	□	Expert opinion can result in immediate, high-quality results, but may introduce bias if rating criteria are not formalized.

	Resources			Capacity			Type of Evaluation						
KEY: ■ Best ◘ Moderate □ Worst	Relatively Low Evaluation Cost	Relatively Low Time Investment	Relatively Fewer Personnel Needed	Relatively Easier to Perform	Relatively Easier to Use	Database Needed (Yes, No, Maybe)	Measures Demographic Characteristics	Measures Knowledge, Attitudes	Generates New Ideas	Provides In-Depth Information	Promotes Participation	Relatively Unbiased	Description and Advantages
Website Tracking	■	■	■	■	□	Y	■	■	■	■	■	■	Straight measurement supplies numbers to support assessment programs (e.g., number of web visits).
Combined Techniques													
Case Study	■	■	■	◘	◘	M	■	■	□	■	■	■	Case studies investigate a program within its natural context and provide in-depth, but biased, data.
Stewardship Monitoring	□	□	□	□	◘	Y	■	◘	■	◘	■	■	Stewardship monitoring can result in data about the effects of human behaviors on ecological indicators, and may show behavior changes.
Longitudinal Study	□	◘	◘	◘	◘	Y	◘	□	□	◘	◘	■	Longitudinal studies can provide a wealth of information about your audience over time. Can suffer from changes in staff and approach.

Type of Evaluation	Resources			Capacity								Description and Advantages
KEY: ■ Best ◘ Moderate □ Worst	Relatively Low Evaluation Cost	Relatively Low Time Investment	Relatively Fewer Personnel Needed	Relatively Easier to Perform	Relatively Easier to Use	Database Needed (Yes, No, Maybe)	Measures Demographic Characteristics	Measures Knowledge, Attitudes	Generates New Ideas	Provides In-Depth Information	Promotes Participation	Relatively Unbiased	
Internal Review	■	■	■	◘	◘	M	◘	◘	◘	◘	◘	□	Internal review can be an easy way to evaluate your program without taking a lot of time or funds. Can very easily lead to biased results.

Evaluation Tools

Surveys

Surveys are a versatile tool and may be designed for the planning, formative, or summative evaluation of all types of aquatic education programs.

Description

Surveys are a versatile tool for gathering information about the knowledge, attitudes, beliefs, intended behaviors, and characteristics of individuals. Surveys are accomplished through questionnaires, telephone interviews, or other means. If a random sample is used, the results can be generalized to the entire group or "population" from which the sample was drawn. See the detailed discussion of sampling in Chapter 3.

Benefits to Aquatic Educators

Depending on the design and sampling, surveys can be used to learn participants' views about aquatic education topics, to find out how much people gain from aquatic education programs, or to sample how strongly they support or oppose specific proposals or management schemes. The cost of survey research depends on the level of effort and formality. In any case, it is almost always appropriate to take a random sample so that your survey results can be generalized for a broader audience. Internet surveys are a new technology that can be inexpensive, but not necessarily representative.

When to Use Surveys

- To gather input from the public or from a specific group (called a "population"), such as licensed boaters or aquatic education program participants.
- To gather statistically valid information that can be generalized to a broader population from which a random sample is taken.
- To get information from people who would not normally participate in programs or provide comments about programs.

When Not to Use Surveys

- When you have a very limited budget for program evaluation, consider a small survey or a web-based survey.
- When it is not necessary to gain a statistically valid sample, in which case you can use a less costly or more descriptive evaluation tool, such as informal interviews, observations, case studies, or content analysis.
- When you need richer qualitative information about a program than a survey can supply.
- When you have a very limited time frame.

Plan and Define the Survey Sample

Identify the target audience and determine the necessary sample size. How will you get access to the addresses, telephone numbers, or e-mail addresses for your target population? Ask a survey expert to help you define the population and the random selection process. See the discussion of sampling in Chapter 3 for more guidance.

All types of surveys require close attention to question design to avoid bias. See Step 6 of evaluation planning in Chapter 3 for more discussion of question design for surveys. ***Telephone surveys*** can be completed fairly quickly and are moderately expensive. ***Mail surveys*** are a bit more expensive and require multiple follow-up mailings. Both types are good for geographically dispersed audiences.

Internet surveys that are e-mailed to a random selection of the population are easy and inexpensive, but require a keen eye to how the selection will be made. Internet surveys work best with a defined audience of people with valid e-mail addresses. ***Web-based surveys*** (e.g., surveys in pop-up windows or surveys voluntarily offered at an organization website) are not random, and thus do not provide results that are representative of the population.

Prepare and Implement the Survey

Based on the evaluation purpose, identify the key concepts to be covered by the survey. Decide what types of questions to use. Survey questions can be "closed-end" (e.g., "yes-no" questions or items on a scale of 1 to 5) or "open-ended" (What do you think of ___?). There are no right or wrong answers in survey research – the desired outcome is to capture the thoughts and opinions of the audience. Draft a set of clear and concise survey questions. Put the most interesting or important issues first, with questions about demographic characteristics last. Limit the length of the survey to less than four pages (print) or less than 20 minutes (Internet or

Case Study: Evaluation of a Camp Sportfishing Program

An example of a survey

Researchers evaluated the New York Sportfishing and Aquatic Resources Education Program (SAREP), which was incorporated into several youth camp programs to promote fishing skills, ethical angling behavior, and aquatic resources stewardship behaviors. Written pre-program and post-program surveys were administered to the camp youth by counselors to assess knowledge and attitudes of participants in the program. The survey included questions to determine the level of sportfishing knowledge (10), fish biology and ecology (8), and awareness of ethical behavior (15) and stewardship behavior (4). The camps returned 127 completed surveys. The overall mean (average) scores of camp participants significantly increased, with participants showing the most gains in sportfishing knowledge and biology/ecology knowledge. No significant change was measured in ethical/stewardship behavior awareness scores; however, responses indicated that youth intended to continue to fish in the future and that "saving the environment" was very important to them. Researchers concluded that increasing knowledge is one key factor in developing stewardship behavior, but that developing knowledge and practice of citizenship action skills is more important in encouraging the shift in behavior.

Source: http://www.joe.org/joe/2003february/rb6.shtml

telephone). Consult with a survey expert for more help in developing and refining survey questions.

Conduct a pilot test of the survey with members of the target population, at least a dozen individuals for larger surveys, and for small groups enough to show that the survey is working. Ask for feedback and revise the survey as necessary.

A higher response rate yields more accurate results with lower rates of statistical error. The response rate to surveys is directly related to the number of contacts with the subjects. Response rates are higher if the survey is announced to the population in advance. For mail surveys, using a total of three or four contacts is recommended: the advance letter or postcard, the survey tool, a thank-you/reminder postcard, and a follow-up letter with an additional survey for those who did not respond. Internet surveys generally follow a protocol with an announcement of the coming survey and several follow-up reminders. Telephone survey organizations will have an established protocol for survey scripts, training, sampling, repeat calling, and help sheets for dealing with questions or unusual situations.

Case Study: Roaring River Fish Hatchery Visitor Evaluation

An example of a visitor exit survey

The Roaring Creek Fish Hatchery is located in the Roaring River State Park and is one of the oldest hatcheries in the state of Missouri. Displays at the hatchery focus on aquaculture and hatcheries, fishing skills, habitat conservation, and development of a land ethic. Evaluators used two brief surveys to ask exiting visitors about hatchery topics and about the educational displays. The results showed that visitors taking the tour were more knowledgeable than general visitors and that pamphlets and signs were the most common media used by visitors to the site. Ideas were gained for program improvement and development of new ideas.

Source: Gregg Krumme and Janice Schnake Greene, Southwest Missouri State University

Evaluation Tools

Interviews

Interviews are useful for all types of program evaluation. Most aquatic educators are already talking to key stakeholders. Create an instant evaluation by systematizing and documenting the discussions!

Description

An interview is a structured or unstructured conversation conducted in person or over the telephone. One person (the interviewer) asks questions of another person (the respondent). Interviews yield information on respondents' knowledge, attitudes, intended behaviors, motivations, and other factors. Interviewing is a skill, and the interviewer may require training to remain unbiased during data collection. In structured settings, interviewers must be consistent in their questioning to achieve valid quantitative consistent results. In less structured settings, questions may vary as long as the evaluation purpose is addressed. In-depth interviews can yield rich qualitative data about the program. Interview results are frequently coded to provide quantitative data and may also be prepared as a qualitative summary.

Benefits to Aquatic Educators

Interviews offer an opportunity to gain a representative sample of the audience's views about a particular topic, program, or proposal, and also to explore more in-depth ideas and feelings with the respondents. For a first evaluation, most aquatic educators can simply transform their existing stakeholder conversations into "interviews" by covering and recording answers to a consistent series of questions about the program.

When to Use Interviews

- To gather information and input from the public or from a specific group (called a "population"), such as licensed boaters or aquatic education program participants.
- To gather information that is representative of the broader population from which the random sample is taken.
- To ascertain the current level of knowledge, understanding, or acceptance of a particular program or issue.
- To gather information through probing or open-ended questions that may not be revealed in other forums.
- To clarify questions or discrepancies in data collected from another evaluation tool.
- To get information from people who would not normally participate or comment.
- To get a higher rate of response than from other forms of evaluation.
- To gather information from people in a casual and comfortable setting. To build a relationship between the organization and stakeholders through personal meeting and discussion (for personal interviews).
- To gather information from a population spread over a broad geographic area (for telephone interviews).

When Not to Use Interviews

When the program evaluation budget is limited, face-to-face interviews may be too expensive. You will be limited to performing informal interviews, telephone interviews, or other less costly evaluation techniques.

Plan and Define Sample Populations

Identify the target population and determine the necessary sample size. If you want to gain information that can be generalized to the entire population, you must define a random selection process. See the discussion of Sampling Design in Chapter 3 for more information. Consult a survey professional for help with sampling for onsite personal interviews, which involves selecting a series of dates and times to capture a random cross-section of visitors at a given site or facility. For telephone interviews (or telephone survey), the research bureau will have access to telephone numbers for a given geographic area, or you can provide telephone numbers for a defined population, such as license holders.

In-person interviews are costly and require careful attention to interviewer training, but they result in in-depth information that can be useful for difficult policy or planning decisions. ***Telephone***

interviews/surveys are less expensive and good for geographically dispersed populations, but may not gain the same depth of information as personal interviews. In-person interviews can be conducted door-to-door, in public areas, or at high-visitation aquatic sites, while telephone interviews are conducted using random dialing within a target population or geographic area. ***Group interviews*** may also be performed – see the related fact sheet on Focus Groups. Interviews can be tightly structured, semi-structured, unstructured, in-depth, or conversational. For consistency and accuracy, any series of interviews will generally follow the same interview script (questions). Interview questions can be "closed-end" types (e.g., "yes-no" questions or items on a scale of 1 to 5) or they can be "open-ended" (What do you think of ___?).

Prepare and Implement the Interviews

Based on the evaluation purpose, identify the key concepts to be covered by the interviews. With the guidance of a survey professional, decide what types of questions to use. Develop a concise interview script. See Chapter 3 for guidelines on writing interview questions. By their nature, interviews usually include more open-ended questions, which are developed with an eye toward the later coding of results. Put the most interesting or important issues first, with questions about demographic characteristics last. Limit the length of the interview to less than 20 minutes.

For consistency, interview questions must be asked in exactly the same way of everybody, even though certain questions seem more interesting or relevant to certain respondents. Training and practice sessions are especially important if you are using multiple or student/volunteer interviewers. Videotape or record some of the practice interviews to provide interviewer feedback. Interview responses are tracked on paper or computer or recorded by an audio or video recorder, allowing for later data coding and analysis.

Conduct a pilot test of the interview with members of the target audience (a dozen for large groups; and for small groups enough to show that the tool is working). Ask for feedback during the pilot test. Revise the interview procedures and questions if necessary. The pilot test can also be used to develop a set of selection items for open-ended questions (i.e., you can turn an open-ended question into a multiple choice question if there is a consistent set of responses during the pilot test).

Response rates to interviews may be higher if the interviews are announced to the population in advance. For example, statewide interviews might be announced through a press release. Onsite personal interviews (e.g., at an aquatic education site or boat ramp) are probably better unannounced, to capture a random sample of the audience. Telephone survey bureaus have protocols for repeat calling and other interview characteristics.

Case Study: Media and Messages on Life Vests

An example of a telephone interview evaluation

BoatUS conducted telephone interviews to determine the effectiveness of various media and messages in encouraging the use of personal flotation devices (PFDs). The sample was a representative selection of participants (boat owners and boat anglers) from four coastal and inland states. A total of 810 interviews revealed that an overwhelming majority of boaters (95 percent) say they know about safe boating and PFDs, and nearly a third (32 percent) have taken a boating safety course. Nearly a quarter of boaters (22 percent) admitted that they relax their safety practices when they are boating with friends as opposed to family, but most boaters did not change their behavior with different groups. A huge majority of respondents (95 percent) agreed that PFDs should be worn while boating in bad weather, but only 55 percent agreed that PFDs should be worn under good conditions. Most respondents said that they use PFDs for the safety and protection of others (80 percent), while only 14 percent said they wear PFDs in response to media messages. A majority of boaters had seen ads or information encouraging PFD use. Respondents provided opinions on various scenarios of future PFD advertising campaigns.

Source: http://www.responsivemanagement.com/download/reports/BOATUS%20_PFD_Report.pdf

Evaluation Tools

Brainstorming / Nominal Group Technique

Brainstorming techniques are used for planning evaluation and anytime new ideas are needed.

Description

Brainstorming is an organized individual or group approach to generating and capturing a large number of ideas. The more ideas generated, the greater the chance of finding a creative and feasible solution. Brainstorming participants are encouraged and expected to share in the flow of ideas, without concern for interruptions, judgments, or criticisms. The key to successful brainstorming is to provide a supportive atmosphere to simulate the development of more and better ideas. A judgment or decision session can be added to the end of a brainstorming session if a solution must be reached that day.

Benefits to Aquatic Educators

Brainstorming develops creative solutions by focusing on a problem and pushing ideas as far as possible. Brainstorming can be used to encourage creative ideas for unique and improved aquatic education programs. Brainstorming techniques are best used when originality of ideas is more important than reaching acceptance or conclusion. Brainstorming requires few resources to perform.

When to Use Brainstorming

- To generate creative, original ideas for developing new or improved products or programs.
- To generate new ideas for marketing, advertising, or outreach campaigns.
- To get at root causes and alternative solutions for sticky problems.
- To find ways to improve the operation of organizational processes.
- To generate sharing and discussion among team members.

When Not to Use Brainstorming

- When expert judgment of or discrimination among alternatives is essential for making a decision. In these circumstances, brainstorming may produce solutions inferior to expert input.
- When there is a pressure to reach a decision – brainstorming in that environment is often non-conclusive.
- When a group needs to build consensus. By developing a wide range of divergent ideas, brainstorming is the opposite of consensus, which is convergent.
- When the group does not have an open mind toward new ideas.

Plan the Brainstorming Session

Identify the participants for the brainstorming session. Include between six and 20 participants with an interest in the subject or problem at hand. You do not necessarily want experts at the brainstorming session – a diverse range of participants will come up with the most creative ideas during this exercise. Have the meeting at a convenient time and comfortable location.

Brainstorming can be open ended, starting with just a simple statement of the problem, or more structured, seeking to develop ideas to address a series of questions. If the problem is large or contentious, consider having an outside/neutral facilitator lead the session. Assign a separate person to record or organize ideas for the participants, using a flip chart, dry-erase board, overhead projector, or electronic projector. If the brainstorming session is consistently documented and the ideas are reviewed and analyzed, it can be a valuable part of a program planning evaluation process.

Implement the Brainstorming Session

Participants are presented with a statement of the problem to be discussed. In some cases, they are given a question to think about a week in advance. During the session, ideas can be stated aloud or written down. One technique involves "group passing" of ideas, where each participant writes down one idea and passes the paper to the next person, who then adds thoughts or ideas to the sheet. At some point, all ideas are organized on paper, flip charts, or wall boards. Ideas should be recorded concisely but without changing their meaning. The recorder restates the idea back to the group to confirm that it expresses the original meaning. All generated ideas are kept in view. For example, as flip chart pages become full, display them on the wall.

The facilitator continues to encourage ideas to be generated without judgment, and waits at least five minutes after participants say they don't have any more ideas. The best ideas sometimes come at the end of the brainstorming session. If there is more than one topic for the brainstorming session, the group can move on to the next topic at this point. At the end of the session, the facilitator and recorder help participants organize their ideas into categories or clusters. It helps to have flexible media at this stage – such as post-it notes or cards with tape – so that ideas can be moved around into different groupings. If there is no more work for the group, thank them for the valuable ideas.

If the brainstorming session is to be followed by a judging or decision-making session, ideas can be organized in a way that addresses the purpose. At this point the facilitator can encourage discussion or criticism of the ideas and can monitor the flow of conversation to be sure that all participants have a chance to provide input. Other facilitation techniques, such as round-robin discussions or ranking/voting exercises, can be used to guide the group toward a conclusion. Smaller working groups may be able to rank the best ideas or solutions on various factors that are important to program success, as shown in the hypothetical ranking table below.

Example Brainstorming Ranking Table

Hypothetical ideas are ranked on a scale of 1 to 5 where 5 is best.

Factors → / Ideas ↓	Low Cost	Feasible/Easy to Implement	Supports Mission	SCORE
Program Idea A	5	3	1	8
Program Idea B	2	5	5	12
Program Idea C	1	3	3	7

Alternate Format: Nominal Group Technique

The nominal group technique is a structured format that generates a ranked list of ideas in about three hours. A facilitator leads the group through a specific flow of steps in the process:

1. Silent idea generation (individual brainstorming) without evaluation/judgment of ideas;
2. Round-robin sharing of ideas, where each person shares one idea and the facilitator keeps going around the room until all ideas are exhausted (ideas are clustered by the recorder);
3. Individual participant scoring of ideas, where ideas are copied onto participants' own sheet of paper and scored, then the facilitator collects and compiles all scores;
4. Group discussion, where ideas are evaluated one at a time;
5. Reassessment/rescoring of ideas by individuals; and
6. Summary/voting/ranking of ideas after revised judgments.

This technique often produces better results than an unstructured group process. The individual brainstorming, the assurance that every participant will have equal input, and the opportunity to re-evaluate ideas after the group discussion all lead to a robust conclusion.

Case Study: Creating an Educational Niche for Chinsegut

An example of brainstorming

Planners with the Florida Fish and Wildlife Conservation Commission (FWC) were developing a strategic plan for the Chinsegut Wildlife Area. The unit is focused on stewardship and recreation education and a new visitor center. FWC managers and planners convened a group of 32 area stakeholders to brainstorm activities and marketing strategies for a unique educational niche for the site. External facilitators led the group to brainstorm potential programs, to categorize those programs into clusters, and to discuss ways to market the programs to target audiences. After lunch, the facilitators and FWC planners used a rating and ranking matrix to guide development of a strategic plan for which programs would be best to pursue. Final recommendations included a unique slate of programs for Chinsegut, along with suggestions to guide the design of the new visitor center.

Source: Pandion Systems, Inc., www.pandionsystems.com

Evaluation Tools

Focus Groups / Focused Conversation

Focus Groups may be used for planning, formative, or summative evaluations. The technique is best for evaluations assessing audience needs, attitudes, and motivations.

Description

A focus group provides an opportunity for participants to share information and discuss views and perspectives on a program or issue. Focus groups originated as a market research tool to determine current social trends or attitudes. Focus groups can be held with people from inside the organization or with external audiences. Although results are not representative of the broader population, they provide in-depth information on the range of audience attitudes, perspectives, or information needs. Focus groups can provide information and new ideas for evaluation or program design. The technique is fairly easy and moderately expensive to perform.

Benefits to Aquatic Educators

Focus groups are less intimidating than either large public meetings or one-on-one interviews. Participants enjoy discussing issues and asking questions in a comfortable, small-group setting. Focus groups also provide a casual learning environment where background information or a range of alternatives can be presented for consideration before the discussion. Focus groups are cost effective for the amount of in-depth information they generate. They are especially useful for needs assessment.

When to Use Focus Groups

- To simultaneously interview a group of people, thus saving on interview time and resources and gaining in-depth information through interactive discussion.
- To encourage interactive dialogue and generate new ideas about a problem or issue.
- To establish the range of different values or perspectives within a community on an issue by including members of different stakeholder groups.
- To have an opportunity to provide background information and ask pointed questions of key stakeholders of an issue or program.

When Not to Use Focus Groups

- When a representative (random) sample of a population is needed.
- When an audience is dispersed, unless you hold several groups around a region.
- When audience members are more likely to appreciate one-on-one interactions.
- When extremely contentious issues are being discussed.

Plan and Define the Focus Group

Define the focus group audience and set criteria for participation. Are you looking for people who hold similar views, or people with a variety of different viewpoints? Consider how to recruit the most appropriate participants. If you are planning to recruit your own participants, you will need to create or acquire a database of appropriate people who can be invited to participate. Follow-up telephone calls will be necessary to finalize recruitment, with reminders in the days before the focus group. Market research firms or universities may be able to provide facilities and help recruit group members, or they may have regularly scheduled groups (i.e., pre-recruited panels). A payment is often provided for participation in a focus group. Non-monetary incentives may also be used (e.g., transportation and refreshments, take-home materials, maps, or souvenirs).

Focus groups usually take between one and two hours. To increase efficiency, several focus groups can be held in a single day. Consider the needs of the participants and whether daytime, evenings or weekends would be best. Hold the group in a convenient and comfortable place. A neutral space – away from organizational headquarters or government buildings – is best for encouraging open discussion. Rooms designed for focus groups (i.e., with one-way mirrors, which is standard practice with permission) are found in universities or marketing firms in most cities. For some audiences, a different venue (e.g., library room, business center, café, nature center, private home) may create a more relaxed atmosphere.

Focus Groups for Children

Focus groups are a useful tool for evaluating children's programs. The focus group approach must be modified when working with children.

- Questions must be age appropriate.
- Total time should be 40 to 60 minutes.
- Each group should have five or six participants of the same age.
- Participation of youngsters who know each other may increase interaction.
- If girls and boys have vastly different reactions to the topic, you may hold separate gender groups.
- Incorporate interactive elements – things to touch or do.
- Avoid dichotomous (yes/no, either/or) questions – ask participants to explain how they think or feel about an issue.
- Focus on only one or two main topics.

Prepare and Implement the Focus Group

Based on your evaluation purpose(s), identify the key concepts to be covered in the focus group and design discussion questions and a concise discussion script. By their nature, focus groups include prompts and open-ended discussion. Nevertheless, prompts are developed with an eye toward later summarization and/or coding of the results.

Limit the length of the focus group to two hours or less, covering five to 10 open-ended discussion questions. Several groups can be performed with one facilitator in a day. To protect privacy, questions about demographic characteristics are answered on a registration form and are not discussed with the group. You may want to present background information about the issue or a range of alternatives before the discussion to ensure that participants are starting with the same basic level of understanding. The group can begin with a background presentation or a brief discussion to build rapport and then proceed to the important issues at hand.

To avoid undue bias, the program coordinator and/or evaluator will be observers. It is always best to have a neutral facilitator. It is often best to have someone who is both trained/skilled in facilitation and knowledgeable about aquatic resources. The facilitator is briefed on the focus group purpose and the script in advance. The facilitator should hold at least one practice session – a videotape of the practice session can provide helpful feedback. The facilitator has several roles: (1) to make sure the discussion covers all of the key issues; (2) to make sure that input is equally solicited from all participants; (3) to further probe interesting or unusual or passionate comments; and (4) to provide non-judgmental guidance for the group or act as a referee in heated discussions.

Always make a visit to the facility in advance. Perform a test of the presentation and recording equipment. Set up the room for a casual discussion – chairs are usually placed around a large table for a focus discussion. Light refreshments assist in creating a relaxed atmosphere for discussion.

The focus group(s) must be carefully recorded for later analysis and reporting. Recording is performed by someone other than the facilitator. To increase accuracy, focus groups are usually recorded by more than one method: possibilities include note taking (paper or flip chart), audio or video recording, observation, or stenography. In a professional facility, video, observation, and note-taking are performed behind a one-way mirror to minimize distraction. While a fixed video camera is less costly, having a videographer provide moving camera work (i.e., focus on participants as they speak) can enhance information gathered from individual expressions.

Notes and recordings of focus group proceedings must be transcribed into text documents as soon as possible after completion. All observers, note takers, and facilitators should review the transcripts to enrich information recall. Once the focus group transcripts are completed, the team can meet to discuss coding of open-ended responses, which may be grouped into a dozen or fewer categories for coherent graphic display and discussion. Results will also be reported in narrative format, so it is not necessary to code all of the results. The narrative portion of the data is developed from the transcripts by team collaboration. Content analysis techniques may also be used with the focus group transcripts.

Alternate Format: Focused Conversation

The focused conversation is a format used for smaller focus groups or internal working groups that need to reach a goal, achieve consensus, resolve a conflict, or solve a particular problem. A facilitator leads the group through a structured series of four types of questions that assist participants in grappling with an issue:

1. Unbiased questions ask about facts and data, such as what participants actually saw, heard, or read about an issue. Questions include What did you see? or What have you read about this?
2. Reflective questions assess personal reactions, responses, or feelings associated with the facts. A reflective question might be What was your gut reaction to that?
3. Interpretive questions go a step further to draw out the meaning or significance to the participants, by asking What new insight did you get from this?
4. Decision questions bring the conversation to conclusion by asking the group to reach a resolution. What do you think should be done? is a typical question. The group need not necessarily agree on all of the issues to reach a resolution for further study or action, such as establishing a citizen advisory group or creating a training or outreach program.

The focused discussion is different in that it seeks to build some level of consensus and to bring a group to resolution. It is more often used for internal organizational discussions rather than for stakeholder or audience assessment.

Case Study: Increasing Participation in Sportfishing

Example of focus group evaluation and development of new marketing strategies

The New York Sea Grant program wanted to identify the social and psychological factors that influence sportfishing participation. Preliminary mail surveys and telephone interviews were used to identify key factors. Following this round of research, two focus group sessions were held in March 2002 for 31 people involved in the sportfishing industry, including state agency personnel, business owners, and tourism promoters, to identify strategies to increase sportfishing participation. The focus sessions had two parts: a presentation of the results of the preliminary study, and a discussion of potential sportfishing management and promotion strategies.

Focus group participants identified new promotional strategies and mentioned a number of national programs that were successful in promoting sportfishing, including Water Works Wonders, Kids All-American Fishing Derby, Camp Fishing Initiative, Fishing Tackle Loaner Program, Hooked on Fishing – Not on Drugs, and Wonderful Outdoor World Program. Other strategies identified at the focus sessions included: expand or establish sportfishing mentor groups, establish a "fishing celebrities" program, increase or improve fishing access, encourage girls to participate in fishing, focus on developing fishing skills in children, encourage family fishing events, and encourage businesses to provide family fishing incentives.

Source: http://www.seagrant.sunysb.edu/ifishny/

Evaluation Tools

Citizen Advisory Group / Public Workshop

Citizen Advisory Groups are best used in planning or formative evaluation for programs that involve a variety of stakeholder groups.

Description

A citizen advisory group engages stakeholder participation in tasks or decision-making processes. The group generally includes a fixed number of members who meet on a regular basis to provide advice, complete a specific task, develop a report or product, or to supervise implementation of a plan or program. Members are carefully selected and their roles and responsibilities are defined when the group is established. Through the process, citizen advisers become much better informed about issues and develop an appreciation and understanding for the sponsor organization. Advisers develop an enhanced sense of stewardship through group service and contributions to organizational decisions. The new understanding, appreciation, and trust are carried back into the community by the advisory participants.

Benefits to Aquatic Educators

Citizen advisory groups can be used to involve program stakeholders in providing guidance, input, and review of evaluation plans. Citizen advisory groups can also be used to involve stakeholders in providing feedback on program plans and to involve stakeholders in program implementation. For the purposes of evaluation, a diversity of citizen advisers can provide valuable input toward creating unbiased evaluations.

When to Use Citizen Advisory Groups

- To gain valuable evaluation guidance from program stakeholders.
- To ensure that evaluators are informed of community interests as represented by the advisory group members.
- To provide cultural, racial, or gender input toward avoiding biased evaluations.
- To engender positive public awareness of the organization and the program evaluation process by providing stakeholders a forum for real and meaningful input.

When Not to Use Citizen Advisory Groups

When general public involvement is needed or when input that is representative of the public's view is required for an evaluation. Only randomly selected advisers can be representative of a larger population.

Determine Citizen Advisory Group Participants

Determine how advisors should correspond to various stakeholder or audience groups. Should advisers represent specific groups (e.g., anglers, boaters, teachers, recreationists), diverse perspectives (e.g., resource users, conservationists, developers), or certain demographic characteristics (e.g., geography, age, ethnicity, gender)? Define the advisory group's role in the decision-making process. Provisions may need to be made for advisory group members to check back with their sponsoring organizations before decisions or recommendations are made.

Implement the Citizen Advisory Group

A charter is developed to summarize the purpose of the group, the group's relationship to the evaluation process, the composition and size of the group, the selection of group leadership, the discussion and decision-making techniques or rules to be followed by the group, the meeting frequency and duration, the roles and responsibilities of group members, and the relationship of the group to the broader public. It is important to state how the input of the group will be used in the evaluation or program planning process. The group should have a name that reflects the main task or assignment. Citizen advisory groups generally need an organizational support team to handle meetings, logistics, and administrative support.

To be most credible in the public eye, the advisory group should be autonomous from the sponsoring organization, by meeting in a neutral space and using a non-organizational facilitator. Advisory groups sometimes elect leadership from within their own ranks. Facilitation of the advisory group typically includes a variety of techniques, such as

brainstorming, nominal group technique, polling or voting, consensus, and focused discussion. The facilitator works with participants to create a comfortable environment, establish a group mission, encourage active listening and creativity, ask appropriate questions to lead the group forward, and assess progress via the group's work calendar.

Minutes of meetings are necessary to document the activities of the group. Advice or recommendations made by the group to the sponsoring organization is formally transmitted through written (paper or electronic) media so that it is available in published format for future reference. When the group has provided their final input, the evaluator(s) discusses the results of the advisory process and decides what actions should be taken based on the feedback or recommendations.

Alternate Format: Public Workshop

A workshop is a public forum where participants work together on an assignment to provide specific input to a process, such as evaluation or program planning. Participants may work in one group or move to different groups throughout the event. Small groups may be facilitated or self-facilitating. Assignments may include sharing perspectives, identifying issues, developing ideas, ranking alternatives, or commenting on a draft evaluation or program plan. Each group documents their work and presents their work to the entire assembly (verbally or in writing) before the end of the forum. Invite the appropriate and necessary participants to the workshop, whether it is a certain group of stakeholders or residents of a given geographic area. If the workshop is to consider a contentious issue, it can be designed so that small groups have a diversity of opinion, rather than small groups being comprised of people from a single perspective. Expert facilitation may be necessary for particularly contentious issues. A workshop may also be presented as one portion of an open house or other format.

Case Study: Biomonitoring for Ecological Complexity

An example of citizen advisory role in selecting indicators for stewardship monitoring

The Department of Energy (DOE) is faced with determining cleanup standards and stewardship options for contaminated lands. There is need for a monitoring plan that can aid in remediation decisions and evaluate the success of remediation, restoration, and stewardship. The DOE Ecological Health Group set out to develop a set of bioindicators for the Savannah River Site (SRS) that could be used to evaluate risks to human and environmental health and could be a source of information on the success of remediation and stewardship activities.

The biomonitoring plan was developed in collaboration with SRS scientists and with input from a variety of stakeholders, including the DOE, U.S. EPA, state regulators, CDC Health Effects Subcommittee, and the Citizen Advisory Board for SRS, which included fishermen and hunters, other recreationists, and the general public. Stakeholder input was crucial in selecting monitoring indicators of relevance to the public, as well as to scientists and regulators.

Source: Joanna Burger, CRESP, http://www.cresp.org/dcwrkshp/posters/biomont2/biomont2.html

Evaluation Tools

Observation

Observation techniques are useful for recording behaviors for formative or summative evaluation of participatory programs, particularly if an unobtrusive technique is needed.

Description

Observation involves directly and unobtrusively witnessing people or events in their natural setting, and measuring behavior as it unfolds. The "observer," whether human or mechanical, generally does not influence the people or events being observed. Observation techniques can contribute to both qualitative and quantitative evaluations. Observation can also be used to inventory physical facilities or resources. Observation is a fairly inexpensive evaluation technique.

Benefits to Aquatic Educators

Observation is a flexible, low cost, unobtrusive evaluation technique. It directly measures behaviors, which are often of interest to aquatic educators. Because the observation subjects are not randomly selected, observation cannot be representative of a larger population. It can, however, be definitive for a given group, such as participants in a class or workshop. Observation may also be useful for measuring behavior in the environment, such as fish catch (creel surveys) or trail use (electronic counters).

When to Use Observation

- To understand any real-life behavior, process, or situation.
- To measure the behaviors of program participants, where it may provide greater insight into actual behavior as compared to surveys or interviews, which can only ask participants about their intended behaviors (notoriously inaccurate).
- When you need an unobtrusive evaluation technique or a technique that can be used without approval (as long as you do not attempt to influence behavior).
- To understand unexplained problems with the implementation of your program.
- To record and evaluate the details of a certain process or event.
- To visually assess physical products, outcomes, or other evidence of program success or to take a physical inventory of facilities or resources.
- To evaluate one individual or a small group at one time.
- To add an unobtrusive evaluation technique for triangulation with interviews, surveys, or other more intensive techniques.

When Not to Use Observation

- When you need insight into the thoughts, emotions, or motivations behind behaviors – conduct interviews or surveys to ask people about their motivations.
- When you need a sample (random) that can be generalized to a broader population.
- When the observer has strongly biased views about or a close relationship with the situation or people being observed.

Plan and Define Observation

Consider whether your observations will take place at one site (e.g., a classroom or visitor center) or at multiple sites (e.g., all the boat ramps in the state). If you are observing at a large number of sites, consider making a random selection of sites, dates, and times to create results that can be generalized to the behaviors going on at those sites. See the discussion of sampling in Chapter 3 for more details.

The other sampling decision with observation is how often to measure the target behaviors. For example, if a behavior is very common, an observer might choose to record the behavior at a random time (e.g., after a random number of minutes) or at systematic time (e.g., between 23 and 35 minutes after each hour). This is called the "snapshot" method because you are taking a mental photograph of the individual or group. In fact, you might take an actual photograph at the time of the observation to duplicate your written record (don't rely entirely on photographs – always take notes as well). If a behavior or event is less common, the evaluator can simply count the behaviors as they occur. For groups, it is useful to record the number of people engaged in a behavior and the total number of people observed, giving you a proportion

(percentage) of those involved in the behavior at each observation.

The observer will also choose a number of different approaches to observation:

- **Participant vs. non-participant observation:** Are the evaluators part of the situation they are studying (e.g., teachers observing their own classes)? The drawback to participant observation is that the evaluator tends to be biased.
- **Obtrusive vs. unobtrusive observation:** Can the subjects being studied detect the observation? In obtrusive observation, the evaluator announces herself and her intentions, which usually leads to some small participant behavior changes. Hidden cameras, garbage audits, or electronic counters are examples of unobtrusive observation. Unobtrusive observation may also be done in "disguised" or "covert" form, such as when an observer pretends to be part of a tour group.
- **Natural vs. contrived settings:** Is behavior observed when and where it is occurring (called a "natural setting"), or in a contrived setting where a situation is created to speed up the process?
- **Structured vs. unstructured observation:** Is a formal checklist used to guide the observation? Structured observation leads to quantitative data and robust evaluation results. Unstructured observation, also called "informal" or "exploratory" observation, is done when an evaluator has little knowledge of expected behavior or is interested in creating a hypothesis about how people will behave.
- **Direct or indirect observation:** Is the behavior observed as it happens (directly) or after the fact (indirectly), as in viewing videotapes from a visitor center?

When people are being observed (whether they are aware of it or not), ethical issues must be considered by the evaluator. See the discussion of ethics concerns in Chapter 4.

What to observe depends on the evaluation purpose. For example, if you are trying to show the success of a casting training program, you will want to observe casting skills in participants near the beginning and then again at the end of the program. Some program components that may be of interest include the following:

- participant characteristics, such as gender, age, ethnicity, skill, verbal statements;
- participant interactions, such as demonstrated interest, level of participation, group problem solving, level of cooperation, level of support;
- nonverbal behaviors, such as facial expressions, gestures, postures, expressions of interest or commitment;
- behavior of program leaders, such as clarity of communication, response to questions, leadership skills, encouragement of participation, awareness of group dynamics, flexibility, adaptability, knowledge of subject, use of educational aids, use of learning techniques, and activity sequencing;
- physical surroundings, such as facilities, learning climate, seating, presence of amenities;
- program products, such as demonstrations, facility development, materials.

Prepare and Implement the Observation

You may select from a variety of technologies to complete the observation, depending on the evaluation resources. Techniques may include paper-and-pencil, mechanical (handheld) counters, calculators (press +1 each time a behavior occurs), audio or video tape, still camera (photographs), automatic counters, and other technological innovations.

Sample Field Day Observation Sheet

Location _____ Date _____ Time: _____

(Make a mark for each occurrence.)

1. Number of people who stopped and looked at the exhibit: _____
2. Number of people who asked a question: _____
3. Number of people who actively used the interactive features: _____
4. Number of people who took brochures: _____
5. Did people seem to struggle to read the exhibit? ❑ Yes ❑ No
6. Were staff members always available? ❑ Yes ❑ No

Comments, questions, problems, notes:

_____

_____

Structured observations require a checklist to guide observations. Because time is often short during an event, it helps to have the checklist ready so that observations may be quickly and unobtrusively recorded. Codes are often developed in advance for the most important or common behaviors to maximize observation time and minimize writing time. An observation sheet can also include some overall questions or ratings for the observer to address after the session is over. It may be helpful to make several pilot observations of the target audience before developing or finalizing the checklist. To make observations more accurate, consider having two people observe the same situation. If you are using multiple sites with multiple observers, you will want to develop a tested checklist and train your observers for consistency.

Case Study: EstuaryLive 2005 Evaluation

An example of a classroom observation evaluation

The National Estuarine Research Reserve Association performed a 2005 evaluation of EstuaryLive, an annual live web-based broadcast designed for teachers to use in the classroom with students. The evaluation included surveys and student assessments. To enrich the information received through the other evaluation methods, the team performed classroom observations during the broadcast in several classrooms around the country. A classroom observation checklist was designed to prompt volunteers to observe certain things about how the broadcast was presented to students and how technology was used in the classroom. The observations provided in-depth information about some of the technological problems faced by educators during the broadcast. The results are being used to improve the application of technology for the interactive broadcast.

Source: Pandion Systems, Inc., www.pandionsystems.com

Evaluation Tools

Content Analysis

Content Analysis techniques are useful for planning, formative, or summative evaluations of programs that include "documents" or have a long history to examine.

Description

Content analysis is a systematic technique for the analysis of the substance of a variety of sources or documents. Content analysis enables evaluators to sift through large volumes of information and reduce them to a meaningful data set (quantitative) or narrative description (qualitative). A "document" is any symbolic representation that can be reviewed for analysis, including print and electronic media, audio media, visual media (photographs, video), letters, artifacts, databases, field notes or diaries, educational materials, advertising content, and all electronic media. Despite the name, content analysis also looks at things in context – not just what was said, but the meaning of how and why it was said.

Benefits to Aquatic Educators

Content analysis is a useful technique to discover and describe the focus of individual, group, institutional, or social communications or records. It is often used to review interview transcripts, summarize the responses to open-ended questions, assess student portfolios, evaluate the history of organizational communications, or assess news coverage of an important topic. Content analysis is a systematic, unbiased, and inexpensive evaluation technique.

When to Use Content Analysis

- To examine large volumes of data or publications for specific information on an important topic or program.
- To accomplish an unobtrusive or non-interactive evaluation, such as, when analysis is needed for a group where personal contact would be intrusive.
- To prepare for a publicity campaign by reviewing news media coverage to assess frequency, accuracy, and bias of coverage.
- To evaluate organizational communications on a certain program.
- To assess a longitudinal or historical series of documents for trend analysis.
- To assess how certain issues are presented to audience members in segmented communications, such as how fishing is presented in women's vs. men's magazines.
- To create both a quantitative and qualitative evaluation based on a set of documents.

When Not to Use Content Analysis

- When an issue or topic is ambiguous or not clearly defined.
- When only poor quality documents (e.g., damaged, ambiguous) are available.
- When you need to understand why something happened, in which case you will have to ask people directly for their thoughts and motivations.

Determine Document Sources

Content analysis generally involves searching the content of documents for occurrences of certain words/phrases, for underlying meanings, or for patterns or information trends. The process involves categorizing and coding the data so that it can be further analyzed, allowing you to see not only what is actually there, but to infer meaning from the content/context.

Decide which documents are to be analyzed and how that sample, or "population," is defined. Select a population of documents that are likely to give you information to address the evaluation purpose(s). Define the type of documents and the dates of the media. For a long-term subject, you will want a spread of years or decades. For a recent topic, you may need a few months or weeks of documents. Online databases and organization files are good sources of documentary evidence. Don't forget about electronic files – much information is now stored in electronic formats.

In advance of performing the content analysis, design a coding scheme for items of interest. Each point to be coded will allow for either a text or numerical code, producing a spreadsheet database with both qualitative and quantitative entries. Items in the sheet might include a document source, year of publication, the occupation of the author, the number of times a key topic is mentioned, the context within which the key topic is mentioned, and the length of the article or document. Numerical codes are easier to analyze, so design

coding schemes that allow for "yes-no" (1 or 0) codes or that ask for codes on a scale of 1 to 5. For example, *On a scale of 1 to 5, how positive was the coverage of the salmon recovery effort?* Other units of measure for written documents might include number of column inches, number of words, or how many times a certain word appears (often expressed as a ratio of number of times:total number of words).

The coding schedule should be written out so that it is unbiased. Consider ways to avoid double measurements from the same source. For example, if you are searching newspapers, skip Sunday sections that contain weekly news roundups. To increase triangulation, analyze two or more document sources, for example, tracking a key issue or theme across several different media. Consider also using another program evaluation tool in addition to the content analysis. Refer to the discussion of accuracy and bias in Chapter 4.

Case Study: Assessment of Great Lakes Fisheries Education

An example of a content analysis needs assessment

The researchers performed a review of opinion surveys, Great Lakes curriculum materials, and other education programs in an effort to identify education needs and opportunities for the area. On the basis of the content analysis, the researchers developed a set of Great Lakes education literacy goals, a review of existing education materials, and an identification of information and education gaps and needs. They also identified potential funding partners for Great Lakes ecosystem education programs.

Source: Michaela Zint and Rosanne Fortner, Great Lakes Fisheries and Ecosystems Education Networking Conference, Great Lakes Fishery Trust

Perform the Content Analysis

If multiple people are coding documents, hold a session to pilot test the coding scheme on a dozen documents to increase accuracy. Revise the coding scheme and test it on several additional documents if necessary. It is acceptable to make revisions to the coding scheme after you have begun coding the documents, in response to some unexpected characteristics, as long as you go back to update previously coded records.

Several problems may occur in your selection of documents. It is important to not only record the content analysis data, but to also record the reasons for any missing or unrecorded content, such as the following:

Missing documents: If a substantial proportion of the documents from the selected population are missing (say over 20 percent), you must abandon the analysis of that document source, because it will produce inaccurate data.

Inappropriate documents in population: If inappropriate records or documents (ones that do not match your definition) are included in the source material, eliminate those documents from the analysis and make a record of the reason(s) for elimination (e.g., "Document 'X' was eliminated because it did not pertain to watershed conservation in Arkansas").

Ambiguous or damaged content: Some documents in the source population will match your definition but will contain missing passages or ambiguous content, so you will make a record of the reason(s) for not coding those documents (e.g., "Document 'X' was not coded because of coffee stains over pertinent content").

Case Study: Coverage of Zebra Mussels in North America

An example of a news coverage content analysis

Zebra mussels are an eastern European species that is now found throughout the Great Lakes, down the St. Lawrence Seaway, along Ontario's Trent-Severn Waterway, throughout New York's Erie Canal-Hudson River, in the Mississippi River and five of its tributaries, and in at least nine inland lakes and reservoirs. Much of the North American continent is eventually expected to be colonized by the mussels, which produce colonies that cover nearly every solid surface and were causing over $5 billion in damages in the early 1990s.

Researchers conducted a content analysis to examine the coverage of zebra mussels in newspapers in five major cities, looking at all of the news items containing the words "zebra mussel" between 1988 and 1993. Zebra mussel coverage was found to be most closely related to geographic proximity of the infestation. In addition, researchers noted that coverage receded from the news as the shellfish came closer to the city where the newspaper was published. The weakening of coverage over time was postulated to be a means of protecting local newspaper circulation by softening bad news, and has the negative impact of leading audiences to believe that local environmental repercussions are negligible, while environmental problems far from home are more urgent. These results suggest that aquatic educations should intensify education and outreach efforts to keep critical local environmental issues in the public eye.

Source: Donny Roush and Rosanne Fortner, http://uidaho.worldcat.org/title/newspaper-coverage-of-zebra-mussels-in-north-america-a-case-of-afghanistanism/oclc/657343151&referer=brief_results

Evaluation Tools

Skills (Performance) Assessment

Skills Assessment techniques are best used for summative evaluation of recreation, outdoor adventure, stewardship, youth development, or non-formal environmental education programs.

Description

Participant performance of any critical thinking, leadership, or physical task may be evaluated against established criteria or against the earlier performance of that individual. The criteria are based on the aspects of performing the task and meeting the learning objectives. Samples of participant performance – descriptions of the performance or actual products (i.e., photographs, journals) – are often included in the evaluation report to demonstrate the link between performance of skills and overall program success.

Benefits to Aquatic Educators

Skills assessment is ideal for assessing any physical training program, such as fishing or outdoor skills training. Properly designed, the assessment provides an unbiased measure of physical performance, which can then be used as evidence for program success or program improvement. Skills assessments can be done at a quick and informal level, or at a more formal level for long-term training programs.

When to Use Skills Assessment

- To evaluate any program that seeks to develop improved critical thinking, leadership, or physical skills in participants.
- To demonstrate the success of a skills training program by measuring the performance of individuals in the program.
- To develop a qualitative "success story" or case study about a skills training program.

When Not to Use Skills Assessment

- When low self-esteem or competition between participants might be a problem, the skills assessment must be designed so that participants are compared only against set criteria or against themselves and not against other participants.

Define Assessment Criteria

Specific learning outcomes are identified for each skill being taught. Once learning outcomes are identified, assessment criteria can be established for the performance of that skill. The criteria might include specific physical activities that address the components of the skill, such as strength, endurance, or flexibility, or targeted skills, such as casting, fly-tying, or water quality measuring. Criteria may also include behavioral or problem-solving objectives for leadership or critical thinking skills. For each activity, the criteria will indicate what evidence is needed to show that the participant has met the learning objectives. Evaluators may want to create a scoring rubric (see tip box) for more unbiased scoring of participants.

Plan and Prepare the Skills Assessment

Two or more sources of evidence are generally used to assess participant performance. In informal settings, the assessment can be presented as a demonstration of what participants have learned, or performed as an unobtrusive assessment (i.e., through photographs or observations). Participants in non-formal programs are generally not subjected to a lengthy formal performance assessment.

A variety of techniques are used for skill performance assessment. The technique(s) selected will depend greatly on the format of the skills training program. For example, one-day programs will use quick or unobtrusive assessment tools, such as photographs, observations, short reflective writing exercises, or demonstrations; long-term programs may use planning worksheets, journals, periodic demonstrations, or portfolios.

- **Observations** may be used by training leaders or evaluators to assess participant skills. Observation is a useful assessment tool for one-day workshops, and can be unobtrusively performed while participants are demonstrating their new skills at the end of the session. See the fact sheet on Observations for more information.
- **Photographs** can be used much like observations, to provide a "snapshot" of participant skill development. Photographs can be taken of participants demonstrating their skills at the end of a workshop, and might also be taken at the beginning of a workshop to provide a before-and-after comparison. If photographs are used, pictures will be taken of all workshop participants, so that there is no bias in the selection of subjects for the

Example of a Scoring Rubric

For unbiased skills assessment, evaluators may find it useful to create a "rubric" (scoring guidelines) that includes descriptions of the different levels of performance. This is a sample rubric with scores ranging from 5 to 1 – the actual rubric should be as specific to the activity as possible. For example, a performance rubric for a youth fishing training program might include ability to cast, casting distance, and duration of casting practice as key criteria.

5 – Substantially Exceeds Criteria: Participant's performance meets and exceeds activity goals. Participant has mastered the skill. Participant's journal is up to date and includes insightful comments about the experience of learning the skill.

4 – Above Criteria: Participant's performance matches the activity goals. Participant can perform the skill. Participant learns from making mistakes. Participant's journal is up to date and identifies difficulties and successes in learning. Participant may need some further training support.

3 – Meets Criteria: Participant's performance sometimes meets activity goals. Participant can perform the skill with support from leaders. Participant's journal is partially up to date. Participant will need ongoing training support.

2 – Somewhat Below Criteria: Participant's performance does not often meet activity goals. Participant can perform the skill only with constant guidance. Participant shows little interest in practicing. Participant's journal is incomplete.

1 – Well Below Criteria: Participant cannot perform or participate in the activity. Participant shows no interest in performing the skill. Participant's journal contains no entries. Participant requires ongoing support to meet minimum criteria.

This sample includes the additional evidence of a participant journal. Other forms of additional evidence might include photographs, leader observations, self-assessments, or interviews. Although not required, additional evidence helps triangulate the data.

photographic record. Ask for permission to photograph by having participants sign a photo release, especially if the photos will be used in reports and presentations.

- **Informal interviews** give participants an opportunity to reflect on the workshop and give leaders a chance to assess the participant's knowledge, attitudes, and feelings about the skill training. Interviews might also be used before a training program to assess the participant's learning needs. See the fact sheet on Interviews.
- **Reflective writing and art exercises** are used to generate creative participant reflection and self-assessment. This tool can be used as a quick assessment at the end of a one-day workshop, or as part of an ongoing journal or portfolio development.
- **Skill demonstrations** can follow the format of a competition, a demonstration, or a test of skills.
- **Planning and goal-setting worksheets** are a motivational strategy for setting individual participant goals for the skill training. The sheets can include reflections on physical interests and abilities, and will list the participant's short- and long-term goals. The goal statement will be dated and included in the participant's journal or portfolio.
- **Participant journals or worklogs** are useful for performance assessment in long-term training programs. Journals encourage reflection and increase motivation of participants. Journal entries may be a broad review of the events of each training session or structured to address a specific topic. Participants may also ask questions, celebrate successes, or identify challenges to learning the skill in their journals.
- **Portfolios** are a purposeful collection of long-term work that demonstrates the participant's efforts and tasks in the training program. Portfolio

entries are dated so that leaders can track participant development and achievement over time. The portfolio may include any of the other tools discussed above, such as participant goal-setting worksheets, photographs, journal entries, reflective writing, art work, or any other items that reflect on the participant's involvement in learning the skill.

Implement the Skills Assessment

Successful skills training programs inform participants of the desired learning outcomes, provide demonstrations of the skills being taught, teach the various aspects of skill performance toward the desired learning outcomes, and allow participants plenty of time to practice the skill before performance assessment takes place. It helps if the training program has several different levels or steps involved in learning the skill, as participants are more likely to be willing to master one level and then move on to the next level.

Using the selected assessment techniques, implement the assessment before, during, and after the training program, as appropriate. Give participants ample opportunity to demonstrate their skills. Record all assessment results in written, visual, photographic, or portfolio forms for later analysis. If appropriate, some record of participation or assessment results – perhaps a certification of completion – may be offered to participants at the end of the program. Journals or portfolios or other documentation are returned to participants after the analysis is completed.

Reflective Writing or Art Prompts

The most surprising aspect of this activity/project for me was _____.

I would like to find out more about _____.

If I were to do this activity again, I would _____.

What I enjoyed most about this activity was _____.

Here is a drawing that represents how I feel when I am having trouble doing this activity.

Here is a drawing that represents how I feel when I successfully perform this activity.

Participant Journal Prompts

Today we learned _____.

I tried to _____.

I asked _____.

I found out _____.

I wish I had _____.

One question I have is _____.

My most successful activity was _____.

I had the most trouble with _____.

I figured out how to solve my problem by _____.

The resources and people I used to help were _____.

Evaluation Tools

Case Study

Case Study techniques are a qualitative tool used for formative or summative evaluations. The case study is a rich program narrative.

Description

Case study is an evaluation tool that examines a program or event in depth, within a "real life" context. Case studies offer a detailed contextual analysis of events or conditions, and can be used alone or to add strength to other evaluation techniques. Case studies may report program processes, present program outcomes, and/or present a critique geared toward program improvement. While they are generally qualitative in nature, case study reports may also incorporate quantitative data to make a point about a program. Case studies may examine multiple events or programs to identify patterns and to contrast/compare cases. Case studies are not representative of entire populations or programs, but can be used to generalize events similar to the ones being studied.

Benefits to Aquatic Educators

Case studies are useful when you are interested in evaluating program and organization processes and explaining the causal links in real-life situations, rather than evaluating the technical aspects of a program. Case studies are useful in answering questions of "how" and "why" something happens. Case study is an inexpensive, systematic, informal, and enjoyable evaluation tool!

When to Use Case Study

- To describe or tell the "story" of a program or of program participants.
- To completely depict the experiences, processes, and lessons learned from a program.
- To explain complex causal links (how and why something happens) within real-life situations, programs, or events.
- To explore situations in which the program or event being evaluated has no defined set of outcomes or objectives.
- To powerfully portray a situation, study, or event to others.
- To describe the real-life context in which the program or event has occurred.

When Not to Use Case Study

- When a random sample is needed to represent your target audience or program population.
- When a more structured evaluation is needed, for example, if quantitative data are required, if specific variables need to be measured, or if strong evidence is needed to justify the continuation of a program.

Define Case Selection Criteria

The case study purpose is established by asking questions about the situation, such as "How…" or "Why…" something happens. Case study questions may be targeted at a single event or at multiple events and their interrelationships.

With questions in hand, determine which of your programs or activities you will review for the case study. If you plan to examine multiple activities, you may want to represent geographic regions, different sizes of events, different types of participants, or other features. Or you may want the selection to be random so that the case study is representative of the overall program. You can also review one program or event and report that as a case study. Refer to the case study questions to focus attention on where to look for evidence that will satisfy the purpose of the evaluation.

There are several different types of case studies. Select the type of study that best fits your evaluation purpose and case study questions. All of these case study types may be applied to a single case or to multiple cases:

- **Exploratory case studies** often are used as an information-collecting prelude to a more in-depth evaluation.
- **Explanatory case studies** seek to explain how or why something happens, to discern which factors contributed to which outcomes.
- **Descriptive case studies** portray the events or processes of a program.
- **Snapshot case studies** provide a detailed view of one case or program at a single point in time.
- **Comparative case studies** are performed on two or more cases for the purpose of cross-case or cross-unit comparison.

- **Pre- and post-case studies** examine the situation before and after an event or program.
- **Longitudinal case studies** look at a case or multiple cases at several points over time.
- **Patchwork case studies** are a collection of several different types of case studies performed to provide a more expansive view of a single program.

There are two main principles for collecting case study evidence: (1) use multiple sources of information and (2) create a case study database or "chain of evidence." Because no single source of evidence can tell the whole story, case studies examine multiple sources of evidence to provide higher accuracy and "triangulation" of the evidence. A variety of evidence sources, including documents, archives, interviews, observations, or physical artifacts, may be used to complete the case study. Each source of evidence provides part of the story; all of the evidence together provides a more complete story of the program.

Implement the Case Study

Exemplary case studies are conducted by well-prepared evaluators. If multiple evaluators will be preparing the case study, establish procedures in advance of the field work and conduct a pilot study with a source of evidence in advance to discover any barriers or problems with the planned evaluation. Evaluators should be able to ask good questions, be good listeners and observers, interpret and record responses to questions, be adaptive and flexible, have a firm grasp on the evaluation purpose, and be unbiased.

Keep the case study evaluation as simple as possible and focused on the evaluation questions and purpose. Because case study research generates a large amount of data from multiple sources, systematic organization of the data is important to prevent overload and loss of focus. Use field notes and databases to catalogue case study data so that it is readily available for analysis and interpretation. The field notes chronicle emotions, questions, testimonies, stories, and illustrations which will be used in later reports.

Analysis of case study data may be as simple as writing the story of the case or as complex as placing data into arrays or matrices to search for connections or relationships. At the very least, the case study team will endeavor to create some graphic representations of the case study data, such as flow charts, illustrations, or other displays. It is painless to tabulate the frequency of events and to present data from observations, interviews, or other evaluation tools used during the case study. Quantitative data collected during the case study provides support for the qualitative "story" and rationale developed for the case.

Exemplary case study reports take a complex issue and transform it into one that can be easily understood by the audience. Case studies can be reported as separate stories or chapters, in chronological order, or as one large body of information. A solid focus on the main evaluation questions and purpose will ensure that the case study report is relevant to the evaluation audiences. The story of the case will effectively convey the experience to the reader, while citing a number of different sources of evidence to bolster the conclusions about program outcomes and success.

Case Study: High School Aquaculture Programs Success Story

An example of a case study presented at an annual fisheries meeting

In Mobile and Baldwin counties, Alabama, there are six school aquaculture programs. These programs use hands-on applications to teach biology, chemistry, physics, math, statistics, water quality, and physiology. Students demonstrate an improved ability to retain and apply the subject matter. Outside of scholastics, students learn responsibility, teamwork, and self-confidence as they build the systems.

Most schools begin raising tilapia in small recirculating systems. As students and teachers become comfortable, additional species can be added. Species that have been grown in Alabama include: tilapia, gambusia, cobia, red snapper, rainbow trout, Australian red claw, pacific white shrimp, and oysters. Program funding comes from grants and donations.

As new programs are established, cross-program interaction is increasing, expanding students' exposure to different species and techniques. The future of school-based aquaculture in Alabama will include access to information from other programs, as well as a central location for teachers to derive ideas and learn from one another.

Source: P.J. Waters, Auburn University Marine Extension and Research Center, Mobile, AL

Stewardship Monitoring

Stewardship Monitoring is the "ultimate" tool for formative and summative evaluation of stewardship program impacts on participants and the environment.

Description

Aquatic stewardship can be defined as *taking personal responsibility to sustain and enhance resources, while accepting an obligation to the environment and future users* (RBFF). Environmental problems require an understanding of how people think, what they care about, and how they behave. Most aquatic educators are attempting to promote stewardship or conservation behavior with an eye toward improving compromised or damaged natural systems. Because stewardship programs are designed to have positive environmental impacts, there are two steps to evaluating the success of stewardship programs: (1) demonstrating that improved stewardship behavior results from the program, and (2) showing that improved resource quality results from the improved behavior. Stewardship monitoring is more of a "process" than a tool – it usually incorporates several of the tools that have already been mentioned (e.g., surveys, interviews, observations) along with biological monitoring of environmental impacts.

Benefits to Aquatic Educators

In addition to evaluating the success of the program itself, aquatic educators are able to monitor the impacts of their stewardship education programs on natural resources. The evaluation can show which stewardship education approaches are most successful and may also reveal what motivates people to participate in improved stewardship behavior. Stewardship monitoring can be a complex and expensive evaluation tool to implement, but it can be accomplished with proper planning. The benefits to the program and the organization can be outstanding.

When to Use Stewardship Monitoring

- To evaluate programs that include stewardship objectives and activities, such as behavioral change and environmental restoration.
- To provide a complete and meaningful evaluation where both environmental and behavioral factors are important indicators of program success.

When Not to Use Stewardship Monitoring

- When no environmental change is expected as a result of the program.
- When there is no good indicator of the environmental improvement impact from the program, in which case the program should be redesigned to provide more tangible or measurable outcomes and impacts.

Review the Stewardship Program and Identify Environmental Indicators

An effective stewardship program will focus on informing participants or audience members about the specific behavioral changes that are needed to achieve the desired improvements in environmental resources. Barriers and motivators for stewardship participation can often be discovered by talking to participants before or during the program. These conversations might take the form of a focus group or interview process to provide input for program design (see related fact sheets). The "Fostering Sustainable Behavior" website features a collection of case studies that examine the motivations behind various environmental behaviors (http://www.cbsm.com; it is free to sign up to read the case studies). Another resource for better understanding the relationship between participants in outdoor activities and stewardship behaviors is the Stewardship Market Research Report available on the RBFF website (http://www.rbff.org).

Characteristics of a successful stewardship program:

- follows best practices or replicate other successful programs;
- overcomes barriers and uses known motivators for participation;
- provides solid information, explaining "why" things are important so that participants or communities can make informed behavioral choices;
- integrates stewardship into overall organization activities by using environmental and sustainable business practices.

Environmental indicators must be carefully selected for the purposes of stewardship program evaluation. The indicators must be readily influenced by the targeted behavioral changes and be able to show a meaningful change during the course of the evaluation. Appropriate indicators provide a direct link between

people and the natural resource. The indicators must be measurable, involve cost-effective data collection and processing, and be easy to interpret. Many stewardship programs use multiple indicators because programs are targeting multiple resource quality goals.

Measurement of stewardship program impacts often requires intensive labor and resource inputs. A fairly significant budget is required if you hope to prove a connection between the program, changes in behavior, and improvements in natural resource conditions.

The table below provides an example process for selecting indicators for several hypothetical stewardship programs. Team up with internal or external experts to help determine (and measure!) the appropriate environmental indicators.

Example Process for Selecting Indicators

Goal/ Ecosystem Attribute	Evaluation Question(s)	Sample Measure(s)	Sample Indicator(s)	Associated Stewardship Behavior(s)
Restore stream community/ Overall stream community health	Has the overall quality of the community improved?	Measures of community attributes, such as structures, compositions, and functions known to be important to key species in the community.	Index of Biotic Integrity (aquatic); number of self-sustaining populations of indicator species; percentage of species at risk; percentage cover by non-native (invasive) species	Participation in restoration behaviors; reduction of damaging behaviors (e.g., stream bank erosion, disposal of trash in streams)
Reduce pollution/ Recovery and health of polluted resources	What is the change in concentration of pollutants in streams? Is the source of the pollutant decreasing?	Soil and water pH, change in sedimentation loads, concentrations of pollutants (fertilizers, pesticides) in streams	Turbidity, suspended sediment, bacteria concentrations, toxins in water, toxins in fish, concentrations of pollutants, fish health, contamination in sediments	Reduction in point source pollution by industry and municipality; reduction in non-point source pollution (e.g., no dumping of oil in storm sewers, no washing cars on driveways, reduced use of yard chemicals)

Implement the Stewardship Monitoring Program

Using the selected environmental indicators, a team of evaluators, scientists, and program participants will monitor for environmental change during and after the stewardship program. Monitoring can include techniques as diverse as photographic evidence, water quality measurements, species sampling, erosion records, and other measurement techniques. Participation of program participants in the monitoring effort may reinforce motivation for appropriate stewardship behaviors.

In addition to monitoring the environmental indicators, the evaluation will also be looking at human behaviors resulting from the stewardship program. Refer to the fact sheets on Surveys, Interviews, Observations, and Skill Assessments for some ideas about evaluating the thoughts, behaviors, and motivations of the stewardship audience.

Case Study: Salmonpeople Campaign

An example of stewardship monitoring

The Salmonpeople Campaign works in a watershed for three years to design and implement a stewardship report card based on the unique vision of each community. Community participants follow a design process with five simple steps: (1) community commitment, (2) coalition building, (3) community asset mapping, (4) a seasonal rhythm of town meetings or "confluences," and (5) report card design. The report card is central to this effort. It reveals the results of measurable outcomes that community members themselves have selected. Indicators for the report card are drawn from a menu of indicators, such as health of salmon populations, pollution levels, deforestation, human health, economic prosperity, population, and energy use. Success is partly determined by improvement in the indicators and by how the stewardship principles are being adopted in policies, programs, and budgets in all sectors of the community, from town hall to the family unit.

Source: http://www.peterdonaldson.net/Salmonpeople/Campaign/overview.html

Evaluation Tools

Other Tools

Expert Opinion / Delphi Group

In some cases, you will need expert input for program design or evaluation purposes. ***Expert opinion*** can be especially valuable for programs involving complex technical problems, sensitive species, or scientific research outcomes. For example, expert input may be helpful in developing stewardship monitoring evaluations. Expert opinion can be gathered by a simple process of telephone conferences with one or more experts on an issue or problem. You might even consider having a pool of internal and external advisory experts whom you call on when you need them.

Delphi technique is a structured form of evaluation for gathering expert input. The technique is used especially for forecasting future trends, predicting future events, or designing new products. The Delphi technique is done at a distance – the group does not meet. Group members answer several preliminary questionnaires and the facilitator summarizes the responses. The synthesis is shared with group members for further feedback. The technique is repeated until the opinions of the group members converge or until enough information is available to complete the program design or evaluation. While it may generate fewer ideas than brainstorming techniques, this technique can generate a consensus over a difficult technical issue or problem. Like the namesake oracles, the best feature of the Delphi process is that a meeting of minds can occur without an actual meeting. It is a cost- and time-effective tool for gathering expert opinion. If the experts are brought together at the end of the process, the Delphi technique is often followed with a nominal group technique (see fact sheet).

Website Evaluation and Tracking

Most website hosting services provide a tracking feature for their clients. Organizations with their own domain or server will have internal experts who can provide similar tracking information. Basic website tracking information should be able to tell you:

- who visits your website;
- how they found the website (what search engine they came from);
- how they navigate your site;
- if they buy licenses or make other transactions on your site; and
- how effective the advertising or marketing is.

Websites or pages on websites that call for registration in order to view content can be an even better source of information. You or your organization can develop a form that collects very specific information from registrants. These registered users can then be a source population for future evaluation efforts. In addition, websites may be used as portals for survey research. A website can be programmed, for example, to ask every 10th user to take a survey, or can guide users to a link to a web-based survey. The possibilities with website technology are endless. Consult with an internal or external Internet expert to explore how the program's website might become a source of evaluation information.

Other Tools Include:

- Expert Opinion/Delphi Group
- Website Evaluation and Tracking
- Longitudinal Study/Panel Study
- Internal Review
- License Sales Tracking
- Cost-Benefit Analysis
- Open House/Public Meeting

Longitudinal Study / Panel Study

Longitudinal study involves the repeated measurement or observation of a program or topic over time. Longitudinal studies generally require careful matching of research methods to the problem under consideration. An informal longitudinal review can be compiled by taking a retrospective look at a series of past program evaluations to glean information about change over time, development of program or participant characteristics, "turning points" in that development, analysis of factors leading to change, and other factors that benefit from a long-term perspective.

Longitudinal studies often used "mixed methods" of evaluation, which means more than one evaluation tool is used. If statistically valid outcomes are desired, it would be best to consult an evaluation expert to properly design and execute the study. See the

discussion of Longitudinal Change in Chapter 4 for more details about different tools.

Similar to a longitudinal study, a ***panel study*** is an evaluation where the same group, or panel, of individuals is repeatedly interviewed over time. A longitudinal study might also take the more relaxed form of a series of "snapshots" – either real photographs or narrative/observational snapshots – of individuals or a group taken over time. This technique is often used during the course of a training program.

Internal Review

Internal review is a useful tool for program evaluation within an organization. It is often used by training organizations and manufacturers as part of a quality improvement process. An internal review would typically be performed by a committee, but the review's resulting statements may also be completed and compiled by individuals or teams within a program. Internal review information may be analyzed by an external examiner (often coupled with a site visit) for additional feedback and recommendations for program improvement.

Questions that might be answered by an internal review process include:

- What are the program's goals and objectives? What activities were undertaken to meet those goals and objectives?
- What has gone well with the program? What is the most successful aspect of the program?
- How has the program changed from what was originally proposed and what were those changes? Why did those changes happen and how did those changes affect the program outcomes?
- Where did the program run into difficulties, and how did you handle those challenges?
- What would you do differently in the program next time?
- What needs to be done right now to improve or bolster the program?

License Sales Tracking

Tracking of license sales is a common technique for showing the level of interest in fishing or other licensed activity in a state. Tracking of sales is a rote technical activity. Many states use automated fishing and hunting license sales systems. The systems often provide computerized, point-of-sale service for private license vendor locations throughout each state, as well as telephone and Internet license sales for individuals. Regular summaries of license sales are provided by the sales contractor or by the organization (usually a state agency).

License sales data is a program output, not an outcome or impact. License sales do not provide an idea of why people participate in outdoor recreation, or *motivations* for that participation. License sales also do not tell us the participant's *behaviors* or levels of *knowledge* about the resource. Nevertheless, license sales may be a better source of information than aquatic educators might imagine. License sales data can be mined to provide information about various subgroups within the larger population. For example, license sales data may be able to provide specific demographic information about license holders, information that can be valuable in designing or targeting future programs. The data may also be cross-referenced with market data from other sources, thus resulting in a rich data-mining opportunity that can enhance the understanding of the license "customers." There is a strong "customer relations management" opportunity in license sales data, one that could be used as the basis of designing customer surveys or other evaluation measures. Another benefit of license sales programs is that they provide a ready database of licensed outdoor recreation participants as a source population for program evaluation research.

Cost-Benefit Analysis

Cost-Benefit Analysis (CBA) is a relatively simple and widely used quantitative evaluation technique that provides information about whether to engage in program change or improvement. As the name suggests, the total costs and total benefits of a program are compared to each other. At its best, CBA can be an analytical way for society to make decisions about complicated issues such as education, health care, transportation, or the environment.

The difficulty with the process is that program costs are often incurred immediately, while program benefits may be intangible and received over time. CBA seeks to translate all relevant considerations into monetary terms. Cost-benefit analysts seek to monetize both the costs of regulation e.g., the money spent to install a water treatment plant) and the benefits of regulation (e.g., preventing pollution, preventing disease, and saving human lives). Program costs can be determined by placing a dollar value on the resources to plan, implement, and deliver a program, such as salaries, travel, materials, supplies, equipment, and facilities. Program benefits are the

Case Study: Targeting License Sales in Ohio

An example of an evaluation of license sales resulting from a marketing campaign

After 14 consecutive years of declining license sales, the Ohio Division of Wildlife joined a special pilot program that combined a marketing approach with RBFF's tested and targeted messages. RBFF's Water Works Wonders/Take Me Fishing ads were already running in Ohio as part of the national campaign. Based on their license sales and market research, the Division of Wildlife targeted 21 communities that would receive 750,000 copies of their 2001 fishing guide newspaper insert. The fishing guide's front cover showed the RBFF advertising. In addition, 60,000 license renewal reminder postcards were mailed directly to lapsed anglers in those same communities. Ohio also made sure that the images on the postcard reflected the ethnicity of the targeted communities. The target markets received the message multiple times, not just once or twice. Fishing license sales increased dramatically in the targeted experimental communities as compared to "control group" communities. In fact, when outreach programs were used, 40,000 more anglers bought a license in 2001 than in the previous year – a difference that totaled $560,000 in new revenues and a $3.50 return on investment for every dollar spent on the campaign.

"The RBFF campaign gave us the effective message we needed to use in our targeted marketing effort. The Water Works Wonders message sold fishing and boating as enhancing the quality of life, and it worked. Here in Ohio, for the first time in 14 years, we increased by six percent the number of fishing licenses sold in a year. That meant 40,000 more people were out on our waters enjoying fishing." – Mike Budzik, Chief, Ohio Department of Natural Resources, Division of Wildlife

positive outcomes and impacts that can be identified as resulting from the program. Some outcomes (i.e., organizational money saved) are easy to quantify. Other benefits (i.e., youth leadership development) may be ascribed a value by "shadow pricing" (i.e., potential for increased earnings with a college degree) or "opportunity costs" (i.e., resources wasted or polluted by people not receiving the program).

As you can imagine, there is much discussion in the economic community about how to put monetary values on life, health, and nature. A number of books have recently been written, for example, about the valuation of ecological services, such as clean water and wetlands. For a good introduction to the subject, see the

Environmental Literacy Council's description at http://www.enviroliteracy.org/article.php/1322.html.

Open House / Public Meeting

An ***open house*** is an informal setting with multiple displays where participants view information and discuss issues on set topics at stations around a room. Participants have the opportunity to talk directly and develop relationships with knowledgeable program or organization staff members. A ***public meeting*** is an organized meeting format involving a presentation and an opportunity for public questions and comments, as well as to present information and displays. Participants can hear all of the questions and comments in a large forum. Having a skilled facilitator is essential, especially if the public meeting topic is controversial. See the fact sheet on citizen advisory groups for more ideas.

Many formats are used to raise public awareness of programs or complex issues, and to engage in public discussion and seek input. Meetings are required by law for certain programs. Public meetings should be considered more often for educational purposes because of their instructive and revealing format. If the meeting or showcase is designed with the evaluation purpose in mind and for a particular target audience, it can often result in a large body of useful program design (needs assessment) or evaluation information, while simultaneously encouraging stakeholder involvement and support. If the program covers a broad geographic area, several meetings or workshops can be held throughout the region. These events may also be presented in a blended format or through virtual (electronic) means. Many organizations have staff members who are experienced in designing and offering public meetings or workshops.

Glossary

analysis of variance – A method for analyzing the differences in the means of two or more groups of cases.

assessment – The gathering and scoring of evidence (both quantitative and qualitative) that reflects learning or behavioral changes for program participants, with the purpose of influencing the learning environment or related programs or policies.

baseline data – Initial information on a program or program components collected prior to receipt of services or participation activities. Used later for comparing measures that determine changes in a program.

best practices – Successful innovations or techniques of top-performing organizations.

bias – A lack of objectivity, fairness, or impartiality on the part of the assessor or evaluator, often based on personal preferences and inclinations.

categorical/nominal – Quantitative measurement whose attributes have no inherent order. Numerals, labels, or names are assigned to the data such as gender, race, religious affiliation, political party, college major, or birthplace.

causality – A relationship between two variables in which a change in one brings about a change in the other.

chi-square – A non-parametric test of statistical significance. Typically, the hypothesis tested with chi square is whether or not two different samples (of people, texts, whatever) are different enough in some characteristic or aspect of their behavior that we can generalize from our samples that the populations from which our samples are drawn are also different in the behavior or characteristic.

cohort studies – A study in which subjects who share a common characteristic or experience within a defined time period and/or receive a particular program are followed over time and compared with another group that represents the general population from which the cohort was drawn or subjects who did not receive the program.

coded, coded data – (See coding.)

code – A symbol, either numeric or alphabetic, used to represent attributes or words (e.g., G3 = Third Grade, M = Math, TQ = Teacher asks question).

coding – The process of converting information obtained on a subject or unit into coded values (typically numeric) for the purpose of data storage, management, and quantitative analysis. See also, code.

correlation – A statistical means of showing a relationship between an intervention and an outcome. The degree of relationship between two variables, scores, or assessments.

cost-benefit analysis – An analysis that compares present values of all benefits less those of related costs when benefits can be valued in dollars the same way as costs. A cost-benefit analysis is performed in order to select the alternative that maximizes the benefits of a program.

culture of inquiry – A culture of learning that incorporates an organizational mindset and atmosphere of openness to asking questions about all aspects of work, careful consideration of the answers, a commitment to considering change, and a willingness to learn.

data analysis – The process of systematically applying statistical and logical techniques to describe, summarize, and compare data.

data mining – The science of extracting useful information from large data sets or databases.

demographic question – A question used in compiling vital background and social statistics.

demographics – Shared characteristics held in common by a group, such as age, sex, income, education, occupation, and geographic dispersion.

descriptive statistics – A statistic used to summarize or describe a set of samples upon which observations were made.

ethics – An internal system that determines correct behavior.

evaluation – The systematic collection and scrutiny of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming.

evaluation culture – A culture that accepts the use of evaluation, understands why the organization uses evaluation, can design or get advice on design of necessary evaluations, and uses evaluation, particularly to support change and development.

evaluation process – All of the steps involved in planning, executing, and analyzing an evaluation. See also, Logic Model.

formative evaluation – An evaluation conducted early in the planning or implementation of a program. It helps to define the scope of a program, identify appropriate goals and objectives, test program ideas and strategies, or provide for program improvement.

goal – The end toward which the project, program, organization, or department efforts are directed, supported by a series of objectives needed to realize it. A goal tends to be more general than an objective.

hypothesis – The assumed statement that is tested in a research process. In evaluation research, this typically involves a prediction that the program or treatment will cause a specified outcome.

hypothesis testing – The use of statistics to determine the probability that a given hypothesis is true.

impact evaluation – Measures the broad and long-term program effects, such as long-term changes (intended or unintended) in ecological, social, economic, or community conditions

impacts – The fundamental intended or unintended change occurring in organizations, communities, or systems as a result of program activities.

implementation evaluation – Assessment of program delivery (a subset of formative evaluation). See also, process evaluation.

indicators – Key pieces of information that let you know when your evaluation questions have been answered.

indirect benefit – Results that are related to a program, but not its intended objectives or goals.

inferential statistics – A statistic used to describe a population using information from observations on only a probability sample from the population. Used to model patterns in data or to draw inferences about the larger population from which the sample was taken, while accounting for randomness and uncertainty in the data.

inputs – The resources needed to create and implement a program including staff, time, money, materials, equipment, facilities, administrative approvals, budget authority, agreement with cooperating agencies, etc.

inquiry minded – (See culture of inquiry.)

Institutional Review Board (IRB) – An appropriately constituted group that has been formally designated to review and monitor research involving human subjects.

institutionalization of evaluation – The act of making program feedback a part of the organization's standard planning and management practices.

interval/discrete – Values that have all the features of integers (whole numbers), with equal values between the numbers. Examples include dates, temperatures (Celsius or Fahrenheit), IQ scores, and scores on many social survey questions.

Logic Model – A flow chart that traces how inputs and activities interact to produce outcomes and impacts.

longitudinal studies – The study of a particular individual or group of individuals followed over a given period to discover changes that may be attributable to the program.

Mann-Whitney U – A non-parametric statistical significance test for assessing whether the difference in medians between two samples of observations is statistically significant.

mean (average) – A statistic which is calculated by adding all the scores for one question or test together and dividing by the total number of tests or answers. This is less precisely referred to as the average.

mission – A statement of purpose by which an organization can measure its success.

model – Describes processes or strategies that are difficult to understand directly. A model may be a description, a representation, or an analogy.

median – The statistical mid-point of a group of answers or test scores.

mode – The most frequent answer or test score. This is determined through a simple count.

needs assessment – An analysis that studies the needs of a specific group and presents the results in a written statement detailing those needs. It also

identifies the actions required to fulfill these needs, for the purpose of program development and implementation.

objectives – Specific results or effects of project, program, organization, or department activities that must be achieved in pursuing the ultimate goal(s). Objectives tend to be more specific than goals.

ordinal/rank – Quantitative measurement whose attributes are ordered but for which the numerical differences between adjacent attributes are not necessarily interpreted as equal. Examples include the results of a race (without time intervals), and most measurements in the social sciences, such as attitudes, preferences, and social class.

organizational learning – (See evaluation culture.)

outcome evaluation – An evaluation that assesses the extent to which a program achieves its outcome-oriented objectives.

outcomes – Measurable results or consequences – both expected and unexpected – of an activity or program in meeting its stated goals and objectives, such as the percentage of participants who gain some knowledge or skill as a result of the program.

outputs – The quantity of products and services delivered by an agency or program to the intended users, such as number of programs, number of participants, geographic area covered, memberships acquired, money earned, etc.

pilot test – A pre-test or trial run of a program, evaluation instrument, or sampling procedure for the purpose of correcting any problems before it is implemented or used on a larger scale.

planning evaluation – Evaluation that occurs before and during the program to get baseline data, collect input, and develop guidance.

privacy – A person's privilege to have his or her information kept confidential and not disclosed to unauthorized parties without their permission.

process evaluation – Identifies the procedures undertaken and the decisions made in developing a program, describes how the program operates, the services it delivers, and the functions it carries out. See also, implementation evaluation.

qualitative evaluation – A process involving detailed, in-depth descriptions of program or participant characteristics, behaviors, cases, and settings.

quantitative evaluation – A systematic approach that attempts to define, measure, and report on the relationships between various program elements using numbers.

random sampling – A procedure for sampling from a population that gives each unit in the population a known probability of being selected into the sample.

ratio/continuous/scaled – A level of measurement which has all the attributes of nominal, ordinal, and interval measures, and is based on a "true zero" point. As a result, the difference between two values or cases may be expressed as a ratio. Examples include distance, length, temperature (Kelvin), age, length of residence in a given place, number of fish caught in a day, or number of events produced in a year.

raw data – Data collected in the evaluation process that have not been analyzed.

recommendations – A set of suggestions derived from the evaluation results.

regression analysis – A method for determining the association between a dependent variable and one or more independent variables.

request for proposal (RFP) – A document that outlines a specific format and requirements for an outside vendor or contractor to follow in submitting a proposal to provide the requested services.

sampling design – The sampling procedure used to produce any type of sample.

scope of work – A document that clearly states the work that is to be performed.

scoring (of evaluation surveys) – The process of determining the value of a performance on an indicator or criterion.

secondary program outcomes – (See unintended outcomes.)

stakeholders – People who have some sort of a stake or interest in the program that is being developed; people that care, are willing to develop a commitment, and/or are best able to offer input.

standard deviation – A statistical measure of the dispersion of a sample.

statistic – A number computed from data on one or more variables.

statistical analysis – Analyzing collected data for the purposes of summarizing information to make it more usable and/or making generalizations about a population based on a sample drawn from that population.

statistical significance – The degree to which a value is greater or smaller than would be expected by chance. Typically, a relationship is considered statistically significant when the probability of obtaining that result by chance is less than 5 percent if there were, in fact, no relationship in the population.

statistically significant – (See statistical significance.)

statistician – An expert or specialist in statistics and the statistical evaluation of data.

Student's t-test – An assessment of the statistical significance of the difference between two sample means, and for confidence intervals for the difference between two population means.

summative evaluation – An evaluation that measures the results or outcomes of a program when the program is completed, or on some regular basis for an ongoing program.

target audience – The individual, organization, company, or population type a communication, program or event is intended for.

time series analysis – Analysis of sequences of measurements that follow non-random orders, based on the assumption that successive values in the data file represent consecutive measurements taken at equally spaced time intervals.

timeline – The designated length of time in which activities will occur and the chronological sequence of these activities.

trend analysis – The analysis of the changes in a given item of information over a stated length of time.

triangulation – The attempt to obtain more valid results by using multiple sources of data about one aspect of performance, multiple methods of collecting data, and/or multiple interpretations of the same data.

vision – The ideal future the organization is striving to achieve.

Resources

This section contains resources that you can use in planning, designing, and implementing your evaluation, including books, journals, research, data sources, and websites.

This section is divided into the following general categories:

- Top Resources
- More Resources
- Evaluation Tools
- Stewardship
- License Sales Tracking
- Websites

Top Resources

Evaluating Extension Program Outcomes
University of Tennessee, Institute of Agriculture, Agricultural Extension Service
Uses Bennett's Hierarchy to develop planning with an emphasis on the "right people, right information, right way" evaluation planning approach.

USFWS Education Program Evaluation Participant Notebook
U.S. Fish and Wildlife Service National Conservation Training Center
304-876-7388
This binder has useful sections on planning (simplified logic model), level of attention (Bennett's Hierarchy) and objectives. It includes checklists and worksheets, descriptions of tools, and literature lists and resources.

Conducting Program and Project Evaluations: A Primer for Natural Resource Program Managers in British Columbia, FORREX Series 6
FORREX Forest Research Extension Partnership
http://www.forrex.org
Useful and concise information on level of attention, tools, and resources.

How to Conduct Evaluation of Extension Programs
Michigan State University Extension, Department of Education and Communication Systems, ANRECS Center for Evaluative Studies
Includes a chart of appropriate evaluation tools for various stages of program development, as well as information on tools and costs.

Water Quality Project Evaluation: A Handbook for Objective-based Evaluation of Water Quality Projects, Bulletin 868-98
Ohio State University Extension
http://ohioline.osu.edu/b868/
Information and examples of concise publications, plus discussions of objectives, barriers, tools, and interpreting and reporting results.

Monitoring and Evaluation: Some Tools, Methods, and Approaches
The World Bank
www.worldbank.org/oed/ecd
Compares various tools.

Designing Evaluation for Education Projects
NOAA Office of Education and Sustainable Development
202-482-2893
Contains useful sections on planning, tools, ethics, plus an appended chart of tools for specific audiences.

User-Friendly Handbook for Mixed Method Evaluations
National Science Foundation Division of Research, Evaluation, and Communications
Useful in all areas, especially key concepts, planning, tools, and data analysis, with emphasis on mixed methods and data triangulation.

Taking Stock: A Practical Guide to Evaluating Your Own Programs

Horizon Research, Inc.

Sections on planning, tools, and interpreting/reporting results.

Measuring the Success of Environmental Education Programs

Canadian Parks and Wilderness Society and Sierra Club of Canada BC Chapter

gthomson@cpawscalgary.org

Of particular note is the ripple diagram, planning checklist, tools for measuring values/behaviors, and tool samples.

Basic Guide to Program Evaluation

Carter McNamara

http://www.managementhelp.org/evaluatn/fnl_eval.htm

A practical and very concise guide.

CIPP Evaluation Model Checklist (Appendix in The CIPP Model For Evaluation)

Daniel L. Stufflebeam, Western Michigan University

http://www.wmich.edu/evalctr/archive_checklists/cippchecklist.htm

Assessing Nonformal Environmental Education: Unobtrusive Data Collection, Infosheet #25

EETAP Resource Library

http://eelink.net/eetap/info25.pdf

Description of observation techniques.

Understanding Evaluations of Environmental Education Programs, Infosheet #12

EETAP Resource Library

http://eelink.net/eetap/info12.pdf

Information for interpreting/reporting results.

W.K. Kellogg Foundation Logic Model Development Guide, #1209

W.K. Kellogg Foundation

http://ww2.wkkf.org/DesktopModules/WKF.00_DmaSupport/ViewDoc.aspx?fld=PDFFile&CID=281&ListID=28&ItemID=2813669&LanguageID=0800-819-9997

Proceedings of the Teton Summit for Program Evaluation in Nonformal Environmental Education

Ohio State University / Teton Science School

http://eric.ed.gov

Covers the "big picture" of evaluation, and includes short sections on politics and culture of evaluation and planning.

What Works: A Guide to Environmental Education and Communication Projects for Practitioners and Donors (Education for Sustainability series)

Martha C. Monroe (Academy for Educational Development and New Society Publishers)

www.aed.org and www.newsociety.com

The editors of this 1999 guide used a series of 41 case studies from around the world to illustrate how educators have successfully taught people about natural resources. From non-formal street theater and games to more formal water conservation campaigns, these case studies give readers a sense of the ingenuity and creativity in modern environmental education practice. Each case study includes a description of the situation, the project, and the results, including a highlight of the best practices of each project. The editors' introduction emphasizes the common features that lead to successful programs. ISBN 0-86571-405-3

Building Evaluation Capacity: 72 Activities for Teaching and Training

Hallie Preskill and Darlene Russ-Eft (SAGE Publications)

www.sagepublications.com

ISBN 0-7619-2810-3

EE Toolbox Workshop Resource Manual: Evaluating Environmental Education Materials

University of Michigan, School of Natural Resources

W.K. Kellogg Foundation Evaluation Handbook

W.K. Kellogg Foundation

www.wkkf.org

Program Evaluation: A Practitioner's Guide for Trainers and Educators

Brinkerhoff, R.O., Brethower, D.M., Hluchyj, T., Nowakowski, J.R. (Evaluation Center Western Michigan University)

Provides information on the fundamentals of program evaluation. Takes evaluators through the key steps of focusing, designing, managing, and evaluation. Provides examples of applied program evaluation at the school, state agency and university level.

Evaluation Thesaurus

Michael Scriven (Edgepress)

Provides practical and understandable definitions to key evaluation terms. The explanations themselves provide an excellent overview for those engaging in program evaluation at any level.

ISBN 0-8039-4364-4

More Resources

From Educational Theory to Conservation Practice (for the IAFWA Summit on Conservation Education)

The Evaluation Folk School and American University

Uses the Logic Model throughout. What is called "conservation education" or "CE" in this paper can be interpreted as "stewardship education" for aquatic educators.

Educating Young People About Water: A Guide to Program Planning and Evaluation

U.S. Department of Agriculture, Cooperative State Research and Extension Service

614-292-6717

Planning and evaluation checklists and resource lists.

Developing and Evaluating EE Programs, #94-66

Washington State Department of Ecology

Includes worksheets, and good sections on planning and assessing evaluation needs.

Measuring Progress: An Evaluation Guide for Ecosystem and Community-Based Projects

University of Michigan, School of Natural Resources & Environment, Ecosystem Management Initiative

http://www.snre.umich.edu/emi/evaluation

Contains a section on planning and an evaluation cycle diagram, and good introductory information.

Footprints: Strategies for Non-Traditional Program Evaluation

National Science Foundation

http://www.nsf.gov/pubs/1995/nsf9541/nsf9541.pdf

Sections on planning and flow, and flow charts p25+.

What Works: Documenting Standard Practices for Aquatic Resource Education

U.S. Fish and Wildlife Service, Region 5

413-253-8506

Information on standard aquatic education methods and stewardship concepts, including case studies.

Evaluation: Parks Project Sample, Section III, How Do We Assess the Informal Education Component? (In: Community Connections for Science Education, Volume II, History and Theory You Can Use)

Ohio State University and National Science Teachers Association Press

www.nsta.org

Best Practices Workbook for Boating, Fishing, and Aquatic Resources Stewardship Education

Recreational Boating & Fishing Foundation

www.rbff.org

Evaluation Tools

Case Study Research Description and Examples

NOAA Coastal Services Center

http://www.csc.noaa.gov/mpass/tools_casestudies.html

Software for Content Analysis (a review)

Audience Dialogue (non-profit)

http://kb.ucla.edu/system/datas/5/original/content_analysis.pdf

Qualitative Content Analysis (article in the online peer-reviewed journal The Forum: Qualitative Social Research, Volume 1, Number 2, June 2000)

Philipp Mayring

http://www.qualitative-research.net/fqs-texte/2-00/2-00mayring-e.htm

Content Analysis Guidebook Online

Cleveland State University

http://academic.csuohio.edu/kneuendorf/content

Guidelines for Survey Research Quality

The Council of American Survey Research Organizations

http://www.casro.org/codeofstandards.cfm

Collecting Evaluation Data: Direct Observation

University of Wisconsin, Cooperative Extension, Program Development and Evaluation Publications

http://www.uwex.edu/ces/pdande/evaluation/evaldocs.html

Stewardship

Stewardship Market Research Report

The Recreational Boating & Fishing Foundation

http://www.rbff.org

Stewardship Programs and Evaluation of Ecosystem Health

The University of Michigan, Ecosystem Management Initiative

http://www.snre.umich.edu/ecomgt

Measuring Results

CoEvolution Institute

http://www.coevolution.org/measuring.html

Identifies methods to measure the impact on attitudes and behaviors from informal learning contexts such as zoos, museums, parks, and other natural settings.

Lake Stewardship Education

Maine Lakes Conservancy Institute

http://www.mainecola.org/

The Place-based Education Evaluation Collaborative

http://www.peecworks.org

Evaluation of environmental literacy programs in New England.

The Theory and Practice of Aquatic Stewardship Education (summary of symposium)

The American Fisheries Society 2005 Annual Meeting

http://www.wdafs.org/Anchorage2005/program/27_aquatic_stewardship.htm

License Sales Tracking

Cooperative License Sales Marketing Program

The Recreational Boating & Fishing Foundation

http://www.rbff.org

National License Trends Data

Automated Wildlife Data Systems provides

http://www.nssf.org/PDF/HuntingLicTrends-NatlRpt.pdf

A system for tracking national license sales, and a library of information.

National License Trends Data

The American Sportfishing Association

http://www.asafishing.org/statistics/participation/

Industry data and statistics.

Websites

Guidelines for Excellence (National Project for Excellence in Environmental Education)

North American Association for Environmental Education

http://www.naaee.org/programs-and-initiatives/guidelines-for-excellence

Planning and Evaluation Resource Center (PERC)

Innovation Center for Community and Youth Development, Institute for Applied Research in Youth Development at Tufts University, Social Policy Research Associates

http://www.evaluationtools.org

Good introduction to the evaluation cycle, and links to many different evaluation tools.

Program Development & Evaluation
University of Wisconsin Extension
http://www.uwex.edu/ces/pdande
Good overview of the logic model, with planning worksheets and useful articles on evaluation practices.

Program Evaluation
University of Kentucky Extension Service
http://www.ca.uky.edu/agpsd/evaluate.htm
Southern Region Evaluation Resource
http://www.ca.uky.edu/agpsd/soregion.htm Links to fact sheets and articles on evaluation topics from extension services around the country.

Program Evaluation
Penn State University
http://extension.psu.edu/evaluation/
Well-written tip sheets on every evaluation topic imaginable, from sample size to reaching new audiences.

Free Management Library for Non-Profits
http://www.managementhelp.org
Useful information on evaluation and other organizational management topics. See also, the subsection on evaluation:
http://www.managementhelp.org/evaluatn/evaluatn.htm

The Center for What Works
http://www.whatworks.org
For information on measuring program performance and benchmarks, click on the "Tools & Resources" link.

Web-based Survey Services
Zoomerang
http://www.zoomerang.com
Survey Monkey
http://www.surveymonkey.com
SurveyKey
http://www.surveykey.com
These web-based survey services allow for the design and distribution of surveys through e-mail or web links. Some of them allow free surveys of up to 50 people.

Organizational Self-Assessment Checklist
The National Endowment for the Arts
http://arts.endow.gov/resources/Lessons/WARSHAWSKI.HTML
Free, helpful for internal evaluations. Can also be downloaded as an Excel file.

Mixed Method Evaluations (free user-friendly handbook)
National Science Foundation
http://www.nsf.gov/pubs/1997/nsf97153/start.htm
PDF version:
http://www.nsf.gov/pubs/2002/nsf02057/nsf02057_1.pdf

Self-Assessment Tool
Drucker Foundation Leader to Leader Institute
http://www.leadertoleader.org/tools/sat/index.html

Center for Social Research Methods
Cornell University
http://www.socialresearchmethods.net
Includes information about the planning-evaluation cycle and social research statistical methods.

American Evaluation Association
16 Sconticut Neck Road, #290,
Fairhaven, MA 02719.
888-232-2275
www.eval.org
Website includes a "find an evaluator" function. They also have an annual conference and summer training institute.

Rutgers Cooperative Research and Extension
http://www.rce.rutgers.edu/evaluation
Program evaluation resources, including survey techniques, sample surveys and evaluation methods.

University of Wisconsin Program Development and Evaluation

http://www.uwex.edu/ces/pdande

Resources written in easy to understand language to help you develop your own programs.

National Extension Water Outreach Education

University of Wisconsin

http://wateroutreach.uwex.edu/use/Evaluate.cfm

Pennsylvania State University

http://www.extension.psu.edu/Evaluation/titles.html

Tip sheets to help with the development of evaluation programs.

New York State Program Evaluation

http://www.programevaluation.org

Geared to the needs of school teachers, there are some useful resources here, including case studies, references and downloadable tools.