

Measuring Results

**Barbara Schneider, Ph.D., and
Nicole Cheslock, M.A.**

*Gaining insight on behavior change
strategies and evaluation methods from
environmental education, museum, health,
and social marketing programs*



A report of the **Coevolution Institute**
Understanding Metrics Project

San Francisco, California
April 2003

About the Authors

Barbara Schneider earned a doctoral degree in science education from Stanford University. Her thesis, “Making the Informal Formal: An Examination of Why and How Teachers and Their Students Leverage Experiences to Informal Learning Environments,” was presented in Winter 2002. She is a part-time lecturer at Stanford and an evaluation consultant for several institutions in the western states. She holds a master’s degree in marine science from the Moss Landing Marine Laboratories.

Nicole Cheslock has been involved in evaluation since 1996, beginning with an internal assessment of discretionary account management for Morgan Stanley in New York City. She is an evaluation consultant for organizations focused on educating youth and adults. Her projects include program evaluation and staff training for Coro Northern California, Hidden Villa, Palo Alto Junior Museum and Zoo, and Yosemite National Institutes. She holds a master’s of education from Stanford University (2001).

Acknowledgments

The Coevolution Institute extends its gratitude to Mr. John Abele and the Barr Foundation for their support in creating the Understanding Metrics Project.

We thank the following individuals for reviewing and providing critical feedback on Measuring Results: Jack Chin; Zahava Doering, Ph.D.; Diane B. Frankel; Alan J. Friedman, Ph.D.; George Hein, Ph.D.; Richard Margoulis, Ph.D.; Sung Ji Nam; Marian Kane; Nick Salafsky, Ph.D.; Carol Saunders, Ph.D.; Larry Suter, Ph.D.; and Steven Yalowitz, Ph.D.

The Coevolution Institute

The Coevolution Institute (CoE), founded in 1997, is a 501(c)3 nonprofit organization that seeks to improve the health and survival of species through conservation efforts and educational interventions. Our current initiatives include the Understanding Metrics Project, the North American Pollinator Protection Campaign, the Bug Mobile, and the Butterfly and Pollinator Gardening Program.

© 2003, Coevolution Institute

423 Washington Street, 4th Floor
San Francisco, CA 94111
415-362-2339 tel
415-362-1137 fax
info@coevolution.org

About the Sources

Research for Measuring Results took place through an extensive academic literature review in the fields of environmental education and museum research and, to a lesser extent, in health programs and social marketing research. A sample of studies representative of common and accepted approaches was selected to describe health and social marketing practices.

Project Team

Project Concept: Paul J. Growald, CoE Chairman
Project Supervision: Laurie Davies Adams, CoE Executive Director
Production Associate: Caitlin Cartwright
Copy Editing: Kate Kelly
Graphic Design: David Gagne

Contents

Contents

Foreword	4
Introduction	6
Chapter 1: Evaluation: What, Why, and How?	9
Chapter 2: Environmental Education and Environmental Psychology	24
Chapter 3: Museums	53
Chapter 4: Health Programs	83
Chapter 5: Social Marketing	101
Chapter 6: Lessons Learned	121

Foreword

*The Coevolution Institute
is a catalyst for
biodiversity stewardship.*

The Coevolution Institute (CoE) works to create meaningful learning experiences that result in positive conservation attitudes and actions. CoE recognizes the complexity and the importance of evaluation in the pursuit of promoting responsible environmental behavior. *Measuring Results*, the first undertaking in CoE's Understanding Metrics Project, examines the ways that those within and outside of environmental education have approached the challenges of effecting and measuring individual behavior change.

As many recognize, measuring the impact of informal education programs is a complex task. Despite the challenges, we believe that evaluation is fundamental. Our commitment to better understanding the outcomes of our work stems from key questions that surfaced from our original idea to create a new constituency for biodiversity stewardship: How would we measure the effectiveness of our programmatic experience? As an organization, how would we measure outcomes? The institute set out to explore evaluation, including research on the theory that guides programs and evaluation, and to study the findings of research and evaluation efforts. Specifically, we asked: What can we learn about evaluation and outcome measurement through a cross-field survey?

Measuring Results is a synthesis of a large and varied body of knowledge concerning behavior change theories, evaluation methodologies, and findings. Although CoE is most interested in behavior modification as applied to conservation and to the protection of biodiversity, this report brings together research on a range of outcomes—including participants' knowledge, attitudes, and behavior change—as tracked by museum, health, and social marketing programs.

Measuring Results is intended for readers who seek background information on what evaluation means and has meant in different fields that focus on behavior change. It should be useful for those who want and need to inform themselves about methods and research findings so they will approach evaluation enlightened by the work of others. *Measuring Results* provides program providers, grantors, and other stakeholders with common terminology and background on how those in different fields have approached

behavior change interventions and assessment. There are many useful “how-to” resources that provide step-by-step evaluation instructions, but that is not our purpose here (see the resources sections of this report for instructional evaluation guides).

Reader feedback is an essential component in the assessment of this project. We welcome comments and suggestions through the reader survey response materials at the end of this report or by e-mail to info@coevolution.org. We wish you—*measurable*—success!

*Laurie Davies Adams
Paul J. Growald
Coevolution Institute
April 2003*

Introduction

Field-Specific Chapter Organization

Theoretical Frameworks

Theoretical underpinnings that can guide programs and evaluation (e.g., evaluation might be guided by theoretical frameworks that

explore effective behavior change just as programs can be a practical manifestation of theory)

Methodology

Data collection strategies

Research Findings

Outcomes of evaluation and research

Chapter Summary

Synopsis of key points

Summary of studies (table)

References

Resources

The Coevolution Institute (CoE) seeks to protect biodiversity through conservation efforts and educational programs. Like most environmental educators, two of CoE's goals are to heighten awareness of environmental issues and to inspire responsible conservation behavior. CoE has a vested interest—some argue, a responsibility—to know whether its work, and that of similar organizations, changes individual behavior. This report is meant as a springboard to a better understanding of behavior change interventions and for considering evaluation practices that can be used now and developed further.

For the purposes of this work, *evaluation* is a general term pertaining to the assessment processes and data collection techniques used to identify program outcomes. We focused on published academic research across fields (and, to a much lesser extent, on individual evaluation reports that dominate evaluation work and that vary substantially from one organization to the next and across fields). References to widely accessible evaluation guidebooks—including how-to manuals and software—are provided at the end of each chapter, as are citations to other resources that can be used in concert with the references cited in the text.

Chapter 1, “Evaluation: What, Why, and How?” introduces readers to evaluation by defining terminology and explaining different models. The next four chapters explain evaluation and core results from specific fields:

- Environmental education and environmental psychology
- Museums
- Health programs
- Social marketing

Each field-specific chapter includes sections on theoretical frameworks, methodology (common data collection and analysis strategies), and findings. Techniques are listed and criticisms are outlined to elucidate current methods and point out areas for development. Research findings in each chapter describe existing research, effective practices, potential uncertainties, and accepted beliefs. The core chapters include summaries that

Databases and Online Journals

Association of Science-Technology Centers
<http://www.astc.org>

Curator
<http://www.altamirapress.com/RLA/Curator/Index.shtml>

Educational Research Information Center
<http://www.eric.ed.gov>

Environment and Behavior
<http://www.sagepub.com>

Museum Learning Collaborative
<http://museumlearning.com/default.html>

Institute of Museum and Library Services
<http://www.imls.gov>

Journal of Environmental Education
<http://www.heldref.org/html/jee.html>

Journal of Research on Science Teaching
<http://www3.interscience.wiley.com/cgi-bin/jtoc?ID=31817>

PsychLit
http://infoeagle.bc.edu/bc_org/apv/ulib/ref/guides/edu/psycq.html

Science Education
<http://www3.interscience.wiley.com/cgi-bin/jtoc?ID=32122>

Smithsonian Institutional Studies Office
<http://www.si.edu/iso/biblio.htm#Exhibition520Assessment%20Studies>

Visitor Studies Association
<http://www.visitorstudies.org>

highlight field-specific theories, evaluation methods, and key findings.

The value of this work lies in the practices that can be used now and developed further based on existing work across fields. “Lessons Learned,” the concluding chapter, synthesizes assessment strategies and the findings from the core chapters. “Lessons Learned” includes a cross-case analysis, describing the implications from the existing literature.

Depending on their expertise and positions, readers might prefer to delve into a specific chapter, to rely on the chapter summaries, or to focus on “Lessons Learned.” Readers seeking a general overview of assessment techniques and findings are encouraged to read Chapters 1 and 6, review the field-specific chapter summaries (in Chapters 2 through 5), and refer to the online evaluation guidebooks noted at the end of Chapter 1.

Research Methods and Information Sources

Our research took place through an extensive academic literature review of environmental education and museum research and, to a lesser degree, in health programs and social marketing. The health field was chosen because of its emphasis on promoting behavior change. A sample of studies that is representative of common and accepted approaches, but by no means comprehensive, describes various interventions and effective practices in health and social marketing.

Chapters 2 and 3 include references from searches in a variety of databases and several prominent journals. Bibliographies by Screven (1999) and Smith (1999), and the meta-analyses by Hines and colleagues (1986/87), Leeming and colleagues (1993), and McManus (1992) and internal research conducted by the Smithsonian’s Institutional Studies Office also were used. The text box lists some online

sources that also were referenced. Our own experience in field research and evaluation rounded out the sources examined for this report.

Barbara Schneider

Nicole Cheslock

April 2003

References

- Hines, J., H. Hungerford, and A. Tomera. 1986/87. Analysis and synthesis of research on responsible environmental behavior: Meta-analysis. *Journal of Environmental Education* 18(2):1–8.
- Leeming, F. C., W. O. Dwyer, B. E. Porter, and M. Cobern. 1993. Outcome research in environmental education: A critical review. In *Essential readings in environmental education*, 209–226. Champaign, Ill.: Stipes.
- McManus, P. 1992. Topics in museums and science education. *Studies in Science Education* 20:157–182.
- Screven, C. G. 1984. Educational evaluation and research in museums and public exhibits: A bibliography. *Curator* 27(2):147–165.
- Smith, S. J. 1999. Bibliography of sources. *Curator* 42(2):174–179.

Chapter 1

Evaluation: What, Why, and How

In This Chapter

Introduction to Evaluation

The Reasons

The Methods

An Example

Chapter Summary

References

Resources

This chapter provides a basic overview of the purposes, products, and common approaches of evaluation. Although many useful “how to” resources that provide instructions for conducting an evaluation are cited throughout the chapter (and provided at the end of Chapters 2–5), a step-by-step guide is not our purpose here.

Introduction to Evaluation

Evaluation is used to measure organizational or program components, needs, structure, or influence. In the past decade, there has been a growing push, as evidenced by legislation and by changes in funding guidelines, to make programs accountable and to improve them through research and evaluation. For this report, the term *program* includes the interventions (educational initiatives and public-health campaigns, and museum visits, for example) that are the object of the evaluation.

The Government Performance and Results Act of 1993 (GPRA)¹ requires that federally funded agencies create performance goals and report progress annually. GPRA has affected federal agencies, and it has changed the way state and local agencies work as well. An increased demand for performance metrics and documented outcomes also has developed in the private sector—especially among nonprofit organizations.

United Way of America, one of the nation’s largest groups of community-dedicated nonprofit organizations, researched and revised its reporting guidelines in 1995. United Way transformed its approach to funding by differentiating between *outputs* (the services rendered, the number of people served, or the number of visitors to a site) and *outcomes* (the results of program participation, including skill acquisition or changes in participants’

¹The White House web site lists the first three purposes of GPRA as follows: “1. improve the confidence of the American people in the capability of the Federal Government, by systematically holding Federal agencies accountable for achieving program results; 2. initiate program performance reform with a series of pilot projects in setting program goals, measuring program performance against those goals, and reporting publicly on their progress; 3. improve Federal program effectiveness and public accountability by promoting a new focus on results, service quality, and customer satisfaction.” (Accessed July 24, 2002: <http://www.whitehouse.gov/omb/mgmt-gpra/gplaw2m.html#h1>.)

Evaluation work was treated initially as a domain of research-oriented social science, but it quickly became apparent that although this was part of the story it was not the whole of it.

(McLaughlin and Phillips, 1991, p. 18)

knowledge, attitudes, or behavior). The United Way outcome evaluation model (<http://national.unitedway.org/outcomes/resources/mpo/model.cfm>) defines inputs, activities, outputs, and outcomes. Inputs are resources dedicated to the program, including staff, funds, material, and facilities. Activities are the program's actions or what is done with the resources, and outputs are the direct products of program activities (the number of people served, number of classes, or number of repeat visits, for example). Outcomes differ from activities and outputs: Outcomes are changes in participants' attitudes, behavior, and knowledge as a result of their experience.

According to "Measuring Program Outcomes: A Practical Approach" (United Way, 1996), outcomes are

Benefits for participants during or after their involvement with a program. Outcomes may relate to knowledge, skills, attitudes, values, behavior, condition, or status. Examples of outcomes include greater knowledge of nutritional needs, improved reading skills, more effective responses to conflict, getting a job, and having greater financial stability.

In "Perspectives on Outcome Based Evaluation for Libraries and Museums," Beverly Sheppard (1999) delineates the outputs–outcomes distinction as "a system of evaluation that replaces 'What have we done to accomplish our goals?' with the question, 'What has changed as the result of [our] work?'" The outcome evaluation model is illustrated in Figure 1.1 as it can be applied to the Coevolution Institute's (CoE) Bug Mobile program.

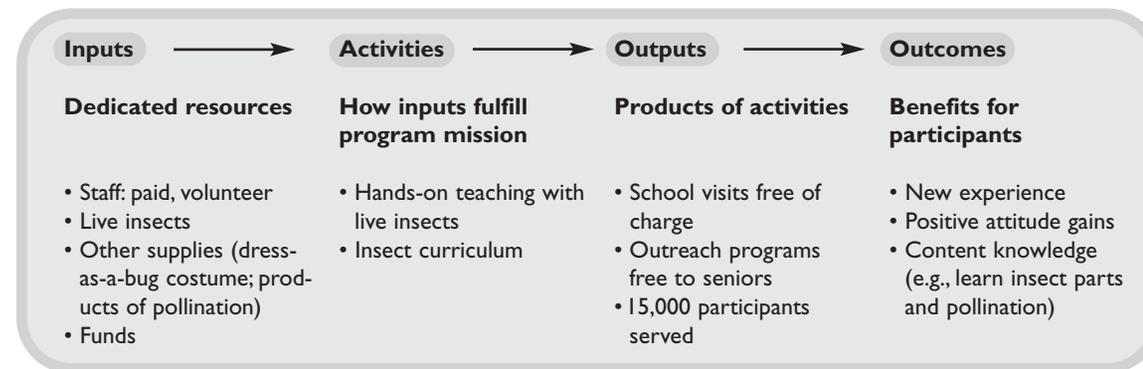


Figure 1.1. The United Way evaluation model as it could be applied to the CoE Bug Mobile.

The Bug Mobile, a hands-on science program piloted in the San Francisco Bay Area beginning in 2000, encourages positive attitudes toward insects and educates participants about the importance of conservation.

The United Way's work and funding guidelines have influenced assessment among non-profit organizations across sectors. Inspired by the United Way's revised reporting guidelines, many foundations now require outcomes-based performance metrics from grant recipients. Beyond the growing requirement for accountability, granting institutions know that evaluation also serves an important role in program comparison, development, and improvement.

Why Evaluate?

Simply put, there are two approaches to and corresponding goals for evaluation

(Katzenmeyer, 1997):

Formative: *to gain direction for improving projects as they are developing*

Summative: *to determine projects' effectiveness after they have had time to produce results*

The Reasons

Assessment serves many functions, from meeting funding requirements to facilitating program development to communicating objectives and results. Evaluations are designed for various audiences: grantmaking organizations, program staff, program participants, academic researchers, and policymakers. In his "General Statement on Evaluation," Ralph Tyler (1991), who has more than 50 years in the field of evaluation (primarily focused on formal education), summarizes the reasons for evaluation:

- Monitor a current program
- Select a better available program
- Assist in developing a new program
- Identify the different effects of a program on different populations
- Estimate the effects and cost-effectiveness of a program
- Test the relevance and validity of a program's defining principles

Although the formative and summative functions of evaluation are different, they serve equally important ends. Formative evaluations most often provide information for program development. The data evaluators glean from formative work (from experiments, focus groups, interviews, and observations) aim to guide the final design of an exhibit or program. For example, the Smithsonian's Institutional Studies Office conducted a formative evaluation before developing an exhibit about oceans. The staff collected demographic information and assessed visitors' knowledge and attitudes about the ocean and conservation through interviews and questionnaires. The formative evaluation was conducted to

predict potential differences in exhibit outcomes at diverse locations (the information gathered suggested that geography does affect awareness of ocean issues; but it is not related to *how* people acquire information about ocean issues). The findings provided useful data for building the exhibit.

Summative evaluation focuses on outputs and outcomes: Did the programs accomplish what they set out to do? Did people learn or change as a result? Experiments, interviews, and observation also are used in summative evaluations. The Smithsonian's Ocean Planet exhibit sought to promote responsible behavior by providing information about ocean heroes, people, and products. The summative evaluation, conducted to determine the exhibit's effectiveness, included personal interviews with entering and exiting visitors. In addition to soliciting demographic information, the interviewers asked visitors about the role of oceans, what their predictions were for the future of oceans, and what they saw as solutions. The exit interview asked visitors to cite specific examples about the exhibit and their overall impressions of the exhibit. They were asked questions like these: "What did you find most interesting? What surprised you? What was most informative?" The summative evaluation provided data on what people learned and experienced at the museum.

The Methods

Evaluation is a complex task that can require extensive human and financial resources, and it can take multiple forms.

The American Evaluation Association has designated "Guiding Principles for Evaluators," five general standards for evaluation across sectors (accessed Oct. 22, 2002: <http://www.eval.org/EvaluationDocuments/aeaprin6.html>):

- Systematic inquiry: Evaluators conduct systematic and data-based inquiries about the topic being assessed.
- Competence: Evaluators provide competent performance to stakeholders.
- Integrity and honesty: Evaluators ensure the integrity of the process.
- Respect for people: Evaluators respect the security, dignity, and self-worth of program stakeholders.
- Responsibility for public welfare: Evaluators consider the general public's diversity of interests and values.

AEA's Guiding Principles

- Systematic inquiry
- Competence
- Integrity and honesty
- Respect for people
- Responsibility for public welfare

Evaluators use various qualitative and quantitative research techniques as well as standard frameworks or models (logic models, for example). According to the National Science Foundation's *User-Friendly Handbook for Mixed Method Evaluations* (Katzenmeyer, 1997, pp. 2, 4),

Experienced evaluators have found that most often the best results are achieved through the use of mixed method evaluations, which combine quantitative and qualitative techniques ... [that] provide a tradeoff between breadth and depth and between generalizability and targeting to specific populations.

Common evaluation methods and data collection tools are listed and explained in Table 1.1. For readers who might be less familiar with those approaches, more thorough explanations of log frames and empowerment evaluation follow the table. Depending on the questions addressed, any and all of those approaches could be appropriate for evaluating a specific program or institution, and each has inherent strengths and deficiencies.

Table 1.1. Evaluation Methods

Method	Explanation
Case study	<ul style="list-style-type: none"> • Captures information from program stakeholders, including the different perspectives and experiences of donors, participants, managers, educators • Includes document analysis, interviews, observation • Investigates phenomena within a specific context
Empowerment evaluation	<ul style="list-style-type: none"> • Encourages input from staff and affiliated interested parties (program donors, participants, clients) • Includes three-step process: <ul style="list-style-type: none"> • Define a mission statement • Take stock, identify, rate, rank the most significant program activities • Plan for the future through goal and strategy statements
Ethnography	<ul style="list-style-type: none"> • Like the case study, provides descriptive information primarily through analysis of multiple sources of data (documents, people, surroundings), interviews, and observation • Emphasizes human behavior, social contexts • Uses field-generated metrics based on data collection
Experiment	<ul style="list-style-type: none"> • Compares treatment and control groups to determine whether differences can be attributed to the program • Assesses outcomes by comparing results or performance in treatment and control groups
Listening (focus group, interview, questionnaire, survey)	<ul style="list-style-type: none"> • Collects data on participant or visitor attitudes, knowledge, expectations, experience through spoken or written forms • Analyzes transcripts to provide salient information, including potential trends in responses
Log frame	<ul style="list-style-type: none"> • A 4 x 4 project-planning matrix for defining program goals, purpose, outputs, and activities and the inputs needed to achieve them • Includes active program staff participation • ZOPP (<i>Zielorientierte Projektplanung</i>) or in English, GOPP (goal-oriented project planning) is an example of a log frame: <ul style="list-style-type: none"> • Most commonly used for international projects • Includes 2 phases: analysis (identification of problems) and project planning • Requires a trained facilitator (1- to 14-day workshop) • Widely used in international projects
Logic model	<ul style="list-style-type: none"> • Visual depiction showing relationships between program components including resources, activities, strategies, target audiences, outcomes, indicators • Multiple uses: facilitate program organization, enhance team building, communicate links among program components, focus the evaluation process by identifying key issues and questions • The United Way outcome model is an outcome-based logic model. Other forms include process theory models, which describe the internal organization of the program (interaction between staff and participants) and program development and monitoring models, such as the log frame

Table 1.1. Evaluation Methods *Continued*

Meta-analysis	Reviews, analyzes, and compares results from multiple studies pertaining to the same subject in assessing the effect of a particular approach or intervention
Observation	Collects data on participant or visitor behavior, interaction, social engagement (conversation, gestures)
Pre / post tools	Collect participants' responses before and after a program
Qualitative methodology	<ul style="list-style-type: none"> • Takes multiple forms: case studies, ethnography, focus groups, interviews, observations, open-ended questions in surveys and questionnaires • Emphasizes context • Addresses atypical results (outlying data)
Quantitative methodology	<ul style="list-style-type: none"> • Takes multiple forms: experiments, pre/post tests, questionnaires, statistical tools, surveys • Seeks to control context through random assignment or multivariate analysis • Generally disregards deviant cases (outlying data)
Results framework	<ul style="list-style-type: none"> • Planning, management, evaluation, and monitoring tool widely used by the U.S. Agency for International Development (USAID) • Staff required to state program hypothesis and define ultimate objective as well as short-term and intermediate objectives • Emphasis on causal linkages and concrete results
Social return on investment (SROI)	<ul style="list-style-type: none"> • Designed by Roberts Enterprise Development Fund • Quantifies the social-economic value of programs or organizations based on a discounted-cash-flow analysis • Includes qualitative descriptions of the intervention, organization, and people served

Logic Models and Log Frames

Logic models are widely used to assess program success across fields, and they serve an evaluative function by specifying intended program outcomes and by describing the strategies for achieving them. Logic models serve these principal aims:

- They build cohesion between people involved in the program (donors, managers, participants, and evaluators).
- They serve as strategic planning tools.
- They facilitate communication of program goals and components.

Logic models depict the relationship between program components (as shown by the Bug Mobile example in Figure 1.1). But by starting with the intended outcomes, logic models

include the steps necessary to achieve the program's stated goals. In essence, by creating a logic model, one is explaining the reason for the program and strategies for success.

Log frames are another example of a logic model (Figure 1.2). Log frames are often used by program managers in Europe, Canada, and Australia. In fact, *Zielorientierte Projektplanung*, or ZOPP, is used for almost all state-funded projects in Germany and is required for funding approval. ZOPP is also required for proposals to the British Overseas Development Agency.

	Narrative Summary What does the project want to achieve?	Objectively Verifiable Indicators How can we tell if we have achieved the goal?	Important Assumptions What else must happen if it is to succeed?
Goal			
Purpose, objective			
Outputs			
Activities			

Figure 1.2. Log frame matrix (adapted from Odame, 2000, p. 3; Farrington and Nelson, in Odame, 2000, p.4).

There are many ways to develop logic models. For the most part, development is meant to be a collaborative effort among stakeholders, and it focuses on defining program goals, strategies, activities, performance measurements, and indicators of success. Kirkpatrick (2001) outlines eight steps for the development of a program logic model:

1. Form a small work group (program staff, evaluators, other stakeholders).
2. Review program reports, planning documents, and relevant literature.
3. List project goals.
4. Define the target group.
5. Outline the program's process or implementation objectives: What will the individuals who implement the program be doing? Outline the process indicators: How will you know if the program activities have been implemented as planned?

6. Determine immediate, intermediate, and long-term indicators or outputs.
7. List the project activities and group program activities into components or strategies. Check your logic: Is each element of the model causally linked to the next? Are causal linkages realistic? Are objectives clear and measurable? Are activities clear and measurable? Are there other potential activities for achieving the outcomes? Are resources adequate?
8. Verify the logic model with stakeholders who were not work group participants and modify it as needed

The resources section at the end of this chapter includes a list of planning guides.

Empowerment Evaluation

Empowerment evaluation, another example of a participatory formal evaluation structure, emphasizes inclusion of and collaboration among multiple program stakeholders (stakeholders often include managers, donors, educators, organizers, board members, and participants) to define mission, take stock (outline strategies and evidence of success), and plan for the future. Empowerment evaluation, as defined by David Fetterman (<http://www.stanford.edu/~davidf/empowermentevaluation.html#summary>, accessed Jan. 6, 2003) is

... part of the intellectual landscape of evaluation. It has been adopted in higher education, government, inner-city public education, nonprofit corporations, and foundations throughout the United States and abroad. ... Empowerment evaluation is the use of evaluation concepts, techniques, and findings to foster improvement and self-determination.

Whatever the method or combination of methods, evaluation is a rigorous process of examining one's work that can take place before, during, and after program delivery.

An Example

Evaluation challenges often involve defining a measurable goal, such as ascribing a value to learning to read or identifying knowledge gained from a museum visit, and actually measuring outcomes, such as tracking environmentally responsible actions after

Nature Conservancy Findings

- Set measurable goals
- Keep measures simple
- Keep measures easy to communicate
- Manage with measures

participation in an outdoor ecology program. The Nature Conservancy's recent grappling with evaluation provides a relevant example of one organization's approach and what it has learned in the process about self-assessment and its mission of stewardship.

The Nature Conservancy's experience and its research involving 20 other nonprofit organizations illustrate how one organization can create and measure targeted outcomes to better assess its work. The metric development issues faced by the Conservancy were outlined in "Surviving Success: An Interview with the Nature Conservancy's John Sawhill" in the *Harvard Business Review* (Howard and Magretta, 1995); in "Measuring What Matters in Nonprofits," in the *McKinsey Quarterly* (Sawhill and Williamson, 2001); and in "Mission Impossible? Measuring Success in Nonprofit Organizations," an unpublished case study by John Sawhill, who was at that time the Conservancy's chief executive officer.

For more than 30 years, the Nature Conservancy self-assessed through a metric it called "Bucks and Acres": The "Bucks" represented the fund-raising side of the organization, and the "Acres" listed the amount of land protected. In the early 1990s, Conservancy managers began to see a divergence between the organization's *mission* (to preserve biologic diversity) and its *traditional metric*, which focused on raising money and sequestering land. In the interview with Howard and Magretta (1995, p. 11), Sawhill described an issue that many organizations encounter as they consider measurable outcomes and the linkages among mission, strategies, and assessment metrics (or tools):

We had to change because while we were doing a lot of good conservation work, there were more and more signs that we were not making significant progress toward accomplishing our mission ... we started to realize that those measures [Bucks and Acres] weren't giving us the right information. For-profit companies can look at their financial statements every day to see how they're doing: they're either making money or they're not. Without the discipline of the bottom line, it's easier for nonprofit organizations to get off track.

The Conservancy changed its strategy from measuring effectiveness by the amount of acquired land that provided refuge for rare species to a more thorough analysis of its efforts in preserving ecosystems. The Conservancy emphasized targeted outcomes (establishing targeted outcomes is often the initial step in evaluation). After extensive

Nature Conservancy's Evaluation Design Excerpt from Conservancy web site

There are four fundamental, related parts to [the Conservancy's] approach: Setting Priorities, Developing Strategies, Taking Action, Measuring Success. The Nature Conservancy defines conservation success as the long-term reduction of critical threats and the sustained maintenance or enhancement of biodiversity health.

- *The Conservancy regularly evaluates the key ecological factors (such as size, condition and landscape context) of selected conservation targets in an area. Scientific assessments of key factors lead to biodiversity health designations of very good, good, fair or poor.*
- *The Conservancy regularly measures the level of threat to the conservation targets at an area, ranking these threats as very high, high, medium or low.*

Collectively, these measures quantify our conservation impact. To hold the organization accountable for results, the Conservancy intends to measure success across full portfolios, not just the focal conservation areas where the Conservancy is taking direct action. After measuring our success in each portfolio, we then set priorities, beginning the process again.

(<http://nature.org/aboutus/howwework/about/art5724.htm>, accessed Oct. 2, 2002)

conversations with staff, leaders of other nonprofit organizations, attempts and frustrations with revised performance systems, analysis of old metrics, work with consulting firms, and an extensive literature review, the Conservancy adopted a revised set of metrics. Since 1997 the Conservancy's conservation mission and evaluation strategies together constitute four parts: setting priorities, developing strategies, taking action, and measuring success.

Sawhill and Williamson (2001) explained that the Conservancy's experiences were far from unique. They also determined that nonprofit organizations should accept the challenge of evaluation and approach assessment with a clear mission and small, manageable goals. The case study put it this way (Sawhill, p. 10):

A number of common themes emerged ... most organizations found it very difficult to measure impact. ... [T]he most effective nonprofit measures are invariably linked to clear, measurable goals and narrowly defined missions.

The Conservancy's process has evolved to find that scientific data are critical to elucidating the impact of its conservation efforts—something Bucks and Acres failed to measure.

Chapter Summary

Since the early 1990s, the growing push for accountability has led to government, foundation, and nonprofit use of evaluation to justify and improve programs. Evaluation is a rigorous process that, according to professional standards, should involve systematic inquiry, competence, integrity and honesty, and respect for people and public welfare.

Participants' experiences and program outcomes are assessed through a variety of methods, including specific evaluation frameworks and social science research methods:

Case studies	Logic models
Diaries	Log frames
Empowerment evaluation	Observation
Experiments	Pre- and post-assessment instruments
Field studies	Scientific data
Focus groups	Surveys
Interviews	Questionnaires

The next four chapters examine interventions and assessment techniques used in environmental education, museums, health programs, and social marketing.

References

- Howard, A., and J. Magretta. 1995. Surviving success: An interview with the Nature Conservancy's John Sawhill. *Harvard Business Review*. Online, accessed Dec. 29, 2002:
http://harvardbusinessonline.hbsp.harvard.edu/b02/en/common/item_detail.jhtml?id=95508.
- Katzenmeyer, C. 1997. Introducing this handbook. In *User-friendly handbook for mixed method evaluations*, ed. J. Frechtling and L. Sharp, 1–8. Arlington, Va.: National Science Foundation. Online, accessed Aug. 22, 2002:
http://www.ehr.nsf.gov/EHR/REC/pubs/NSF97-153/CHAP_1.HTM.
- Kirkpatrick, S. 2001. The program logic model: What, why and how? Online, accessed Dec. 29, 2002:
<http://www.charityvillage.com/charityvillage/research/rstrat3.html>.
- Odame, H. H. 2000. Engendering the logical framework. International Service for National Agricultural Research (ISNAR). Online, accessed Dec. 29, 2002:
<http://www.isnar.cgiar.org/gender/hambly.htm>.
- McLaughlin, M. W., and D. C. Phillips, eds. 1991. *Evaluation and education: At quarter century*. Nineteenth yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press.

- Sawhill, J. C. Mission impossible? Measuring success in nonprofit organizations. Unpublished draft, Nature Conservancy, Arlington, Va.
- Sawhill, J., and D. Williamson. 2001. Measuring what matters in nonprofits. *McKinsey Quarterly* 2:98–107. Online: accessed Aug. 29, 2002: http://www.mckinseyquarterly.com/article_page.asp?ar=1053&L2=33&L3=95.
- Sheppard, B., ed. 1999. Perspectives on outcome based evaluation for libraries and museums. Washington, D.C.: Institute of Museum and Library Services. Online, accessed July 20, 2002: <http://www.ims.gov/pubs/pdf/pubobe.pdf>.
- World Bank Group. 1999–2001. Interactive community planning: ZOPP: Goal oriented project planning. Online, accessed Oct. 10, 2002: <http://web.mit.edu/urbanupgrading/upgrading/issues-tools/tools/ZOPP.html>.
- Tyler, R. 1991. General statement on program evaluation. In *Evaluation and education: At quarter century*. Nineteenth yearbook of the National Society for the Study of Education, ed. M. McLaughlin and D. C. Phillips, 3–17. Chicago: University of Chicago Press.
- United Way. 1996. Measuring program outcomes: A practical approach. Online, accessed July 20, 2002: <http://national.unitedway.org/omrn/resources/mpo/model.cfm>.

Resources

Online Resources: Evaluation

American Evaluation Association <http://www.eval.org/>

AEA is a professional organization devoted to improving and promoting evaluation. Links on the web site provide access to online handbooks for designing and conducting evaluations: <http://www.eval.org/EvaluationLinks/default.htm>.

ERIC Clearinghouse on Assessment and Evaluation <http://ericae.net/pare/>

The online journal *Practical Assessment, Research & Evaluation* provides educators with information about research and evaluation at the local education agency level. The ERIC Assessment and Evaluation web site also includes meta-analyses of education research findings and meta-analysis tools: <http://ericae.net>.

Empowerment Evaluation

<http://www.stanford.edu/~davidf/empowermentevaluation.html#summary>

This site provides information about collaborative, participatory, and empowerment evaluation. It also offers links to other Internet sites and to free-software sources, associations, and other reference works to assist in evaluation.

Online Resources: Logic Models

The resources below include clear descriptions, guidelines for developing logic models, and examples of frameworks.

Israel, G. D. Using logic models for program development: *http://edis.ifas.ufl.edu/BODY_WC041* (requires a browser with frames and JavaScript abilities).

Kirkpatrick, S. 2001. The program logic model: What, why and how?

<http://www.charityvillage.com/charityvillage/research/rstrat3.html>

This article contains the steps for creating a logic model.

McNamara, C. Guidelines and framework for designing basic logic model:

http://www.managementhelp.net/np_progs/np_mod/org_frm.htm

Rockwell, K., and C. Bennett. Hierarchy for targeting outcomes and evaluating their achievement: *<http://deal.unl.edu/TOP/english/index.html>*. Description of Targeting Outcomes of Programs (TOP).

Team Technologies, Inc., logical frame software: *<http://www.zopp.com/>*.

Commercial software for the ZOPP model.

United Way of America. Outcome measurement resource network:

<http://national.unitedway.org/omrn/>.

W. K. Kellogg Foundation. Logic model development guide:

<http://www.wkkf.org/pubs/Pub3669.pdf>.

Online Resources**United Way of Toronto Program Effectiveness Clearinghouse**

www.unitedwaytoronto.com/PEOD/index.html.

This site identifies print and online guides and information sources on outcomes-based program effectiveness and organizational development.

Harvard Family Research Project. Learning from logic models in out-of-school time.

Word file:

http://www.gse.harvard.edu/hfrp/content/projects/afterschool/resources/learning_logic_models.doc.

Portable document format: *http://www.gse.harvard.edu/hfrp/content/projects/afterschool/resources/learning_logic_models.pdf*

American Museum of Natural History

<http://research.amnh.org/biodiversity/center/publ/pubdwnld.html>

Interpreting Biodiversity: A Manual for Environmental Educators in the Tropics is an easy-to-read guide that focuses on program development. Chapter 5 is about evaluation. Visitors to this site will need to register to download the publication.

Mark, M., G. T. Henry, and G. Julnes. 2000. *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs*. San Francisco: Jossey-Bass.

Scriven, M. 1991. Beyond formative and summative evaluation. In *Evaluation and education: At quarter century*. Nineteenth yearbook of the National Society for the Study of Education, ed. M. McLaughlin and D. C. Phillips, 19–61. Chicago: University of Chicago Press.

Stufflebeam, D. 2001. Evaluation models. *New Directions for Evaluation* 89:7–98. Abstract online, accessed Dec. 29, 2002:

<http://www3.interscience.wiley.com/cgi-bin/abstract/89016405/START>.

Taylor, S., and B. Serrell. 1991. *Try it! Improving exhibits through formative evaluation*. Washington D.C.: Association of Science–Technology Centers. New York: New York Hall of Science.

Toffolon-Weiss, M. M., J. T. Bertrand, and S. S. Terrell. 1999. The results framework—An innovative tool for program planning and evaluation. *Evaluation Review* 23(3):336–359.

Chapter 2

Environmental Education and Environmental Psychology

In This Chapter

Theoretical Frameworks

- Knowledge, attitudes, behavior
- Reasoned action and planned behavior
- Activator, behavior, consequence
- Responsible environmental behavior

Methodology

- Quantitative methods
- Qualitative methods
- Experimental designs
- Criticisms

Research Findings

- Responsible environmental behavior
- Knowledge and attitudes

Chapter Summary

- Theories and methods
- Key findings
- Summary of studies

References

Resources

General acceptance of Volker's criticism of environmental education research persists nearly three decades after the statement above (Bocarro and Richards, 1998; Henderson and Fox, 1994; Hoody, 1995; Leeming et al., 1993; Marcinkowski, 1998a). Ways to connect research results, organizations, and educators still have not developed fully.

Most program evaluation in environmental education focuses on one of three areas: individual programs, participants' predispositions (experience, knowledge), or participants' attitudes. There has been a theoretical shift away from the knowledge–attitude–behavior (KAB) model (which asserts that knowledge leads to greater awareness, which ultimately leads to responsible behavior) toward more complex models that suggest multiple factors influence behavior change. The behavioral theories imply that fostering responsible environmental behavior consists of targeting specific actions and teaching those actions in a context that is appropriate to program participants (relevant to their age, experience, knowledge, socioeconomic status, geography). As noted above, there is a weak link between theory and practice—meaning programs tend to emphasize attitude change and awareness more than behavioral skills training.

Traditional evaluation and research methods have included quantitative data collection techniques, including the use of self-reported rating scales and various pre- and post-intervention surveys. Increasingly, evaluations are based on mixed methods (using quantitative and qualitative designs). Often, multiple methods are used to provide a more complete picture of the effects of a program or intervention. Traditional research theories describe effective evaluation as a process that consists of criteria, data, and judgment. Inconsistencies and weaknesses in those three areas are pervasive in studies of environmental education programs. The questionable quality of research has emerged as a consistent theme in the literature.

[W]e are at that point in time when rhetoric and opinion must be substantiated by consolidating existing research efforts and focusing future efforts. ... [W]e must now be about the business of validating the assumptions and utilizing a research base if environmental education is to continue to advance.

(Volker, 1973, in Hoody, 1995, p. 2)

This chapter examines the theoretical frameworks that can be used to guide environmental education programs and program evaluation. Methodologies are described, and there is a discussion of current criticisms. The research findings include information gleaned from studies of programs and participant characteristics. A summary highlights the key ideas presented in this chapter. As noted in the introduction, studies in peer-reviewed publications make up the bulk of the literature reviewed for *Measuring Results* (given this criterion, in-house evaluation work with little supervision by a governing body or peer-review system receives little attention in this report).

Theoretical Frameworks

The goal of environmental education is to change the way people think about and interact with the environment, so it is not surprising that behavioral and educational theories have dominated explanations of environmental behavior formation (Bogner, 1998; Hwang et al., 2000; Marcinkowski, 1998a).

Environmental psychologists have encouraged the use of theories concerning social and developmental psychology, attitude and behavior formation, behavior modification, learning theory, and the study of personality factors (Bell et al., 1978). Specifically, elements of the model of reasoned action and planned behavior (Ajzen and Fishbein, 1986) provide the theoretical basis for many environmental education interventions (Hines et al., 1986/87; Leeming et al., 1993).

Four theoretical frameworks are frequently used to consider environmental education programs and there are clear similarities among them. Each model includes components that are concerned with knowledge and attitudes. *Knowledge* includes the mastery of facts, the understanding of action strategies, or issue information. *Attitude* can be used to refer to opinions about the environment or a particular place or issue, one's sense of responsibility, or the amount of confidence one has that creating change is within one's power. Most environmental educators and psychologists believe that education is the key to improving environmental behavior (Gigliotti, 1990; Zelenzy, 1999).

The Frameworks

- Knowledge, attitudes, and behavior
- Reasoned action and planned behavior
- Activator, behavior, consequence
- Responsible environmental behavior

Knowledge, Attitudes, Behavior

The confluence of environmental education and psychology is guided by the idea that knowledge will lead to changes in attitude, which in turn will influence behavior (Birch and Schwaab, 1983; Hwang et al., 2000; Hungerford and Volk, 1990; Newhouse, 1990; C. Ramsey and Rickson, 1976). Figure 2.1 is the KAB model.

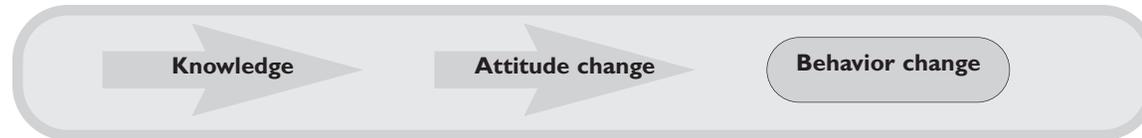


Figure 2.1. The KAB model posits that knowledge (or understanding of environmental issues as a result of becoming educated about them) leads to greater awareness (and to a change in attitude), which leads to more responsible behavior with respect to the environment.

The KAB model suggests that education about issues can influence behavior, but the empirical links between knowledge, attitudes, and behavior are tenuous: Some researchers believe there are strong correlations; others suggest weak connections if any at all (Gillett et al., 1991; Hines et al., 1986/87; Marcinkowski, 1998a,b).

Reasoned Action and Planned Behavior

The theory of reasoned action and planned behavior (Figure 2.2) emphasizes intention as the key determinant of behavior. In their work, Ajzen and Fishbein (1986) posited that intention precedes action, and, more specifically, that intention is based on one's attitude toward a behavior and is influenced by social factors. Participants in an environmental education program, for example, who voice an intention to act could be more likely to exhibit environmentally responsible behavior. The implication for evaluation is that statements of intention may serve as indicators of future action.

This theory is similar to the linear KAB model. The difference is that the theory of reasoned action identifies the personal and social influences that determine and influence attitudes. The theory presents "intention" as both a predictor and a determinant of action (Hwang et al., 2000, p. 20):

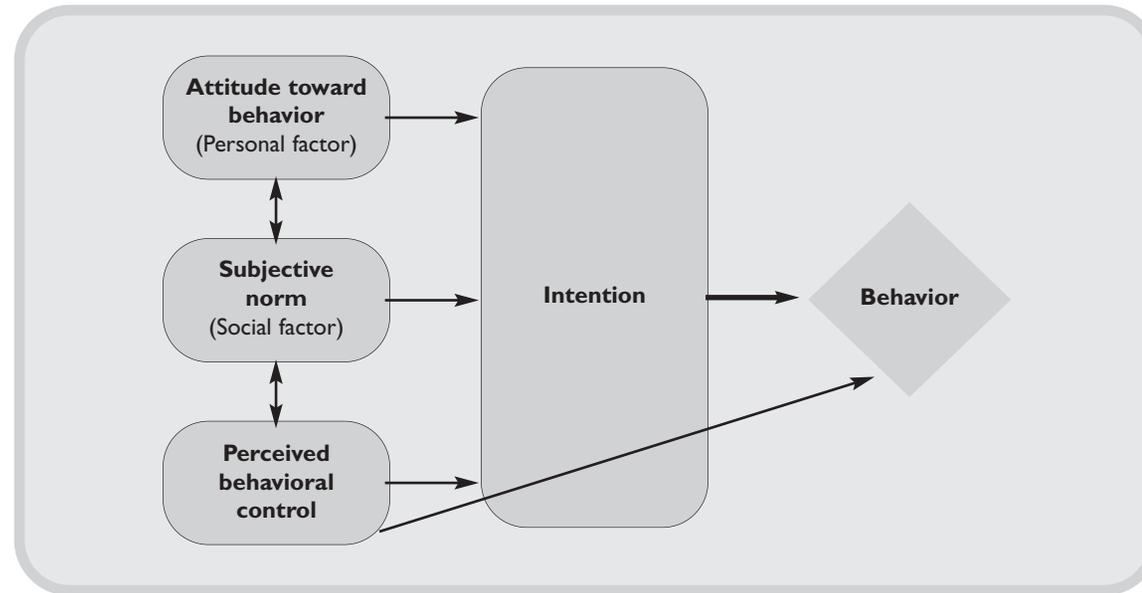


Figure 2.2. The theory of reasoned action and planned behavior proposes that someone's intention to act can be an effective predictor of action.

Intention to act has a direct effect on behavior, and can be predicted by attitude. The attitude is formed by subjective norms and belief, and the importance of those variables is decided by situational factors.

The theory of reasoned action and planned behavior credits several major influences on intentions to act and actual behavior, and a preponderance of studies privilege attitude as the guiding factor in behavior. Using the theory as a framework for environmental education suggests that programs that encourage the positive intention to act bring about responsible environmental behavior.

Over the past few decades researchers have explored whether attitudes alone are sufficient to affect behavior. Attitudes can have a causal relationship with behavior, but their influence is diminished in the face of other causal variables:

- Competence in carrying out behavior
- Experience with targeted behavior
- Sense of responsibility and subjective norms

Behavior analysts do not deny that attitude change can lead to behavior change, but they claim it is usually more cost-effective to target specific behavior directly than it is to expect a change in attitude to occur as a result of an intervention (Geller, 1992).

Activator–Behavior–Consequence

The ABC framework (Figure 2.3) shows multiple antecedents to behavior modification; they include (but are not limited to) attitudes. The emphasis is on the activators, to bring about action directly. Activators—antecedents or stimuli for targeted behavior—include products, such as recycling bins; informative pamphlets and signs; education; positive modeling; or petitions, for example. The ABC model addresses behavior change by defining action items in terms of relevant overt behaviors.

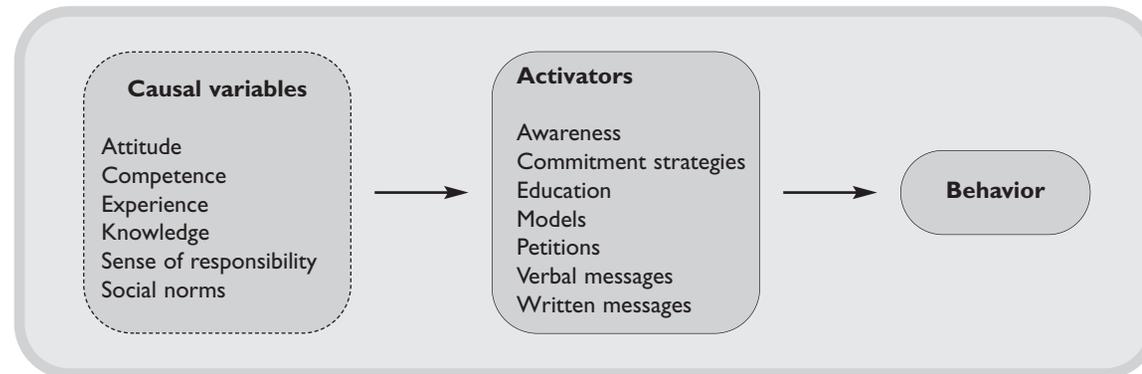


Figure 2.3. ABC framework.

The ABC model suggests that environmental educators should design interventions that focus directly on action. An intervention could be designed to decrease an undesirable behavior (such as littering) or to increase a desirable one (recycling) by focusing on the activators listed in the figure above.

Responsible Environmental Behavior

Hines and colleagues (1986/87) based their model of responsible environmental behavior (REB) (Figure 2.4) on results from a meta-analysis of 128 empirical studies. The

model represents three areas of development or change: affective, cognitive, and situational. The researchers' state (Hines et al., 1986/87, p. 7):

Abilities alone are not sufficient to lead to action. In addition, an individual must possess a desire to act. One's desire to act appears to be affected by a host of personality factors ... locus of control, attitudes, and personal responsibility. Thus, an individual with an internal locus of control, positive attitudes toward the environment and toward taking action, and with a sense of obligation toward the environment will likely develop a desire to take action. If the requisite abilities to act are also present, action will likely follow.

The REB model is clearly similar to the ABC model and the model of reasoned action and planned behavior, given the emphases on intention, individual characteristics, and social factors.

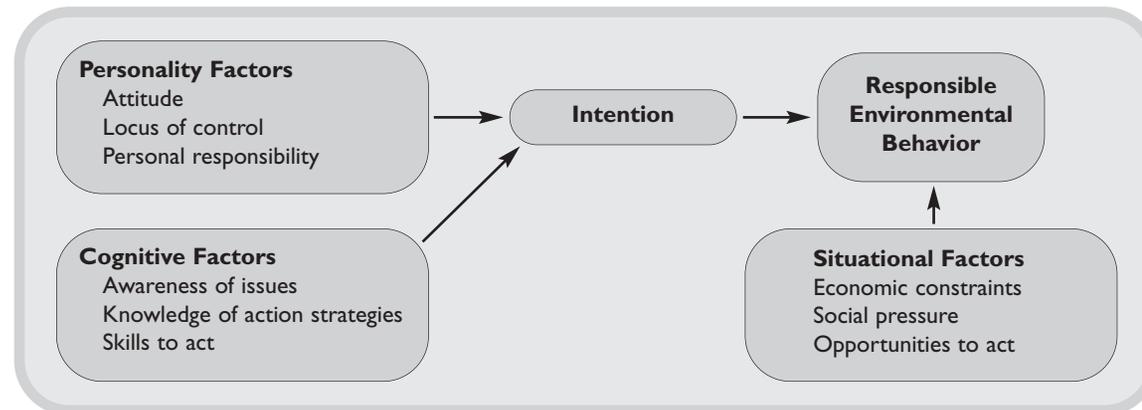


Figure 2.4. REB model (adapted from Hines et al., 1986/87).

Using the model in Figure 2.4 as a framework for environmental education suggests that educators can target change by affecting participants' cognitive and personality factors. In a practical sense, interventions that focus on developing skills, that prompt issue recognition and understanding, and that encourage specific action can lead to the intention to act and subsequently to responsible environmental behavior.

Simmons (1991) analyzed the stated goals of 1225 nature and environmental centers and reported that two-thirds used some variant of the KAB model. Simmons's review reported that only 36% consider action skills influential in behavior formation.

Methodology

Bell and colleagues (1978) attribute the methodologic “diversity, eclecticism, and innovation” used in environmental psychology research to field studies that focus on real issues and on connections among attitudes, behavior, and the environment. Environmental research methodology is driven primarily by the research topic, and it generally is based on multiple strategies (Bell et al., 1978). Data collection relies heavily on self-reported information that is obtained through interviews, questionnaires, and surveys, for example. Observation, consumption habits, and authentic assessment also can be used as sources of data.

Assessments of environmental education programs generally use quantitative, qualitative, or experimental designs, depending on the kind of information sought and the research goal. More often, multiple methods are used to provide a more complete picture of the effects of a program or intervention. For example, one study on how camps conduct self-assessment reported that most evaluation data were collected through questionnaires, interviews, and observations (Henderson and Bialeschki, 1994). The use of quantitative measurements and open-ended questions on the same topics gave researchers qualitative data that provided a framework for and depth to the numbers gained from Likert scales and “yes–no” queries. Virtually all of the camps in the study reported that staff, administration, facilities, and programs were evaluated at least once per year. Evaluations were conducted by program directors, fewer than half of whom had any formal training in evaluation.

Direct observation would probably be the most valid source of information for assessing whether behavioral changes have occurred. However, this is difficult, expensive, and time-consuming. Tracking gas, electricity, and water consumption or purchases (types of food or packaging selection, for example) are ways of measuring behavior without relying on self-stated data.

What Is Measured?

Knowledge

- *Basic ecology*
- *Environmental issues*
- *Environmental problem solving*
- *Specific actions or skills*

Attitude

- *Toward science*
- *Toward the environment*
- *Toward taking action*

Behavior

- *Intention*
- *Actions*

Measuring change often consists of pre- and post-activity assessments.

Quantitative Tools**Bipolar adjective scale**

Negative–positive;
fascinating–boring

Likert scales

Strongly agree–strongly
disagree;
1 (poor) – 5 (excellent)

Standard assessment forms

- *Coopersmith Self-Esteem Inventory*
- *Millward Gitner Outdoor Attitude Inventory*
- *Rosenberg Self-Esteem Scale*
- *Secondary School Environmental Literacy Instrument*
- *Tennessee Self-Concept Scale*

Quantitative Methods

Quantitative methods that use standardized measures to fit diverse information into pre-determined response categories have dominated the analysis of environmental education (Hines et al., 1986/87; Hoody, 1995; Leeming et al., 1993). Quantitative techniques most often take the form of pencil-and-paper instruments administered before and after an intervention as part of an experimental or quasi-experimental design. Hence, much of the information reported could best be described as the product of quantitative psycho- or sociometric research (Hoody, 1995).

Quantitative analysis of environmental education seems like a straightforward process, but the dimensions of and approaches to measuring knowledge, attitudes, and behavior changes are numerous and varied. The results of a particular evaluation might not be widely applicable, and numbers can obscure the larger context in which programs operate and in which participants live. Several quantitative tools are used to measure the effects of programming on participants, generally reducing those effects to data that can be used to identify variables of interest to a particular program.

Multiple valuation rating scales ask respondents to evaluate the extent to which they agree or disagree with a specific referent, generally along a dichotomous or Likert (5-point) scale. Ma and Bateson (1999) measured attitudes about the environment using a Likert scale and the following statements:

- Highway speed limits should be made lower so that we can save gasoline.
- Farmers and ranchers should be able to use any chemical sprays they think are necessary.
- We can use all the natural gas, oil, and gasoline we need now because future generations will find new forms of energy.

The investigators examined the correlation between attitudes toward science and attitudes toward the environment. Measured in a similar fashion, attitudes toward science were based on reactions to the following statements:

- I like to study science in school.
- I do not enjoy science.
- Everyone should learn about science.

With such instruments, participants are asked to indicate knowledge, attitudes, or behavior related to environmental issues and to identify their perceived relationships to or roles in the environment before and after they participate in an educational program.

Authentic Assessment Opportunity

Many environmental education programs include journal-writing, and students create portfolios by selecting pieces from their journals, explaining the significance of those assignments, and documenting what they have learned. The portfolios show individual growth and change, provide information on program goals and results, and can even identify other outcomes. For example, Moorcroft and colleagues (2000, p. 24) describe the value of authentic assessment in the evaluation of an ecology center program:

We have been able to document increases in students' knowledge and application skills, while identifying program areas in need of improvement. We also have discovered that our preconceptions of what students would learn and what they actually did learn were at times vastly different.

Portfolios provide information that charts participant growth and learning during a program. Tracking progress over time provides information on how students build on their experiences over the course of a program.

Qualitative Methods

Qualitative methods are used to capture data through case studies, correlation studies, ethnographies, observations, personal interviews, and open-ended questions on surveys, which seek to explain relationships between factors. Qualitative methods are used to identify the results of specific interventions, explain the participant experience, and supply longitudinal information.

Thomas (1990) reported that the case study is one effective means of collecting information to answer “how” and “why” questions about a program’s effectiveness. Because the case study allows all stake-holders’ views to be expressed, it can provide a more complete picture of a program and its components. For instance, a case study evaluation of a summer environmental program for young people could involve gaining an understanding of the program through interviews with staff members, organizations that provide funding, and past participants; analyzing documents (program marketing material, newspaper articles); observing the conduct of the program and interviewing program managers and participants during the program and afterwards; and analyzing the data and reporting the program in the context of the viewpoints expressed by various stakeholders, in this case, granting agencies, managers, and participants. According to Thomas, there are two steps in undertaking a case study: First, the data must be collected and a record of the data must be established. Then, the data are analyzed and reported.

In addition to the methods described above, several authentic-assessment tools—including student portfolios—provide information about participants’ experiences and can be used to evaluate learning and changes in knowledge, skills, attitudes, or behavior. Authentic assessment, as it is adopted from formal education evaluation (traditional classrooms), consists of student-centered activities that encourage learning. Unlike standardized testing and other traditional quantitative evaluation techniques, authentic assessment consists of learning. The instructor reviews and charts progress by examining work student work in portfolios, essays, scientific investigation, and knowledge-and-process answers to teacher-driven questions. Because this technique is not in widespread use in environmental education programs, the literature on the subject is limited.

Experimental Designs

The evaluation of environmental education programs generally does not involve traditional experimental design, given the complexity of establishing treatment and control groups, isolating independent variables, and following up with participants. However, some experiments have been reported.

Bryant and Hungerford (1979) created an experiment to evaluate the outcomes of an environmental education intervention for kindergarten students. The intervention lasted four weeks. The first week included an introductory module focused on the term and concept “environment” and the following three weeks included activity-based lessons on environmental issues and action strategies (pollution, waste management). All of the students were interviewed after the introductory module, at the end of the final week of the program, and again three weeks after the program ended. Responses were rated and

Description and Dialogue

One component of Schneider and colleagues’ (2002) evaluation of a short-term residential environmental education program used observation, interviews, and focus groups to better identify and explain students’ learning experiences. The evaluators relied on descriptive characterizations of program activities and students’ responses to illuminate the teaching and learning experience. Observation protocols highlighted the types of questions and lessons initiated by field science instructors and students’ reactions and comments.

Using a quantitative assessment—such as counting the number of questions generated by instructors versus the number generated by students—would not have provided as clear a picture of the student experience.

Design Challenge

Leeming and colleagues’ (1993) meta-analysis of 34 experimental studies related to participants’ knowledge, attitudes, and behavior included 17 classroom interventions and 17 informal programs. They concluded that the use of experimental design is fraught with shortcomings (p. 223):

Investigators in the field must become more sensitive to the importance of rigorous designs that include meaningful control groups, controls for expectancy effects, use of reliable and valid dependent measures, follow-up data, and appropriate unit analyses.

analyzed. The design used a modified rotation that allowed first the treatment group, and later the control group, to participate in the educational intervention. This provided additional data to support research findings. Bryant and Hungerford (1979) concluded that kindergarten children can understand environmental issues and distinguish between responsible actions that they can take and that others can take.

Criticisms

Traditional research theories describe effective evaluation as a process that consists of criteria, data, and judgment. Inconsistencies and weaknesses in those three areas are pervasive in studies of environmental education programs. The questionable quality of research has emerged as a consistent theme in the literature. The attention to personal variables (age, gender, education, socioeconomic status) that shape environmentally responsible behavior ignores the larger social, historical, and political contexts in which people operate. Other weaknesses include temporal stability, the heavy reliance on self-reported data, and the all too common poor design of assessments. These weaknesses are described below.

Temporal Stability

Evaluation of environmental education has been dominated by one-time outcome studies (Bocarro and Richards, 1998; Henderson and Fox, 1994) that make it difficult to understand findings in context because they provide little basis for comparison with other studies, interventions, or outcomes. More work on the systemic aspects of the evaluation process could generate widely applicable findings that would better inform program providers and the field at large.

Room for Improvement

A report prepared by the National Environmental Education Advisory Council (NEEAC, 1998), criticizes environmental education evaluation and research for a lack of rigor it attributes to several shortcomings:

- *Complexity of measuring long-term effects (There are few longitudinal studies—or even follow-up studies—that would provide information on long-term outcomes.*
- *Difficulties in identifying quantitative objectives*
- *Limited funding*
- *Poor design*

Overreliance on Self-Reported Data

Hungerford and Payton (1986) are quoted in Thomas (1990, p. 4):

However, since the kinds of behaviors targeted by [environmental education] are in most instances autonomously applied, direct observation is usually impossible due to an enormous expenditure of time and money it would necessitate.

Given the challenges of monitoring participant actions and potential behavior changes, self-stated intentions are often used as a measure of behavior. Self-reports often are subject to recall bias—subjects remember the facts incompletely, or interpret them in ways that are inconsistent with what they actually have done—and so cannot always be considered accurate (Bocarro and Richards, 1998; Hines et al., 1986/87; Hwang et al., 2000; Leeming et al., 1993; Marcinkowski, 1998a). Very few studies actually measure behavior (Hines et al., 1986/87; Leeming et al., 1993; Zelenzy, 1999).

Poor Study Design

By its nature, environmental education evaluation is plagued by poor design: It is difficult to measure human behavior because human behavior is complex and tracking participants is a difficult task. Poorly designed studies—characterized by the lack of a valid control group, the failure to consider multivariable influences, and the use of weak measurement tools—confound the ability to draw conclusions from several studies as a group. There are shortcomings in attempting to use traditional scientific experimental standards to evaluate participant behavior or program outcomes beyond the confines of the laboratory. By way of example, the NEEAC report (1998) cited an assessment of LIVE (Living in Volunteer Experiences), an adventure-based counseling program in Nova Scotia, Canada. The LIVE evaluation (Bocarro and Richards, 1998) did not consider pedagogy or specific program components. The result was to limit the view of the program's ability to change participants' behavior and attitudes. Moreover, the control group consisted of young people who had applied but who had not been selected to participate in the program.

What Affects KAB?

- *Age*
- *Attitude*
- *Education*
- *Experiences in nature*
- *Gender*
- *Income*
- *Knowledge*
- *Locus of control*
- *Occupation*
- *Parental behavior*
- *Parental values*
- *Personal responsibility*

Research Findings

Research on environmental education has explored many interventions, including programs at camps, in the classroom, on field trips, in short-term residential programs, and in extended-stay outdoor programs. Research also has examined the results of product tagging (using product labels that identify a commodity as recycled, organic, or in some other way better for the environment) and the influence of informational pamphlets on the knowledge, attitudes, and behavior of particular groups of people.

This section summarizes major findings concerning participant variables and the results of environmental education interventions. Knowledge, attitudes, and behavior toward the environment are influenced by a constellation of individual and social factors. The most salient tend to be age, education, gender, knowledge, and experience, although others intervene as well.

There are many studies that indicate the importance of individual variables, but on the whole, there is no systematic collection within an indicator or across studies to make definitive statements about the variables. It is clear, though, that individual characteristics are important and play a role in people's experiences and in how they learn. Table 2.1 provides examples from the literature that are suggestive of how individual variables influence program outcomes.

Table 2.1. Individual Variables Can Affect Program Outcome

Age	Few studies measure interventions that work with children in primary grades (K–2) despite positive results in that age group (Basile, 2000; Bryant and Hungerford, 1979).
Gender	Females exhibit more responsible environmental behavior than do males. Men scored higher in environmental knowledge and reported more negative attitudes toward the environment than did women (Tikka et al., 2000).
Education	Biology students reported more positive attitudes and were the most active outdoors-people among university students (Tikka et al., 2000).
Experience	A resident survey about Michigan groundwater reported significant differences in awareness and perceptions of environmental issues correlated with respondent profession. For example, responses on a Likert scale showed that more nonfarmers than farmers believed that a specific land use (agrochemicals) endangers groundwater (Suvedi et al., 2000).

The literature includes studies that have sought to better explain the genesis of responsible environmental behavior. Louise Chawla (1999), for example, interviewed 56 environmental professionals in Kentucky and Norway and reported that the most prominent factors affecting environmental action and commitment were the following:

1. Experience in an outdoor setting as a young person
2. Family values and actions
3. Membership in outdoor and environmental groups
4. Observation and awareness of destructive events or conditions (habitat destruction, pollution, radiation)
5. Education (memorable teachers, university classes, hands-on activities)

Almost all of those interviewed referred to multiple sources, and the average number of factors was four. Chawla's study was based on 1- to 2-hour-long interviews, and it provides personal anecdotes and reflections on how respondents developed a personal sense of responsibility for the environment.

Responsible Environmental Behavior

Changes in behavior are the ultimate goal of most environmental education programs. The literature devoted to environmentally responsible behavior has expanded dramatically

Predictors of Responsible Environmental Behavior

- *Environmental sensitivity*
- *Knowledge of action strategies*
- *Skill in action strategies (ability to perform)*
- *Locus of control*

Knowing How to Act

A study by De Young in 1988–1989 (in De Young, 1996) reported that procedural knowledge distinguished people who recycle from those who did not. Both those who did and those who did not recycle held similar pro-recycling attitudes and had the same extrinsic motivations to recycle. According to Linn and colleagues (1994), people adopt responsible environmental behavior, such as recycling, environmentally conscious consumption, and resource reduction, based on:

- Convenience (proximity to recycling bins, for example)
- Social pressure (personal contact, public commitment)
- Intrinsic motivation

in the past three decades (Hines et al., 1986/87; Hwang et al., 2000; Marcinkowski, 1998a). Programs that have a positive effect on behavior change actually train participants for specific behaviors (Horsely, 1977, in C. Ramsey and Rickson, 1976). People need to know *why* and *how* to behave in environmentally responsible ways (De Young, 1996; Geller, 1992; Hanna, 1995; Hines et al., 1986/87; Mittelstaedt, et al., 1999; C. Ramsey and Rickson, 1976).

Chawla's findings support conclusions from other studies. For example, Marcinkowski's (1998b) critical analysis of three dissertation studies reported that environmental sensitivity, knowledge of action strategies, skills to act, and locus of control are important determinants of behavior. In fact, according to Marcinkowski, the traditional KAB model does not accurately explain the core variables that lead to responsible environmental behavior: Environmental sensitivity, appreciation of and concern for natural surroundings, and knowledge of action strategies are important, but it also is necessary to understand the specific behavior required to address environmental issues (to know that creating healthy habitats, using resources wisely, and restoring degraded areas help restore biodiversity, for example). It is necessary also to have some skill in using action strategy—it is not enough to *know* the strategy, one must be able to perform an action (one skill related to the action strategy of creating healthy habitats to promote biodiversity, for example, would involve the integration of a diverse array of native plant species). Finally, locus of control, or the perceived ability to change a situation, is important.

Antecedents to Environmentally Responsible Behavior

Antecedents, or activators, as Geller (1992) refers to them, are prompts or triggers that increase the frequency of desirable behaviors or decrease the frequency of undesirable ones. Antecedents can educate people about action strategies or provide them with the tools they need to take action. As illustrated in the ABC model, activators prompt behavior directly. Antecedents can come in any one of several forms:

- Spoken or written messages
- Awareness or education interventions
- Role modeling or demonstrations
- Goal setting or commitment strategies
- Engineering or design strategies

Research supports the effectiveness of influencing responsible behavior through the use of prompts (Bell et al., 1978; Geller, 1992; Marcinkowski, 1998b). Written and spoken prompts include flyers; personal reminders; and informational feedback systems, such as utility bills that show individual household usage; or public commitment cards (decals, pledge cards). In recycling efforts, such interventions seem to work beyond experiment: Recycling increased among those who had signed commitment cards (Burn and Oskamp, 1986). Individual and social factors undoubtedly influence the salience and effectiveness of prompts.

Knowledge of action strategies and the ability to use them can promote responsible environmental behavior.

Locus of Control

Locus of control, or self-efficacy, is the perceived ability to bring about change through one's own actions. Hwang and colleagues (2000) reported that locus of control and attitudes were more important than knowledge in forming an individual's intention to act. In addition, in "Predictors of Responsible Environmental Behavior," Marcinkowski (1998b) replicates earlier studies and concludes that each depicts the role of high self-efficacy in taking responsible environmental action. Marcinkowski's conclusions are based on data from responses by members of the Sierra Club, Elderhostel, Ducks Unlimited, and the Audubon Society. The findings suggest that educators can improve programming by promoting participants' confidence in their ability to make an impact. Researchers have repeatedly shown that, in addition to self-efficacy, specific information about action strategies is an important predictor of responsible environmental behavior (Bell et al., 1978; Geller, 1992; Hsu, 1997, Lierman, 1996, and Sia, 1985, in Marcinkowski, 1998b).

Knowledge and Attitudes

Program effectiveness and participant outcomes can be classified into two categories: knowledge and attitudes, and responsible environmental behavior. Even though the ultimate goal of most interventions is to motivate responsible environmental behavior, the evaluative research has generally focused on acquisition of knowledge and changes in attitude. Table 2.2 lists the multiple variables explored in the research.

The outcomes often evaluated by environmental educators appear to be connected with knowledge or literacy (knowledge of ecology and environmental problems, problem-solving skills, sensitivity toward the environment, students' intention to participate, and their actual participation in service projects). A meta-analysis of environmental education studies (Leeming et al., 1993) showed that only 34 studies published in a 20-year period measured changes in environmental knowledge, attitudes, or behavior. Most of the 34 studies investigated the effects of educational interventions on participants' knowledge and attitudes; just five assessed the impact of interventions on participants' behaviors. The findings noted below depict what has been studied and what has been found to be effective or inconclusive.

Attitudes

The general acceptance of the KAB model has meant that interventions and studies of interventions have focused extensively in one area: Attitude is the most-studied variable in environmental education. Simmons (1991, p. 19) explains:

Respondents from some centers saw attitudes and attitude change as essential to the role played by nature [centers] and environmental education centers.

However, Leeming and colleagues (1993) reported that 11 of the 14 experimental studies that focused exclusively on attitude change showed mixed or negative results. The studies involved various interventions—field trips, camp activities, formal lectures, and a college course.

Some studies show that interventions can change attitudes toward the environment.

Changing Attitudes

Dettman-Easler and Pease (1999) advocate the use of residential environmental education programs because they offer flexibility, more time for programming (evenings are available), and extended time in nature. In their study of six residential programs in the Midwest, the investigators learned that students had significantly more positive attitudes toward wildlife after participating in residential programs than did a control group of students who either did not participate at all or who had minimal involvement with wildlife programming. The finding of improved attitudes (derived from pre- and post-program questionnaires and follow-up interviews) is supported by the work of other researchers (Armstrong and Impara, 1991, in Leeming et al., 1993; Bogner, 1998; Hines et al., 1986/87; JAVS 1982, 1984, in Leeming et al., 1993).

Positive-attitude results were reported in studies that evaluated multiple variables, including attitude. Outdoor experiences, with classroom reinforcement, increase knowledge and foster positive attitudes toward the environment. Classroom reinforcement often includes pre- and post-trip activities and curricula. For instance, if students attend a residential program to study water ecology, they can return to school and test water in nearby streams.

Ma and Bateson (1999) found no research literature on the correlation between attitudes toward science and attitudes toward the environment. Using data collected from more than a thousand 9th-grade students, the investigators compared attitudes toward science and attitudes toward environmental issues. Two significant correlations emerged:

- Students who voiced positive attitudes toward the environment also had positive attitudes toward science.
- Students who valued preservation and use of resources also noted the importance of learning science but had a lack of interest in learning science (that is, they believed it was important to learn science even though they were not interested in doing so).

The researchers posited that evaluators and educators could use attitudes toward science as an indicator of participants' attitudes toward the environment.

Knowledge

A meta-analysis by Hines and colleagues (1986/87) showed a positive correlation between knowledge and taking environmentally responsible action: The more people knew about issues and about how to take action, the more likely they were to exhibit responsible behavior.

Another study reflected gains in content knowledge: Bogner (1998) reported that participants had retained information about ecology and conservation issues in an assessment 1 month (and, for a subsample, 6 months) after 1- and 5-day outdoor programs. There was a significant increase in knowledge 1 month after the program for participants in both programs. Similar gains were reported in the assessment of a 6-day-long wilderness program (Gillett et al., 1991). The knowledge component of that study included

Knowledge Gain?

Researchers and educators have many reasons for assessing knowledge, including the need to determine the effectiveness of an intervention in promoting an understanding of facts, concepts, and skills. Knowledge assessment can improve programs because its results can be used to identify the strategies that contribute to participants' understanding.

information on environmental principles, National Park regulations, and camping procedures. In post-program surveys, the treatment group—12th-grade students on the wilderness trip—showed a larger gain in knowledge than did the control group, which did not participate in the program.

Evaluators often use pre- and post-program assessments or experimental designs to quantify participants' knowledge. One study used a case-control design and pre- and post-program assessments to examine the outcomes of a 2-week-long water conservation program for 6th-grade students (Aird and Tomera, 1977, in Leeming et al., 1993). The students in the treatment group received direct instruction, monitored and recorded their water use, presented oral reports on water conservation, and discussed personal conservation methods. The students who participated in the program showed significant knowledge gain evidenced by listing threats to the water supply and naming conservation actions (the control group could not).

Another study used post-program questionnaires a year after students had visited a conservation program site (Ryan, 1991). Unlike the evidence given above (Aird and Tomera, 1977, in Leeming et al., 1993; Bogner, 1998), Ryan's follow-up showed no knowledge gain. One section of the post-program assessment form tested concept knowledge by asking the 5th- and 6th-grade participants to draw a picture of a food chain. The topic was one the program educators had emphasized during the program. About 40% of the students drew nothing, and half of the responses from the rest of the group were incorrect. Ryan's finding showed the program did not achieve the intended goals as they pertained to the concept of the food chain. As noted throughout this chapter, much of the work in environmental education is based on one-time or program-specific assessments that make it difficult to draw overall connections between topics or to compare findings.

Chapter Summary**Theories and Methods**

Evaluation in environmental education is theoretically based. In recent years, the tenets of the traditional KAB model have been challenged in recognition of the complex factors impacting human behavior and action.

Evaluation Challenges

- *Measuring actual behavior*
- *Emphasis on self-reported measures*
- *Lack of funding*
- *Poor construct validity: few valid control groups, small sample sizes*
- *Personal variables ignore larger context*
- *Temporal stability*

- *Knowledge, attitudes, and behavior.* The KAB model asserts that education leads to greater awareness and attitude change, which ultimately leads to responsible behavior.
- *Reasoned action and planned behavior.* Azjen and Fishbein (1986) posit that intention predicates behavior. Intention to act directly influences behavior.
- *Activator–behavior–consequence.* The ABC construct suggests that it is usually more cost-effective to target behaviors directly than it is to expect attitude change to occur as a result of an intervention.
- *Responsible environmental behavior.* The REB model is based on results from a meta-analysis of 128 empirical environmental behavior studies (Hines et al., 1986/87). The model represents three areas of development: affective, cognitive, and situational.

Methods have traditionally focused on quantitative data collection techniques. More recently, evaluators have recognized the value of mixed-method designs to gain deeper insights into participants' experience and program outcomes. The questionable quality of environmental education studies has emerged as a consistent theme in the literature reviewed. Ways to connect research results, organizations, and educators have not been fully developed.

Key Findings

Environmental education research cites the following in bringing about positive environmental behavior:

- Environmental sensitivity
- Knowledge of specific targeted behaviors
- Skill in acting out specific behaviors
- Strong self-efficacy

People need to know *why* and *how* to act in environmentally responsible ways. Effective programs train participants for specific behaviors. In addition, antecedents (prompts or triggers) increase the frequency of desirable behaviors and decrease the frequency of undesirable ones. Examples of antecedents include

- Goal setting and commitment strategies
- Spoken messages or written materials such as flyers, personal reminders, and informational feedback systems (utility bills that show individual household energy use, for example)
- Role modeling or demonstrations

In a practical sense, interventions that focus on developing skills, that prompt issue recognition and understanding, *and* that encourage specific actions can lead to responsible environmental behavior.

Summary of Studies

Table 2.2 summarizes environmental education evaluation studies and the effects of interventions on environmental knowledge, attitudes, and behaviors. The table is sorted by the focus of the study (in many cases also the core goal of the intervention).

For example, if an intervention sought to affect a community's purchasing behavior, the study is listed under "behavior." Some studies examined more than one variable; they are listed under "multiple factors."

Table 2.2. Summary of Environmental Education Studies

Focus of Study	Effect	Intervention	Population	Method	Follow-Up	Reference	
Knowledge	+	Nature center, pond ecology	Elementary school students	Authentic assessment: portfolios, essays (authentic evaluation tools such as structured essays and portfolios consider performance over time)	No	Moorcroft et al., 2000	
Attitudes	+	5-Day residential program	Elementary school students	Millward Gitner Outdoor Attitude Inventory	12 Months	Mittelstaedt et al., 1999	
Attitudes	+	Residential camps and school	Grades 5–6	Pre-, post-activity questionnaire (compare camp and school)	3 Months	Dettman-Easler and Pease, 1999	
Attitudes	+	Conservation, teacher preparation	Grade 5	Likert scale	6 Months	Ryan, 1991	
Attitudes	+	School	Grade 9	Correlational study: science and environmental attitudes; Likert scale	No	Ma and Bateson, 1999	
Behaviors (reported)	+	Environmental education program	Grade 8	Issue investigation	No	J.M. Ramsey, 1993	
Multiple factors ^a Knowledge, attitude, locus of control, behavior (intention)	Mixed ^b	National park, nature trail (Korea)	Adults	Likert scale; semantic differential scale	No	Hwang et al., 2000	
Knowledge Intentions Media credibility	+	+	Media (print and television)	Adults	Telephone surveys, media analysis	No	Fortner et al., 2000
Correlational effects: KAB-gender	Mixed	College	Undergraduate students	Questionnaire (KAB–environment), interviews	No	Tikka et al., 2000	

Table 2.2. Summary of Environmental Education Studies *Continued*

Focus of Study	Effect	Intervention	Population	Method	Follow-Up	Reference
Attitude, behavior, self-esteem	Mixed ^c	Outdoor, inner city	At-risk, low-income youth	Pre-, post-activity surveys, qualitative interviews	Yes	Bocarro and Richards, 1998
Knowledge Attitudes Behaviors (reported)	+ + + ^d	National park, residential program	Grade 8	Experiment: pre-, post-activity surveys, case-control design	1 and 6 Months	Bogner, 1998
Attitudes Intentions Behavior	Mixed ^e	Residential: Audubon and Outward Bound	Adults	Pre-, post-activity surveys, interviews, observation, document analysis	6 Months	Hanna, 1995
Knowledge Attitudes Behaviors (reported)	+ + +	College	Undergraduate students	Course work	No	Smith-Sebasto, 1995
Awareness, self-reports of consuming behavior	-	Grocery stores	Adults	Pre-, post-activity telephone interviews	No	Linn et al., 1994
Attitudes, intentions	+	College	Undergraduate students	Course work, interviews	No	Gigliotti, 1992
Knowledge Attitude Self-concept	+ - +	6-Day trip	Grade 12	Experiment; pre-, post-program Tennessee Self-Concept Scale (TSES); Coopersmith Self-Esteem Inventory (SEI); environmental attitude and knowledge questionnaire	No	Gillett et al., 1991
Behaviors Attitude Locus of control Spoken commitment Sense of responsibility Knowledge of issues and action	+ + + + + +	Meta-analysis	NA	Meta-analysis of environmental behavior research	NA	Hines et al., 1986/87

Table 2.2. Summary of Environmental Education Studies *Continued*

Focus of Study	Effect	Intervention	Population	Method	Follow-Up	Reference
Other Evaluation practices	Mixed ^f	Camp	Camp directors	Questionnaires, Likert scales, qualitative analyses	No	Henderson and Bialeschki, 1994
Case study as evaluative tool	+ ^g	University (Australia)	Graduate students	Case study	NA	Thomas, 1990

Table key: *Focus of Study*, topic of study; *Effect*, positive (+), negative (-), or mixed; *Intervention*, type of environmental education program; *Population*, study participants; *Method*, data collection technique; *Follow-Up*, if and when collection took place after the intervention's conclusion; *Reference*, author and date of study; NA, not applicable.

^a This section sorted by reference date; all studies noted multiple emphases.

^b Largest correlations between locus of control, attitudes; locus of control, intention to act.

^c Program, research, design issues.

^d 5-Day only.

^e Audubon participants exhibited more pro-environmental intentions.

^f Evaluation used to improve program and staff; more training on methodology required.

^g Greater understanding among study population; all stakeholders represented.

References

- Armstrong, J. B., and J. C. Impara. 1991. The impact of an environmental education program on knowledge and attitude. *Journal of Environmental Behavior* 22(4):36–40.
- Azjen, I., and M. J. Fishbein. 1986. Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of Experimental Social Psychology* 22:453–474.
- Basile, C. G. 2000. Environmental education as a catalyst for transfer of learning in young children. *Journal of Environmental Education* 32(1):21–27.
- Bell, P. A., J. D. Fisher, and R. J. Loomis. 1978. *Environmental psychology*. Philadelphia: W. B. Saunders.
- Birch, S. K., and K. E. Schwaab. 1983. The effects of water conservation instruction on seventh grade students. *Journal of Environmental Education* 14:26–31.
- Bogner, F. 1998. The influence of short-term outdoor ecology education on long-term variables of environmental perspectives. *Journal of Environmental Education* 29(4):17–29.
- Bocarro, J., and A. Richards. 1998. Experiential research at-risk: The challenge of shifting traditional research paradigms. *Journal of Experiential Education* 21(2):102–107.
- Bryant, C. K., and H. R. Hungerford. 1979. An analysis of strategies for teaching environmental concepts and values clarification in kindergarten. *Journal of Environmental Education*, 9:44–49.
- Burn, S. M., and S. Oskamp. 1986. Increasing community recycling with persuasive communication and public commitment. *Journal of Applied Social Psychology* 16:29–41.
- Chawla, L. 1999. Life paths into effective environmental action. *Journal of Environmental Education* 31(1):15–26.
- Dettman-Easler, D., and J. Pease. 1999. Evaluating the effectiveness of residential environmental education programs in fostering positive attitudes toward wildlife. *Journal of Environmental Education* 31(1):33–39.
- De Young, R. 1996. Some psychological aspects of reduced consumption behavior, the role of intrinsic satisfaction, and competence motivation. *Environment and Behavior* 28(3):358–409.
- Fortner, R., J. Lee, J. Corney, S. Romanello, J. Bonnell, B. Luthy, C. Figuerido, and N. Ntsiko. 2000. Public understanding of climate change: Certainty and willingness to act. *Environmental Education Research* 6(2):127–141.

- Geller, E. S. 1992. Solving environmental problems: A behavior change perspective. In *Psychology and social responsibility: Facing global challenges*, ed. S. Staub and P. Green. New York: New York University Press.
- Gigliotti, L. (1990). Environmental education: What went wrong? What can be done? *Journal of Environmental Education* 22(1):9–12.
- Gillett, D. P., P. Thomas, R. L. Skok, and T. F. McLaughlin. 1991. The effects of wilderness camping and hiking on the self concept and the environmental attitudes and knowledge of twelfth graders. *Journal of Environmental Education* 22(3):33–44.
- Hanna, G. 1995. Wilderness-related environmental outcomes of adventure and ecology education programming. *Journal of Environmental Education* 27(1):21–32.
- Henderson, K., and K. Fox. 1994. Methods, measures, and madness: Possibilities for outdoor education research. *ERIC* 383:487.
- Henderson, K., and D. M. Bialeschki. 1994. The status of evaluation in ACA accredited camping programs. In *Coalition for Education in the Outdoors research symposium proceedings*, *ERIC ED* 383:494.
- Hines, J., H. Hungerford, and A. Tomera. 1986/87. Analysis and synthesis of research on responsible environmental behavior: Meta-analysis. *Journal of Environmental Education* 18(2):1–8.
- Hoody, L. 1995. The educational efficacy of environmental education: An interim report. State Education Environmental Roundtable.
- Hungerford, H., and Payton, R. B. 1986. Strategies for developing an environmental education curriculum. Paris: UNESCO–UNEP International Education Programme. *Environmental Education* 22:70–71.
- Hungerford, H., and T. Volk. 1990. Changing learner behavior through environmental education. *Journal of Environmental Education* 21(3):8–21.
- Hwang, Y., S. Kim, and J. Jeng. 2000. Examining the causal relationships among selected antecedents of responsible environmental behavior. *Journal of Environmental Education* 31(4):19–25.
- Leeming, F. C., W. O. Dwyer, B. E. Porter, and M. Cobern. 1993. Outcome research in environmental education: A critical review. In *Essential readings in environmental education*, 209–226. Champaign, Ill.: Stipes.
- Linn, N., J. Vinning, and P. A. Feeley. 1994. Toward a sustainable society: Waste minimization through environmentally conscious consuming. *Journal of Applied Social Psychology* 24(17):1550–1572.

- Ma, X., and D. Bateson. 1999. A multivariate analysis of the relationship between attitude toward science and attitude toward the environment. *Journal of Environmental Education* 31(1):27–32.
- Marcinkowski, T. 1998a. Assessment in environmental education. In *Essential readings in environmental education*, 171–208. Champaign, Ill.: Stipes.
- Marcinkowski, T. 1998b. Predictors of responsible environmental behavior: A review of three dissertation studies. In *Essential readings in environmental education*, 227–256. Champaign, Ill.: Stipes.
- Mittelstaedt, R., L. Sanker, and B. VanderVeer. 1999. Impact of a week-long experiential education program on environmental attitude and awareness. *Journal of Experiential Education* 22(3):138–148.
- Moorcroft, T., K. Desmarais, K. Hogan, and A. Berkowitz. 2000. Authentic assessment in the informal setting: How it can work for you. *Journal of Environmental Education* 31(3):20–24.
- NEEAC (National Environmental Education Advisory Council). 1998. Report assessing environmental education in the United States and the implementation of the National Environmental Education Act of 1990. EPA 171-R-96-001. Washington, D.C.: U.S. Environmental Protection Agency.
- Newhouse, N. 1990. Implications of attitude and behavior research for environmental conservation. *Journal of Environmental Education* 22(1):26–32.
- Ramsey, J. M. 1993. The effects of issue investigation and action training on eighth-grade students' environmental behavior. *Journal of Environmental Education* 24(3):31.
- Ramsey, C., and R. Rickson. 1976. Environmental knowledge and attitudes. *Journal of Environmental Education* 8(1):10–18.
- Ryan, C. 1991. The effect of a conservation program on school children's attitudes toward the environment. *Journal of Environmental Education*.
- Simmons, D. A. 1991. Are we meeting the goal of responsible environmental behavior? An examination of nature and environmental education center goals. *Journal of Environmental Education* 22(3):16–21.
- Schneider, B., N. Cheslock, and S. Funke. 2002. Evaluation of Yosemite National Institute's program institute, "Changing lives: A comprehensive action plan for science-based environmental education." Unpublished manuscript.
- Smith-Sebasto, N. J. 1995. The effects of an environmental studies course on selected variables related to environmentally responsible behavior. *Journal of Environmental Education* 26(4):30–35.

- Suvedi, M., D. Krueger, A. Sherestha, and D. Bettinghouse. 2000. Michigan citizens' knowledge and perceptions about groundwater. *Journal of Environmental Education* 31(1):16–21.
- Tikka, P. M., M. T. Kuitunen, and S. M. Tynys. 2000. The effects of educational background on students' attitudes, activity and knowledge concerning the environment. *Journal of Environmental Education* 31(3):12–19.
- Thomas, I. G. 1990. Evaluating environmental education programs using case studies. *Journal of Environmental Education* 21(2):3–8.
- Zelenzy, L. C. 1999. Educational interventions that improve environmental behaviors: Meta-analysis. *Journal of Environmental Education* 31(1):5–14.

Resources

Online Resources

Nature Conservancy http://www.consci.org/scp/download_the_handbook.htm

The Nature Conservancy's handbook, the Five-S Framework for Site Conservation, is a practitioner's guide designed to provide direction for determining conservation targets, analyzing threats, planning conservation strategies, and measuring success.

North American Association for Environmental Education <http://www.naaee.org/>

The NAAEE web site offers a variety of curriculum and program evaluation resources for educators.

City of Toronto Environmental Impact Assessment & Policy Development Office

<http://www.city.toronto.on.ca/eia/index.htm>

The Toronto, Ontario, municipal web site provides links to Internet sources and a questionnaire for determining an individual "ecological footprint" (a concept developed by William Rees, Mathis Wackernagel, and the Task Force on Healthy and Sustainable Communities at the University of British Columbia).

Other Resources

- Domroese, M. C., and E. J. Sterling. 1999. Interpreting biodiversity: A manual for environmental educators in the tropics. New York: American Museum of Natural History. Online: <http://research.amnh.org/biodiversity/center/publ/pubdwld.html>.
- Hart, P., and K. Nolan. 1999. A critical analysis of research in environmental education. *Studies in Science Education* 34:1-69.
- Hattie, J., H. Marsh, J. Neill, and G. E. Richards. 1997. Adventure education and Outward Bound: Out of class experiences that make a lasting difference. *Review of Educational Research* 67(1):43-87.

Chapter 3

Museums

In This Chapter

Theoretical Frameworks

- Contextual model of learning
- Conversational elaboration

Methodology

- Observation techniques
- Listening techniques
- Experimental designs
- Longitudinal studies

Research Findings

- Visitor characteristics
- Exhibit characteristics
- Family behavior and learning
- Field trips

Chapter Summary

- Theories and methods
- Key findings
- Summary of studies

References

Resources

For the purposes of this report, “museum” refers to such cultural institutions as aquaria, museums, nature centers, zoos, and other science learning settings. There is a great deal of heterogeneity among visitors to these venues. Understanding and measuring the learning that happens in and as a result of museum visits is difficult (Falk and Dierking, 2000; Hein, 1998; Lucas, 1991; Roberts, 1997). Stephen Weil (1999, p. 14) succinctly describes at least one challenge of museum evaluation:

[T]he impact of a museum tends to be subtle, indirect, frequently cumulative over time, and often intertwined with the impacts of such other forces of formal and informal educational experiences as schools, churches and various social and affinity groups. Museums will not only have to educate themselves as to how their impact can be captured and described, they will also have to educate those to whom they may be accountable as to what may and may not be possible in rendering their accounts. Daunting as those tasks may be, they will be essential.

This chapter provides a selected review and critique of museum research and evaluation to highlight the issues under analysis and describe the methods that have been used to assess visitors’ experiences. The studies noted include seminal pieces that illustrate the field and examples of current trends in research. Although a large portion of the work in this area is done in house and not shared or published in peer-reviewed journals, for the most part, the findings in this report are from published work in the field.

Theoretical Frameworks

Most work in the museum field has been in visitor studies. Visitor studies have relied on educational theories, and quantitative and qualitative methods have traditionally guided the work. In “The Maze and the Web” Hein (1997) clarifies how the opposing perspectives on how people learn (educational theories) influence program strategies and the work of evaluation.

[M]useum learning has many potential advantages: nurturing curiosity, improving motivation and attitudes, engaging the audience through participation and social interaction, and enrichment. By nurturing curiosity, the desire to learn can be enhanced.
(Ramey-Gassert et al., 1994, p. 351)

Contextual Model of Learning

Falk and Dierking (2000) use a framework to study the effects of learning in personal (agenda, experience, knowledge), physical (objects, setting, design), and sociocultural (sometimes collaborative, cultural setting) contexts. The contextual model of learning differs from most museum research, which generally examines only one context at a time—personal, physical, or sociocultural—in assessing exhibit effectiveness. In addition to the personal, physical and sociocultural elements of the model, Falk and Dierking include the importance of considering time, or the past, in better elucidating the learning that takes place in museums.

Conversational Elaboration

Leinhardt and Crowley (1998) and others have noted the lack of a common theoretical model to guide museum evaluation and research. Those researchers, who have examined museum learning in the context of visitors' conversations, also have noted that the multiple definitions of learning have guided different studies—making it difficult to compare studies or to identify the factors that are most effective or influential in museum visitor learning. *Learning* means many things, from how people retain discrete bits of factual information (along the lines of traditional transmission-absorption models) to “meaning-making” (applying prior knowledge and thinking to gain understanding) to more affective experiences (emotional responses to a setting; whole experiences). In response to this lack of a common definition or standard framework, Leinhardt and Crowley (1998) propose the conversational elaboration model.

Conversational elaboration (Figure 3.1) is a sociocultural model based on the belief that multiple factors affect learning and that the complex relationship of those factors must be considered in the assessment of learning. Conversational elaboration accounts for variations among individual visitors and in the physical environment. The evaluation or research model is based on collecting information about the learning environment, on group or individual identity (motivation, amount of interest, prior knowledge), and on visitors' interpretation and understanding of what they see. Leinhardt and Crowley (1998, p. 6) describe the role of conversation as “one of the primary means through which past experience is incorporated into current activity, and current experiences are carried to shape future activity.”

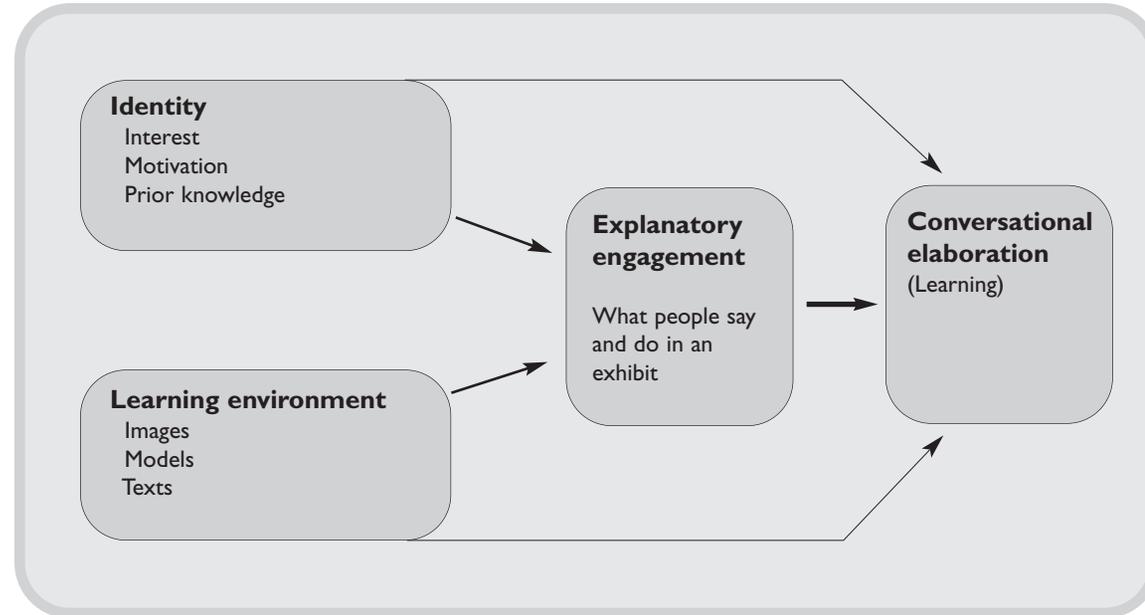


Figure 3.1. The conversational elaboration model is a sociocultural framework to explain learning in museums and to guide evaluation and research.

In an evaluation of participants' experiences in various museum settings, Leinhardt and colleagues (2000, p. 2) describe how conversational elaboration can be used as an indicator of learning:

By conversation we mean talk, occurring during and after a museum visit, that focuses on the way meaning, experiences, and interpretation develop. By elaboration we mean an extension of details, exemplification, emotive connection, and specificity ... by adding the diary to our interconnected web of studies we intend to illuminate some of the ways in which more completely developed elaborations of meaning unfold.

Underpinning this theory is that conversation and visitors' approaches to informal learning settings shape their experience and can indicate what learning or potential learning takes place on the visit. The conversational elaboration model shows one potential direction of the field.

Methodology

Visitor studies aim to explain who visitors are, what they think, and how they interpret their experiences (Hein, 1998). Museum research and evaluation are based primarily on observation and interviews. Assessments of visitors' experiences depend on what researchers observe and on what visitors tell them (in writing and orally). In addition to the formative and summative functions of evaluation described in the introductory chapter, museum work includes front-end evaluation. Called needs-assessment in other fields, that approach is used to in the development of a museum exhibit or program itself (see the Association of Science–Technology Centers web site, <http://www.astc.org/>, for examples of front-end studies).

Naturalistic studies and experimental designs alike have been used to capture information about visitors' knowledge, attitudes, and behavior (Hein, 1997). Common evaluation methods involve measuring the amount of time visitors spend at an exhibit (its holding power), monitoring visitors' movements (the exhibit's attracting power), and observing visitor interactions (both with the exhibit and with one another). Some data collection strategies (diaries, personal interviews, surveys, questionnaires, pre- and post-visit assessments) require direct visitor access and participation. Other methods are completely unobtrusive: tracking the number of fingerprints on glass, quantifying store purchases, or charting tread marks on floors. Table 3.1 summarizes in museum evaluation methods.

It is difficult to determine the long-term influence of museum experiences. Evaluation is constrained by various factors:

- Complexity in defining, measuring, and determining causal links between the experience and its effects on visitors
- Short duration and infrequency of visits
- Likelihood that time, distance, or a trigger outside the experience will be required before the full effects of the experience are realized (Bitgood, 1992; Falk and Dierking, 2000; Hein, 1998; Rennie and McClafferty, 1996; Schauble et al., 1997)
- Variations among individual visitors

Table 3.1. Museum Research and Evaluation Methods

Collection Method	Type of Data
Observation Tracking, time allocation	<ul style="list-style-type: none"> • Behavior patterns • Level of engagement • Duration of engagement
Naturalistic observation	<ul style="list-style-type: none"> • Behavior patterns • Interaction between visitors • Interaction between visitors and exhibit
Online	Web pages searched, pages reviewed, repeat visitors, referring URLs, content analysis
Listening techniques Diaries, interviews, surveys	<ul style="list-style-type: none"> • Demographic information • Visitors' perceptions, thoughts, feelings, content knowledge
Noting	Visitors' actions or evidence of their experiences during their visit (comment cards, sketches, changed computer entries, for example)
Experiments	Examining changes by exhibit type, contrasting results with a control group or between different groups
Longitudinal studies	Visitor outcomes beyond the immediate experience

Different ideas about learning further confound the effort to estimate how museums affect visitors' knowledge and attitudes. Recent work has demonstrated a move toward a multidimensional and sociocultural approach to learning and understanding visitors' museum experiences.

Observation Techniques

Museum research has often focused on the observable aspects of museum visits—visitors' movements and social interactions. Many and varied tools, including videotaping, photography, and tracking of fingerprints and tread marks, are used to assess visitors' experiences and exhibit effectiveness. Special printed forms are sometimes used to record specific observations systematically. For example, some studies collect data using checklists of behaviors or maps for tracking visitors' typical routes. Timing and tracking studies and naturalistic observations are two basic techniques that constitute the bulk of behavior studies. Hein (1998) provides a detailed analysis of the methods and tools used in the history of visitor research.

Timing and Tracking Studies

Traditional tracking studies provide information about visitors' basic movements and are considered the workhorse of visitor studies. Information about which routes visitors take, how many stops they make, where they stop (exhibit attracting power) and for how long (exhibit holding power) helps researchers identify patterns of exhibit use and is useful for elucidating patterns in visitor behavior. The correlation between time spent at an exhibit and learning was examined in a study that combined pre- and post-visit testing and unobtrusive observation with a hidden camera (Falk, 1983, in Dierking and Falk, 1994). A highly significant correlation between duration of viewing and learning was found. Common sense supports the finding that spending more time with an exhibit will lead to increased understanding of its subject matter.

Collections of quantitative data are an advantage of tracking studies; the numbers allow researchers to find generalizations and to make comparisons across sites. Serrell's (1993) summary of tracking studies describes a successful museum as one in which half of the visitors visit half the exhibits—a criterion that few museums or science centers meet. One assumption of tracking studies is that the amount of time spent at the exhibit is a marker of the influence the exhibit could have on visitor knowledge and attitudes.

How Much Time? Findings from Tracking Studies

- *The average museum visit lasts about 2 hours (Dierking and Falk, 1994; Hein, 1998; Taylor, 1986).*
- *Visitors experience "object satiation" or museum fatigue within 30-45 minutes (Falk et al., 1985, in Patterson and Bitgood, 1988).*
- *Family behavior patterns in museums are predictable (Dierking and Falk, 1994).*

Naturalistic Observations

Like tracking studies, naturalistic (or descriptive) observation studies monitor patrons' movements. Naturalistic studies, however, place greater emphasis on examining visitor interactions and participant reactions to and behavior toward specific exhibits. Evaluators often use trained observers and sometimes hidden video cameras to assess visitor behavior and interactions. Naturalistic studies are particularly good at revealing the social elements of a visit (Diamond, 1986). Evaluators collect data about interactions among visitors and about exhibit attracting and holding power. The family behavior and learning section in this chapter highlights findings of naturalistic observation studies.

Online Behavior

The widespread use and availability of the Internet has provided new opportunities for informal educational centers to communicate with a broad population and to measure usage patterns. The Museum Learning Collaborative (MLC) obtained information about users' interests and online behavior (Crowley et al., 2001b) by monitoring web sites for almost two years to assess the viability of disseminating research results online. The evaluators collected such information as the addresses of users' Internet connections, the URLs requested, the referring Internet service providers' addresses, and dates and times of requests. The researchers wanted to develop a profile of users of the MLC web site (<http://museumlearning.com/default.html>) and to learn about visitors' web use by tracking the number of hits, repeat visitors, and pages viewed.

Internet sites also offer a place to extend the museum environment. In "Supporting Scientific Inquiry Through Museum Websites," Cennamo and Eriksson (2001) evaluate the capacity of museum web sites to prompt inquiry. The researchers initially reviewed 100 sites, and they concentrated their study on 36. Given the focus on inquiry, the evaluation was based on examining the existence and depth of content that supported scientific inquiry (material for observation, prompting of questions, data analysis guides, for example). They reported that museum web sites support inquiry-based learning on various levels and in different ways. Namely, web sites can provide information, demonstrate an inquiry-based process, explain what is known about a particular topic, post on- and offline activities, and encourage collaboration. The evaluation provided formative

Formative Evaluation: Ocean Planet Exhibit

Researchers gathered information about visitor demographics, knowledge, and attitudes about the ocean and conservation before they developed the exhibit. A questionnaire was used for data collection by two sampling methods: Systematic interviews were spaced at pre-determined time intervals to reach every “nth” visitor. Quota sampling was used to provide information that the evaluation team could use to compare and contrast responses from target groups. The research was done to predict differences in exhibit outcomes at diverse locations. Findings suggest that geography affects visitor awareness of, but is not related to how they acquire information about, ocean issues.

information for the creation of the Muse-it web site, (<http://www.vtmnh.vt.edu/muse-it/default.htm>), which includes an online collection, natural history and scientific information related to the collection, and questions to prompt online exploration and investigation.

Listening Techniques

Interviews and Surveys

Interviews and surveys can illuminate visitors’ expectations and their reasons for visiting museums (Kaplan, 1993; McManus, 1992; Pekarik et al., 1999) as well as their reactions to and reflections on the museum experience. Trained staff and professional evaluators generally conduct interviews, which often are done on site. Interview protocols can include open-ended and objective questions. For example, the Smithsonian’s Institutional Studies Office asked exiting visitors, “What did you find most interesting? What surprised you? What was most informative?” The Ocean Planet exhibit was designed to reinforce attitudes and educate visitors. Ocean Planet sought to promote environmentally responsible behavior by providing information about ocean heroes, people, and products. The summative evaluation included personal interviews with entering and exiting visitors. In addition to soliciting demographic information, the interviewers asked visitors about the role of oceans, what their predictions were for the future of oceans, and what they saw as solutions for problems.

Diaries

Diaries are yet another data collection instrument used to elucidate visitors’ thoughts, attitudes, and behavior. As with interviews and surveys, diaries can be used to collect qualitative and quantitative information. One study (Leinhardt et al., 2000) examined the learning that takes place in museum settings by reviewing diaries based on participants’ visits to the museums of their choice. The researchers read and coded each diary entry. They noted the following information:

- Purpose of visit: Did visitors “float” through the exhibit? Did they have a clear focus? Were visitors “challenging” themselves in an unfamiliar situation?

- Response to the setting: How did visitors describe the environment? Responses included “neutral” (no mention), frame (described as the setting for the exhibit), “content” (museum described in context of exhibit), or “conflict” (disruptive, confusing)
- Main point of diary entry

The researchers acknowledged that the process of completing a diary affects a visitor’s experience because it encourages reflection. This study illustrates how researchers used the theoretical framework to guide their understanding of personal identity in museums.

Experimental Designs

Experimental designs assess outcomes between visitors and a control group. For example, Peart (1984) studied changes in visitors’ knowledge and attitudes based on exhibit types in the Living Land, Living Sea Gallery of the British Columbia (Canada) Provincial Museum. The study evaluated the effectiveness of five exhibit types: word (label only), picture and label, object without label, object and label and object, label and sound. A post-test-only experimental design was used (visitors to each of the five exhibits completed a post-trip questionnaire). The study assessed knowledge gain by comparing immediate results from those who had seen the exhibits with data collected from visitors before they entered the exhibit (the control group). The results of the post-exhibit questionnaires were compared with those from the questionnaires completed by the control group to determine the effect of viewing the exhibit. In addition to the questionnaire, the visitors were observed to determine exhibit attracting and holding power and visitor interaction with the exhibit. Peart found that knowledge gain resulted from all exhibits except the object-only exhibit, and the exhibit with object, label, and sound had the most significant (positive) outcome on knowledge gain.

Dierking and Falk (1994) recommend that the experimental design be replaced by more naturalistic studies to provide a clearer picture of the motivation and free choice inherent in visits to museums. They argue that, by constructing an experimental study, researchers might not be able to capture the affective learning and experience of visitors.

Exhibit Experiment: Does Exhibit Type Matter?

Exhibit type affects knowledge gain, attracting power, holding power, and interaction. Concrete exhibits—those that offer more visual cues and that attract different senses—are more likely to produce positive learning outcomes and attract visitors for longer periods (Peart, 1984).

A Memory Trip

In one of the few examples of an effort to identify the long-term effects of informal experiences, photographs were used as prompts to “investigate the memories” visitors had of their experiences at an interactive science center (Stevenson, 1991).

Participants demonstrated good factual recall: They gave clear descriptions of exhibits and were able to connect exhibit content to something in their lives. Specifically, Stevenson found that 99% of family members reported talking to one another or someone else about their science center experience, 27% had a spontaneous memory of trip, and 61% were able to remember something about the exhibit when shown a photograph.

Longitudinal Studies

Evaluators have sought to understand the retention or application of exhibit content and concepts beyond the immediate experience through follow-up interviews and questionnaires. Few long-term studies have been attempted, and most of them are limited to visitors' memories of their experience (Ramey-Gassert et al., 1994). The methods used in evaluation employ varying degrees of direct participant involvement. One inherent difficulty of longitudinal evaluation is access to visitors.

Research Findings

Three areas of research and evaluation have dominated museum studies:

- Visitor and exhibit characteristics
- Family behavior and family learning
- Field trips

The core museum findings are summarized in Table 3.2.

Table 3.2. Research Snapshot

Category	Key Findings
Visitor and exhibit characteristics	<ul style="list-style-type: none"> • Years of formal education influences visitor attendance • Knowledge, experience, and agenda affect visit outcomes • Exhibit labels and level of interactivity are significant in attracting visitors
Family behavior and learning	<ul style="list-style-type: none"> • Family behavior is fairly predictable • Fairly equal amounts of time are spent at each exhibit • Children interact more frequently with hands-on exhibits than do accompanying adults • Learning, defined as identifying, describing, or interpreting content, is taking place
Field trips	<ul style="list-style-type: none"> • Can increase students' content knowledge, conceptual understanding • Most effective when there is advance preparation, including novelty reduction • Provide professional development opportunities for classroom teachers, most significant in science education

How people change could have as much to do with their social values, personal attributes, and cultural norms as it does with the learning environment itself (Dierking, 1991; Doering, 1999; Falk and Dierking, 2000; Hein, 1998; Roberts, 1997; Schauble et al., 1997). The studies discussed below highlight visitors' demographic characteristics as well as research on outcomes that are based on visitor characteristics.

Visitor Characteristics

Visitor characteristics include age, agenda, attitudes, education, gender, and prior knowledge. The belief that each visitor has a unique experience within the same setting has prompted studies that evaluate outcomes based on visitor variables. Visitors' characteristics influence their experiences and affect what they will take away from the exhibits (Bitgood, 1992; Bitgood and Bishop, 1991; Falk and Dierking, 2000; Leinhardt et al., 2000; Rennie and McClafferty, 1996).

Population studies account for most of the research that involves visitor variables. For example, results taken from a Smithsonian Institution study (Doering, 1995) show attendance demographics for the Smithsonian Institutions. In a population of 185 million people over the age of 18, 65% (122 million people) made at least one visit to a museum or museum-like facility in 1994. Natural history museums were the least-visited and zoos and aquaria were the most-visited venues. Approximately one-third of the population (72 million people) visits two or more venues annually. The biggest influence on attendance appears to be education: People with more years of higher education visit more often (race and ethnicity are not significant when education is included as a visitor variable).

Agenda

Visitors have specific expectations and motives for attending exhibits, and their agendas go a long way toward determining the quality of the experience and the visit outcomes (Doering, 1999; Kaplan, 1993; Leinhardt et al., 2000; Miles and Tout, 1991; Pekarik et al., 1999).

Pekarik and colleagues (1999) developed a list of features that visitors say contribute most to a "satisfying experience" in museums. Satisfying trips generally fall into one of

Visitor Expectations

- Cognitive experiences involve gaining knowledge and understanding.
- Introspective experiences include feeling, imagining, reflecting, and recalling past experiences.
- Object experiences are based on visitors' interest in seeing objects, defined as the "real thing, seeing [real, uncommon, or valuable] things, being moved by beauty, thinking about what it would be like to own things, and continuing professional development." (Pekarik, 1999, online summary).
- Social experiences are based on spending time with other people.

four main categories: cognitive, introspective, object, or social. Other evaluators and researchers use different terms in describing agendas. In *Talking to Oneself: Diaries of Museum Visits*, Leinhardt and co-workers (2000) analyze diary entries and define visitor agendas in three distinct ways: "floating," or open to experience, with few preconceptions; "focused," or having a clear purpose, including the intention to learn something specific; or "challenging," with the purpose of expanding one's view. Roberts (1997) classifies agendas as "restorative," "social," for purposes of "reminiscence" or "fantasy," "personal," or "introspective."

According to Pekarik and colleagues, visitors will choose a museum with the expectation of meeting one of the four category experiences. The researchers took the work one step further to examine the relationship between the expectation for a specific experience within a specific type of museum. For example, visitors to craft museums seek object experiences proportionately more often than do visitors to other types of museums. In addition, the researchers drew connections between visitor characteristics (including age, gender, and experience, for example) and agendas. They reported that first-time visitors are more likely to choose object experiences and less likely to choose social experiences than are repeat visitors. This implies that visitor agendas influence the learning that takes place in museums.

Clearly, the purpose of a visit can strongly affect visitors' overall experience and the effectiveness of the exhibit. *Talking to Oneself* (Leinhardt et al., 2000) showed a correlation between age and purpose: Older participants tended to have more "focused" motives than did younger and middle-aged visitors; the youngest group more often had "challenging" agendas. However, purpose does not necessarily correlate with age. Rather, it is a factor related to the individual, the exhibit, and the visit. Leinhardt and colleagues (2000) and Pekarik and colleagues (1999) provide information that suggests that visitors do indeed have different agendas.

Does Visitors' Sex Play a Role?

Some studies report traditional gender stereotypes in visitor behavior: Mothers are submissive, rarely choosing exhibits or initiating movement (Diamond, 1986; McManus, 1992); fathers act as teachers (Koran et al., 1988).

Some studies show girls receive less attention (Crowley and Callanan, 1998); others suggest that girls are favored (Blud, 1990).

Several researchers caution that those findings could have more to do with the exhibit type than with any real gender issue (Dierking and Falk, 1994; McManus, 1992; Rennie and McClafferty, 1996). Still others argue that conclusions are an artifact of sampling.

Gender

Much of the information in the published literature on gender comes from family research. The message about gender is mixed. There is a concern that a “snapshot” at one exhibit is not reflective of the family or of interpersonal interactions throughout a center (McManus, 1992).

Knowledge and Attitudes

Museums serve as grounds to reinforce existing beliefs (Falk and Dierking, 2000) as well as to provide information that visitors can use to form opinions, have conversations, increase their knowledge, and—potentially—change their behavior. Prior knowledge influences visitor experiences. In a study of knowledge about *and* attitudes toward animals at the Birmingham (Alabama) Zoo, Bitgood (1992) reports that visitors *entering* the zoo exhibited a positive correlation between self-rated knowledge and the belief that animals are “worth saving” and “attractive.” A high level of knowledge corresponded to positive attitudes. Remarkably, upon *exiting* the zoo “high-knowledge” individuals reported lower ratings for “worth saving” and “attractiveness” than did “low-knowledge” individuals. One conclusion from the data is that zoo visits might have the most positive impact on people who are less knowledgeable about animals (Bitgood, 1992, p. 10). The less knowledgeable visitors appeared to benefit most from the visit in terms of positive attitude change since they increased their favorable ratings as a result of the visit. This is encouraging because there is always a danger of “preaching to the choir.”

Doering (1999) reported that museum exhibits confirm and reinforce existing attitudes, indicating that there may be little or no attitude change as a result of a visit. Doering (1999, p. 80) cited a lack of knowledge gain:

[P]eople tend to frequent the museums and exhibitions that they think will be congruent with their own attitudes. ... They respond best to exhibitions and themes that are personally relevant and with which they can easily connect. Consequently, we found that most museum visitors acquire little new factual knowledge.

What Makes a Good Exhibit?

Interactive Components

- *Improve holding power*
- *Enhance social interactions*
- *Increase knowledge gain*
- *Lead to more questioning*

Labels

- *Textual references improve holding power*

Exhibit Characteristics

There has been extensive research on the relationship between exhibit characteristics and visitor experiences (Bickford, 1993; Bielick and Doering, 1997; Blud, 1990; Doering et al., 1993; Eratuuli and Sneider, 1990; Hein, 1998; Hilke and Balling, 1985; Koran et al., 1984; Patterson and Bitgood, 1988; Peart, 1984; Stevenson, 1991; Tunnicliffe, 1996). Some characteristics contribute more than others to visitor learning. Hands-on exhibits are more likely to attract and hold visitors' attention (Bitgood, 1988; Eratuuli and Sneider, 1990; Koran et al., 1984; Peart, 1984), enhance social interactions (Bitgood, 1988; Eratuuli and Sneider, 1990) increase knowledge (Peart, 1984; Tunnicliffe, 1996), and lead to more questioning and less "explanatory" or telling behavior (Hilke and Balling, 1985). Interactive exhibits are more influential in reinforcing desirable attitudes and altering unwanted perspectives than are static exhibits (Finlay et al., 1988; Fronville and Doering, 1989; Peart, 1984; Swanagan, 2000; Tuckey, 1992). Behavioral psychology supports the finding that active participation heightens learning.

Swanagan (2000) compared an interactive elephant "show" that included a fact-based program with one that involved viewing the animals only. There was a greater effect on conservation knowledge and attitudes from participation in the interactive exhibit than resulted from simply viewing the animals. The more effective intervention combined three important elements: interaction, information, and opportunity to express intention. Participants expressed their behavioral intentions or attitudes by filling out solicitation cards, which asked them to explain their feeling about the ivory trade, and visitors were encouraged to support U.S. sanctions on ivory imports.

Perhaps is not surprising that the content and method of presentation on exhibit labels influences the effectiveness of a museum experience. A simple but important finding is that making explicit references to the exhibit, within the label text, improves holding power (Serrell, 1981, 1996). Jenni Martin, education director at the Children's Discovery Museum in San Jose, California, explains that the inclusion of a "Power Girl" icon on labels for an energy exhibit remedied gender inequities that were previously noted in parents' responses to their children (Schneider, personal communication, 2001).

The Monterey Bay (California) Aquarium reported that when humor, questions, and

second voice (messages that address the visitor as “you”) were incorporated in exhibit labels, visitors read aloud more often, pointed more often to objects within an exhibit, and were more likely to share relevant information (Rand, 1985). Overall, labels appeared to affect visitor behavior and result in better comprehension.

Family Behavior and Learning

Families represent approximately 60% of museum visitors (Dierking and Falk, 1994) and have been an important focus of museum research (Diamond, 1986; Dierking and Falk, 1994; Hein, 1998; McManus, 1989, 1992; Rennie and McClafferty, 1996). McManus (1992), Rennie and McClafferty (1996), and Dierking and Falk (1994) provide thoughtful reviews of the most important and most frequently cited research conducted on family behaviors and learning. Families are generally defined as groups containing at least one adult and one child (but not exceeding more than four adults or five children). Two broad themes direct studies:

- Family behavior is examined in terms of group interaction, movement, and inferences about how these activities relate to the notion of a family agenda.
- Family learning has been investigated with respect to child–adult interaction and family interactions with the exhibit.

Family Behavior

Time allocation and tracking studies have been a principal source for what we know about visitors’ behaviors within informal settings. As the phrase suggests, timing and tracking studies provide data about the extent to which, and how, and where families spend their time. The assumption underscoring this work is that time spent is a reliable indicator of learning.

There is evidence that families do exhibit predictable patterns in where and how they spend their time during a museum visit (Diamond, 1986; Falk, 1991; Hilke and Balling, 1985; Lakota, 1975, in Dierking and Falk, 1994; Taylor, 1996, in Dierking and Falk, 1994). Related studies at various venues, including museums and an aquarium, describe four distinct phases in family behavior patterns (Table 3.3).

Researchers claim that the exhibit-viewing and cruising phases reflect the presence of a family agenda (Diamond, 1986; Falk, 1991). Tracking studies also show little variation in the time family groups spend from one exhibit to the next. Family groups spend an equal amount of time at each exhibit (Lakota, 1975, in Dierking and Falk, 1994), and children within family groups physically interact with hands-on exhibits significantly more often than do accompanying adults (Dierking and Falk, 1994).

Table 3.3. Family Behavior in Museums

Orientation	Viewing	Cruising	Departure
3–10 minutes	25–30 minutes	30–40 minutes	5–10 minutes
Begin viewing, become familiar with surroundings	Intense viewing, characterized by reading labels, interacting with exhibit	Scan exhibits	Prepare for departure

Learning in Family Groups

Another major focus of family research examines the extent to which families learn or exhibit behaviors that are likely to lead to learning (learning behaviors include asking and answering questions, commenting and explaining on exhibit content, and reading labels). Family members' behaviors and conversations while interacting with exhibits and with one another can suggest when learning is taking place (Falk and Dierking, 2000; Hein, 1998; Ramey-Gassert et al., 1994; Rennie and McClafferty, 1996). Borun and colleagues (1996) identified three levels of learning in their study of four science centers: identifying, describing and interpreting, and applying exhibit content.

The content of family exchanges typically centers on the factual information being conveyed by the exhibit. Adults and children identify and describe what they see. For example, in her study of family visitors to the Lawrence Hall of Science, Diamond (1986) likened the interactions among family members to that of classroom “show and tell.” It is common for family members to acquire and exchange information spontaneously and out loud, so all family members might have access to the information (Hilke, 1989; Hilke and Balling, 1985). There is an inherently social component of learning in informal settings (Hilke, 1989, Hilke and Balling, 1985; Semper, 1990).

In some cases, learning is manifest as teaching behaviors between family members. Teaching behaviors, including pointing, showing, and explaining, are typical of family behavior (Crowley and Callanan, 1998; Diamond, 1980). Crowley and colleagues (2001a, abstract, p. 712) summarized their findings from a study of child–parent interactions at the Children’s Discovery Museum in San Jose, California:

When children engaged an exhibit with parents, their exploration of evidence was observed to be longer, broader, and more focused on relevant comparisons than children who engaged the exhibit without their parents. ... Parents also sometimes assumed the role of explainer by casting children’s experience in causal terms, connecting the experience to prior knowledge, or introducing abstract principles.

These teaching interactions occur more often between parents and children than among siblings (Crowley and Callanan, 1998; Diamond, 1986; Hilke, 1989; McManus, 1992; Rennie and McClafferty, 1996). Individual characteristics, such as age and prior knowledge, as well as specific exhibit characteristics and social interaction play a role in the learning that occurs during family visits to science centers.

Field Trips

Evaluations of school trips to museums conclude that field trips can support both factual and conceptual science understanding as well as affective objectives (Bailey, 2002; Prather, 1989; Price and Hein, 1991). Orientation to setting, clear learning objectives, and follow-up activities are factors that have proven most influential in successful field trips (Bitgood, 1989; Prather, 1989; Price and Hein, 1991; Ramey-Gassert et al., 1994).

Preparing students for the environment they will encounter appears to be particularly important. Pretrip orientations that address students’ personal agendas for the experience have a significant effect on their learning gains (Falk and Dierking, 1992). Although some novelty can stimulate exploration and motivate learning through curiosity (Martin et al., 1981), too much can be distracting and can impede factual and conceptual understanding (Falk and Balling, 1978; Kubota and Olstad, 1991). Likewise, multiple field trip experiences seem to reduce the novelty of the setting and enhance opportunities for learning (Griffin, 1998; Price and Hein, 1991).

Teachers' agendas (intentions and perceptions of the field trip experience) also impact the overall effectiveness of the visit because teachers' agendas influence student perceptions of the visit (Griffin and Symington, 1998; Schneider, 2002). In their meta-analysis, Dierking and Falk (1994) note a correlation between prior knowledge and learning (Shettel et al., 1968, in Dierking and Falk, 1994). People with more science knowledge learn more than do visitors with less prior knowledge. This finding suggests that teachers' framing and preparation for field trips can significantly increase the learning opportunity for students. Field trips linked to the school curriculum result in higher learning gains for students (Anderson, 1999, in Falk and Dierking, 2000; Griffin and Symington, 1997; Price and Hein, 1991). Findings such as these, however, should not be interpreted to mean that field trips are inherently effective (Prather, 1989), because field trips differ.

In addition to serving as educational experiences for students, visits to museums are professional development opportunities for classroom teachers. It is estimated that close to one-third and maybe as much as one-half of professional development for science teachers occurs in informal centers (Bartels, 2001). Furthermore, teachers who visit informal sites are more enthusiastic about and teach more science (Price and Hein, 1991), signifying possible changes in teachers' knowledge, attitude, and behavior as an outcome.

Chapter Summary

Theories and Methods

Much of what is known about museum experiences has been influenced by a view of learning that is determined by the degree to which the intended curriculum is understood. This view of learning has expanded to include a sociocultural approach. In recent years, researchers have sought to define learning as more than just acquiring content knowledge to include some measure of conceptual and affective change. Contextual models are examples of theoretical frameworks that account for multiple factors, including visitor variables, exhibit characteristics, and interaction among visitors and between visitors and the learning environment.

Key Findings

Learning goes on in museums all of the time, and the extent to which visitors' knowledge, attitudes, and behavior are affected by museums depends on many variables that are related both to the individual visitors and to the setting.

Three areas of research and evaluation have dominated museum studies: visitor and exhibit characteristics, family behavior and learning, and field trips. The belief that each visitor has a unique experience within the same setting has prompted studies that evaluate outcomes based on visitor variables.

- Education appears to have the most influential impact on attendance: Individuals with more years of formal education visit museums more often than people do with less formal education.
- Textual references (exhibit labels) increase knowledge gain and holding power.
- Interactive exhibits increase holding power, enhance social interactions, lead to increases in knowledge, and promote more questioning.
- Family behavior can be divided into four distinct phases: orientation, preparation, viewing, and cruising.
- Field trips are most effective when there is advance preparation and follow-up.
- Field trips can increase students' content knowledge and conceptual understanding.
- Field trips provide professional development opportunities for classroom teachers.

Summary of Studies

Visitor and Exhibit Characteristics

This section summarizes museum studies for three research topics (visitor and exhibit characteristics, family behavior and learning in family groups, and field trips). Table 3.4 shows the results of individual studies that have examined visitors' experiences and exhibit outcomes. The table is sorted by the focus of the study, or the core goal of the intervention. Differences in visitors'

prior knowledge, agendas, age, and education affect what and how they experience informal learning environments. Furthermore, evidence suggests that visitors' personal attributes and cultural norms have as much to do as the learning environment itself in determining visit outcomes.

Table 3.4. Summary of Studies: Visitor and Exhibit Characteristics

Focus of Study	Effect	Venue	Intervention	Population	Method	Follow-Up	Reference
Knowledge, agenda	+ ^a	Museum	Museum visits of participants' choice	Adults	Diaries	No	Leinhardt et al., 2000
Attitudes, Knowledge	+ -	3 Museums	Formative study, oceans exhibit	Adults	Surveys	Yes	Bickford, 1993
Attitudes	+	Museum	Ocean exhibit	General visitor	Interviews, observations	No	Bickford et al., 1996
Knowledge, attitude	+	Museum	Smithsonian maps exhibit	General public	Experiment; pre-, post-visit survey	No	Doering, et al., 1993
Behavior: attention and curiosity (attraction and holding power)	+	Museum	Hands-on exhibit versus view only	General public	Experiment, observation	No	Koran et al., 1984
Exhibit characteristics Knowledge Attitude Behavior	+ Minor Mixed	Museum	5 exhibits: label only; picture and label; object; objects and label; object, label, sound	General public	Experiment, observation, questionnaire	No	Peart, 1984
Time allocation: behavior, learning	+ ^b	Museum	Interactive exhibit	School-aged children	Pre-, post-visit observation	No	Falk, 1983, in Dierking and Falk, 1994
Attitudes	+	Science center	Exhibit type, interactive exhibits	Elementary school children	Observation, interview	No	Tuckey, 1992
Behavior, knowledge	+	Science center	Visit to science center with and without descriptive orientation preceding visit	Grade 6	Experiment, observation, post-visit test	Yes	Kubota and Olstad, 1991

Table 3.4. Summary of Studies: Visitor and Exhibit Characteristics *Continued*

Focus of Study	Effect	Venue	Intervention	Population	Method	Follow-Up	Reference
Attitude toward science, technology, society (STS)	Mixed ^c	Science center	Visit to science center	Grades 6–8	Pre-, post-intervention test, pre-study questionnaire to determine prior classroom use of STS	No	Finson, 1985
Attitudes, behavior toward conservation policies	+	Zoo	Elephant exhibit	General public	Survey, petition	No	Swanagan, 2000
Attitude correlation to animal environment	+, Mixed	Zoo	Slide show of different zoo environments compared with no slides	Undergraduate students	Experiment, semantic differential scales	No	Finlay et al., 1988
Attitudes, social interaction	Mixed ^d	Zoo	Exhibit type	Adults	Survey, questionnaire, observation	No	Bielick and Doering, 1997
Attitudes, visitor characteristics, self-reports of knowledge	Mixed	Zoo	Zoo visit	General public	Pre- and post-visit surveys	No	Bitgood, 1992
Content of conversation by exhibit type and visitor age	Mixed	Zoo	Exhibit type: live animals, animated dinosaurs, preserve animals	Elementary school children	Observations, recorded conversations	No	Tunncliffe, 1996

Table key: *Focus of Study*, intended effect of the exhibit or focus (topic) of the study; *Effect*, positive (+), negative (-), or mixed study findings; *Venue*, intervention site; *Intervention*, type of exhibit, program, or activity; *Population*, participants in the study; *Method*, data collection technique; *Follow-Up*, if and when data were collected after the participant visit; *Reference*, author and date of study.

^a Outcomes based on visitors' knowledge, reflection, and agenda.

^b High correlation between attention to exhibit and learning.

^c Differences according to grade level, prior experience, teacher preparation.

^d Although the exhibit did not change visitors' attitudes toward science or the belief that "animals think," viewing the exhibits did affect the way some visitors think about animals. College-educated visitors and visitors with children were most likely to report changes in their perception of animals based on the exhibit.

Family Behavior and Learning

The studies listed in Table 3.5 investigated the role of parent–child variables or exhibit characteristics. Dierking and Falk (1994, pp. 64, 67) summarize the findings in their essay on families in informal science settings:

[T]here is considerable evidence to indicating that families are adopting “learning agendas.” ... [F]amilies utilize informal science settings to facilitate learning, but historically the effort to “structure” and to “measure” the family museum experience in a narrow quantitative fashion has prevented researchers from truly understanding the nature of this learning.

Table 3.5. Summary of Studies: Family Behavior and Learning

Focus of Study	Effect	Venue	Intervention	Method	Follow-Up	Reference
Knowledge, learning behavior	+	Museum	Interactive science exhibits	Observation	No	Diamond, 1986
Learning behavior by age, gender	Mixed ^a	Museum	Exhibit type (static to walk-through to interactive)	Observation	No	Dierking, 1987, in Dierking and Falk, 1994
Gender; parent–child interaction	Mixed	Science center	18 Interactive science exhibits	Observation	No	Crowley et al., 2001a
Learning behaviors (social interaction)	+	Science center	Exhibit types, labels	Observations	No	Eratuuli and Sneider, 1990
Learning behavior	+	Science center	4 Science museums	Observation, post-visit interviews	No	Borun et al., 1996
Knowledge	+	Science center	Interactive science exhibits	Observation, interview	6 Months	Stevenson, 1991
Knowledge (memories, influence of exhibit)	+	Science center	Exhibit at interactive center	Observation, exit interviews	6 Months (interviews)	McManus, 1992

Table key: *Focus of Study*, intended effect of the exhibit or focus (topic) of the study; *Effect*, positive (+), negative (-), or mixed study findings; *Venue*, intervention site; *Intervention*, type of exhibit, program, or activity; *Method*, data collection technique; *Follow-Up*, if and when data were collected after the visit; *Reference*, author and date of study.

^a Frequency of learning behaviors dependent on age of children, sex of parents and children.

Field Trips

Studies of school visits to museums conclude that field trips support both factual and conceptual science understanding as well as affective objectives.

Table 3.6. Summary of Studies: Field Trips

Focus of Study	Effect	Venue	Intervention	Population	Method	Follow-Up	Reference
Novelty, learning behavior	Mixed ^a	Museum	Field trip	Grade 6 public school students	Observation, post-visit test, control group	No	Kubota and Olstad, 1991
Learning attitudes	+	Museum	Field trips	Various	Literature review	NA	Ramey-Gassert et al., 1994
Knowledge, curiosity	+	Museum	Field trip	Elementary, middle school students	Observation	No	Gottfried, 1981

Table key: *Focus of Study*, intended effect of the exhibit or focus (topic) of the study; *Effect*, positive (+), negative (-), or mixed study findings; *Venue*, intervention site; *Intervention*, type of exhibit, program, or activity; *Population*, participants in the study; *Method*, data collection technique; *Follow-Up*, if and when data were collected after the participant visit; *Reference*, author and date of study; NA, not applicable.

^a Exploratory behavior positively correlated with learning, novelty reduction highly effective for boys.

References

- Bailey, Elsa. 2002. Review of selected references from literature search on field trips/school group visits to museums. Association of Science–Technology Centers, ASTC Resource Center, Education. Online, accessed Aug. 24, 2002: <http://www.astc.org/resource/educator/ftrips.htm>.
- Bartels, D. M. 2001. Science centers and the future of educational reform. *Ties* 1,9.
- Bickford, A. 1993. Visitors and ocean issues: A background study for the National Museum of Natural History Ocean Planet exhibition. Institutional Studies Office Report 93-7. Washington, D.C.: Smithsonian Institution.
- Bickford, A., A. J. Pekarik, Z. D. Doering, and S. S. Yalowitz. 1996. Ocean views: A study of visitors to the Ocean Planet exhibition at the National Museum of Natural History. Institutional Studies Office Report 96-5. Washington, D.C.: Smithsonian Institution.
- Bielick, S., and Z. D. Doering. 1997. An assessment of the “Think Tank”: Exhibition at the National Zoological Park. Institutional Studies Office Report 97-1. Washington, D.C.: Smithsonian Institution. Online summary, accessed Aug. 24, 2002: <http://www.si.edu/iso/97-1.htm>.
- Bitgood, S. 1989. School field trips: An overview. *Visitor Behavior* IV(2):3–6.
- Bitgood, S. 1992. The impact of a zoo visit on attitudes: A preliminary report on interaction effects. *Visitor Behavior* 7(3):7–10.
- Bitgood, S., and S. Bishop. 1991. The role of a current visit, prior visits, and gender on visitor perception of the Natural History Museum. *ILVS Review: A Journal of Visitor Behavior* 2(1):49–65.
- Blud, L. M. 1990. Sons and daughters: Observations on the way families interact during a museum visit. *International Journal of Museum Management and Curatorship* 9(3):257–264.
- Borun, M., M. Chambers, and A. Cleghorn. 1996. Families are learning in science museums. *Curator* 39(2):124–138.
- Cennamo, K. S., and S. C. Eriksson. 2001. Supporting scientific inquiry through museum websites. *Educational Technology* (May–June):50–55.
- Crowley, K., and M. Callanan. 1998. Identifying and supporting shared scientific reasoning in parent–child interactions. *Journal of Museum Education* 23:12–17.

- Crowley, K., M. A. Callanan, J. L. Jipson, J. Galco, K. Topping, and J. Shrager. 2001a. Shared scientific thinking in everyday parent-child activity. *Science Education* 85(6):712-732.
- Crowley, K., G. Leinhardt, and C. Chang. 2001b. Emerging research communities and the World Wide Web: Analysis of a Web-based resource for the field of museum learning. *Computers and Education* 36:1-14.
- Diamond, J. 1980. The ethnology of teaching: A perspective from the observation of families in science centers. Unpublished doctoral dissertation, University of California, Berkeley.
- Diamond, J. 1986. The behavior of family groups in museums. *Curator* 29(2):27-33.
- Dierking, L. D. 1987. Parent-child interactions in a free choice learning setting: An examination of animation of attention-directing behaviors. Doctoral dissertation, University of Florida.
- Dierking, L. D. 1991. Learning theory and learning styles: An overview. *Journal of Museum Education* 16(1)4-6.
- Dierking, L., and J. Falk. 1994. Family behavior and learning in informal science settings: A review of the research. *Science Education* 78(1):57-71.
- Doering, Z. 1995. Who attends our cultural institutions? Institutional Studies Office Research Note 95-5. Washington, D.C.: Smithsonian Institution.
- Doering, Z. 1999. Strangers, guests or clients? Visitor experiences in museums. *Curator* 42(2):74-87.
- Doering, Z. D., A. E. Kindlon, and A. Bickford. 1993. The power of maps: A study of an exhibition at the Cooper Hewitt National Museum of Design. Institutional Studies Office Report 93-5. Washington, D.C.: Smithsonian Institution.
- Eratuuli, M., and C. Sneider. 1990. The experiences of visitors in a physics discovery room. *Science Education* 74(4):481-493.
- Falk, J. 1991. Analysis of the behavior of family visitors in natural history museums: The National Museum of Natural History, Washington, D.C. *Curator* 34(1):44-50.
- Falk, J. M., and J. Balling. 1978. The novel field trip phenomenon: Adjustment to novel settings interferes with task learning. *Journal of Research in Science Teaching* 15(2):127-134.
- Falk, J. and L. D. Dierking. 1992. The museum experience. Washington, D.C.: Whalesback Books.
- Falk, J., and L. Dierking. 2000. Learning from museums. Walnut Creek, Calif.: Alta Mira.

- Finlay, T., L. James, and T. Maple. 1988. People's perceptions of animals: The influence of zoo environment. *Environment and Behavior* 20(4):508–528.
- Finson, K. O. 1985. A study of student attitudes toward science–technology–society resulting from visitation to a science–technology museum. Unpublished doctoral dissertation, Kansas State University, Manhattan, Kan.
- Fronville, C. L., and Z. D. Doering. 1989. Visitor perspectives on tropical rainforests: A report based on the 1988 Tropical Rainforests: A Disappearing Treasure. Institutional Studies Office. Washington, D.C.: Smithsonian Institution.
- Gottfried, J. L. 1981. Do children learn on school field trips? *Curator* 23(3):771–781.
- Griffin, J. 1998. School–museum integrated learning experiences in science: A learning journey. Unpublished doctoral dissertation, University of Technology, Sydney, Australia.
- Griffin, J., and D. Symington. 1997. Moving from task-oriented learning strategies on school excursions to museums. *Science Education* 81(6):763–780.
- Hein, G. 1997. The maze and the Web: Implications of constructivist theory for visitor studies. Online, accessed Jan. 5, 2003: http://www.lesley.edu/faculty/ghein/papers_online/mazeweb.html.
- Hein, G. 1998. Learning in the museum. London: Routledge.
- Hilke, D. D. 1989. The family as a learning system: An observational study of families in museums, ed. B. H. Butler and M. B. Sussman. New York: Haworth.
- Hilke, D. D., and J. D. Balling. 1985. The family as a learning system: An observational study of families in museums. Unpublished manuscript, Smithsonian Institution, Washington D.C.
- Kaplan, S. 1993. The museum as a restorative experience. *Environment and Behavior* 25(6):725–742.
- Koran, J., L. Morrison, J. R. Lehman, M. Koran, and L. Gandara. 1984. Attention and curiosity in museums. *Journal of Research in Science Teaching* 21(4):357–363.
- Koran, J. J. Jr., M. L. Koran, L. D. Dierking, and J. Foster. 1988. Using modeling to direct attention in a natural history museum. *Curator* 31(1):36–42.
- Kubota, C. A., and R. G. Olstad. 1991. Effects of novelty-reducing preparation on exploratory behavior and cognitive learning in a science museum setting. *Journal of Research in Science Teaching* 23(3):225–234.

- Leinhardt, G., and K. Crowley. 1998. Museum learning as conversational elaboration: A proposal to capture, code and analyze museum talk. Museum Learning Collaborative Technical Report MLC-01, Learning Research and Development Center. Pittsburgh, Pa.: University of Pittsburgh. Online, accessed Aug. 24, 2002: <http://museumlearning.com/paperresearch.html>.
- Leinhardt, G., C. Tittle, and K. Knutson. 2000. Talking to oneself: Diary studies of museum visits. Museum Learning Collaborative Technical Report MLC-04, Learning Research and Development Center. Pittsburgh, Pa.: University of Pittsburgh. Online, accessed Aug. 24, 2002: <http://museumlearning.com/paperresearch.html>.
- Lucas, A. M. 1991. "Info-tainment" and informal sources for learning science. *International Journal of Science Education* 13(5):495–504.
- Martin, W. W., J. H. Falk, and J. D. Balling. 1981. Environmental effects on learning: The outdoor field trip. *Science Education* 65(3):301–309.
- McManus, P. 1989. Oh yes they do! How visitors read labels and interact with exhibit texts. *Curator* 32(2):174–189.
- McManus, P. 1992. Topics in museums and science education. *Studies in Science Education* 20:157–182.
- Miles, R., and A. Tout. 1991. Impact of research on the approach to the visiting public at the Natural History Museum, London. *International Journal of Science Education* 13(5):543–549.
- Patterson, D., and S. Bitgood. 1988. Some evolving principles of visitor behavior. In *Visitor studies—1988: Theory, research, and practice*, Vol. 1, 40–50, ed. S. Bitgood, J. Roper, and A. Benefield. Jacksonville, Ala.: Center for Social Design.
- Peart, B. 1984. Impact of exhibit type on knowledge gain, attitudes, and behavior. *Curator* 27(3):22–237.
- Pekarik, A., Z. Doering, and D. Karns. 1999. Exploring satisfying experiences in museums. *Curator* 42(2):117–129. Online summary, accessed Jan. 5, 2003: http://museumlearning.com/scripts/search_display.php?Ref_ID=2228.
- Prather, J. P. 1989. Review of the value of field trips in science instruction. *Journal of Elementary Science Education* 1(1):10–17.
- Price, S., and G. Hein. 1991. More than a field trip: Science programmes for elementary school groups at museums. *International Journal of Science Education* 13(5):505–519.

- Ramey-Gassert L., H. J. Walberg III, and H. J. Walberg. 1994. Museums as science learning environments: Reexamining connections. *Science Education* 78(40):345–363.
- Rand, J. 1985. Fish stories that hook readers: Interpretive graphics at the Monterey Bay Aquarium. In 1985 conference proceedings. Columbus, Ohio: American Association of Zoological Parks and Aquaria.
- Rennie, L. J., and T. P. McClafferty. 1996. Science centres and science learning. *Studies in Science Education* 27:53–98.
- Roberts, L. C. 1997. From knowledge to narrative: Educators and the changing museum. Washington, D.C.: Smithsonian Institution Press.
- Swanagan, J. 2000. Factors influencing zoo visitors' conservation attitudes and behavior. *Journal of Environmental Education* 31(4):26–31
- Schauble, L., G. Leinhardt, and L. Martin. 1997. A framework for organizing cumulative research agendas in informal learning contexts. *Journal of Museum Education* 22(2,3):3–8.
- Schneider, B. 2002. Making the informal formal: An examination of why and how teachers and their students leverage experiences to informal learning environments. Unpublished doctoral dissertation, Stanford University.
- Semper, R. J. 1990. Science museums as environments for learning. *Physics Today* (November):50–56.
- Serrell, B. 1981. Zoo label study at Brookfield Zoo. *International Zoo Yearbook* 21:54–61.
- Serrell, B. 1993. Using behavior to define the effectiveness of exhibitions. In Museum visitor studies in the '90s: Conference proceedings, 140–144., ed. S. Bicknell and G. Farmelo. London: Science Museum.
- Serrell, B. 1996. Exhibit labels: An interpretive approach. Walnut Creek, Calif.: Alta Mira.
- Sheppard, B., ed. 1999. Perspectives on outcome based evaluation for libraries and museums. Washington, D.C.: Institute of Museum and Library Services. Online, accessed Aug. 24, 2002: <http://www.ims.gov/pubs/pdf/pubobe.pdf>.
- Shettel, H. H., M. Butcher, T. S. Cotton, J. Northrup, and D. C. Slough. 1968. Strategies for determining exhibit effectiveness. AIR E95-4-68-FR. Washington, D.C.: American Institutes for Research.
- Stevenson, J. 1991. The long-term impact of interactive exhibits. *International Journal of Science Education* 13(5):521–531.
- Taylor, S. 1986. Family behavior at the Steinhart Aquarium. Unpublished doctoral dissertation, University of California, Berkeley.

- Tuckey, C. 1992. School children's reactions to an interactive science center. *Curator* 35(1):28–38.
- Tunncliffe, S. D. 1996. The relationship between pupil's age and the content of conversations generated at three types of animal exhibits. *Research in Science Teaching* 26(4):461–480.
- Weil, S. 1999. Transformed from a cemetery of bric-a-brac. In *Perspectives on outcome based evaluation for libraries and museums*. Washington, D.C.: Institute of Museum and Library Services. Online, accessed Aug. 24, 2002: <http://www.ims.gov/pubs/pdf/pubobe.pdf>.

Resources

Online Resources

Association of Science–Technology Centers <http://www.astc.org/resource/index.htm>

The ASTC Resource Center contains articles and reports in the museum field. The Visitor Studies section highlights the role of visitor studies and provides excerpts from front-end studies, evaluation guides, and other online resources.

Museum Learning Collaborative <http://museumlearning.com>

The MLC web site includes a comprehensive annotated bibliography, research studies, and syllabi from university classes about informal learning and learning in museums. The work is a combined effort among museums and a core group of researchers.

Other Resources

- Bickford, A., and M. M. Riddick. 1996. Little scholars: A study of students at the Smithsonian early enrichment center. Institutional Studies Office Report 96-4. Washington, D.C.: Smithsonian Institution.
- Bickford, A., Z. D. Doering, and S. J. Smith. 1992. The spiders are coming! An exhibition background study for the National Museum of Natural History. Institutional Studies Office Report 92-4. Washington, D.C.: Smithsonian Institution.
- Bitgood, S., and H. H. Shettel. 1996. An overview of visitor studies. *Journal of Museum Education* 21(3):6–10.

- Bitgood, S., and R. J. Loomis, eds. 1993. Environmental design and evaluation in museums: Special issue. *Environment and Behavior* 25(6):683–820.
- Crowley, K., M. A. Callanan, H. R. Tenenbaum, and E. Allen. 2001. Parents explain more **often to boys than to girls during shared scientific thinking**. *Psychological Science* 12:3.
- Falk, J., and L. Dierking. 1995. Public institutions for personal learning: Establishing a research agenda. Washington D.C.: American Association of Museums.
- Leinhardt, G. & Crowley, K. 2002. Objects of learning, objects of talk: Changing minds in museums. In *Perspectives on object-centered learning in museums*, ed S. Paris. Mahwah, N.J.: Lawrence Erlbaum Associates.
- McManus, P. 1993. Memories as indicators of the impact of museum visits. *Museum Management and Curatorship* 12(4):367–380.
- Meyers, G., and C. Saunders, C. In press. Animals as links toward developing caring relationships with the natural world.
- Museums for a new century: a report of the commission on museums for a new century. 1984. *ERIC* 256:647.
- Petty, R. E., and J. T. Cacioppo. 1986. Communication and persuasion: Central and peripheral routes to attitude change. New York: Springer-Verlag.
- Ramey-Gassert, L. 1997. Learning science beyond the classroom. *Elementary School Journal* 57(4):433–450.
- Screven, C. G. 1984. Educational evaluation and research in museums and public exhibits: A bibliography. *Curator* 27(2):147–165.
- Smith, S. J. 1999. Bibliography of sources. *Curator* 42(2):174–179.
- St John, M., and D. Perry. 1993. A framework for evaluation and research: Science infrastructure and relationships. In *Museum visitor studies in the '90s: Conference proceedings*. ed. S. Bicknell and G. Farmelo. London: Science Museum.
- Taylor, S., and B. Serrell. 1991. Try it! Improving exhibits through formative evaluation. Washington D.C.: Association of Science–Technology Centers; New York: New York Hall of Science.
- Yalowitz, S., and M. Wells. 2000. Mixed methods in visitor studies research. *Journal of Interpretation Research*, 5 (1):45–52.

Chapter 4

Health Programs

In This Chapter

Theoretical Frameworks

- Health belief model
- Social learning theory
- CDC framework

Methodology

- Case studies
- Experiments and field studies
- Listening techniques

Research Findings

- Program effectiveness
- Changing behavior through attitudes and modeling

Chapter Summary

- Theories and methods
- Key findings
- Summary of studies

References

Resources

Health programs—such as family-planning interventions, interventions that encourage sexual abstinence or contraceptive use, anti-smoking campaigns, and programs that promote healthy living—are based on informing and changing the behavior of participants. Interventions are as varied as the targeted outcomes and participants.

Documentation of success in youth behavior change (or prevention programs) dates back to the 1970s (Perry, 1999), although it could be argued that the lack of more historical results could be attributed to the relatively small number of valid assessments conducted prior to the 1970s. For example, there were several hundred teenage-pregnancy-prevention programs in the 1990s yet “few good evaluations existed” (Card, 1999). Few evaluations of health programs in general have been published in the peer-reviewed literature (Perry, 1999), and the few published evaluations show a small number that have noticeably influenced health behavior among young people (Perry, 1999). Today, it is standard to dedicate 10% of program resources to evaluation (CDC, 1999b).

This chapter reviews a small selection of articles from health literature, primarily related to youth programming and family-planning practices, to highlight the program strategies, assessment tools, and evaluation results in the health field. Successful interventions are as varied as their targeted outcomes and participants. Youth programs that consider multiple components (school environment, family, and the community, for example) appear to be most effective. The most successful programs tend to incorporate multiple educational and informational techniques. The studies noted are examples and do not represent a comprehensive review of health literature (there was no review of medical journals, for example). As a group, however, the sample illustrates effective behavior change techniques that are applicable to environmental education and other fields.

The way the problem is framed is crucial because such framing determines the possible solutions. It determines “what works” to solve the problem.
(Card, 1999, p. 264).

Theoretical Frameworks

The health literature shows strong program theory for bringing about change. The health belief model and social learning theory are fundamental to many interventions, and the theory of reasoned action and planned behavior, described in Chapter 2, provides direction for program and evaluation structure in the health field. In addition, agencies such as the Centers for Disease Control (CDC) have encouraged assessment guidelines based on traditional evaluation metrics such as logic models.

Health Belief Model

The health belief model asserts that people need to understand the value of changing their behavior in order to change. It is a product of behavioral-science research that accounts for behavior based on individual knowledge and beliefs (Salazar, 1991): One must believe negative consequences will result from *not* changing behavior in order to be motivated to act. For example, people are more likely to quit a destructive behavior if they are knowledgeable about the negative impact of the behavior and believe that they are subject to the negative consequences.

The health belief model is manifested in several ways:

- Behavioral skills development
- Case studies
- Fact-based education
- Goal setting
- Role-playing
- Video presentations

In a review of behavior theories, Salazar (1991) notes the model's potential shortcomings: It is based on the belief that people value health and act accordingly to preserve or promote their own health. It does not account for sporadic, spontaneous action but relies instead on logical thought.

Health Belief Model

Basis of action and behavior

- *Understanding the consequences of an action or behavior*
- *Belief that one can influence outcomes (self-efficacy)*
- *Belief of susceptibility to disease or other negative consequences of behavior*

Social Learning Theory*Modeling*

- *Seeing adverse or positive results in others influences one's behavior*
- *Same-sex or similar models are especially salient*
- *Coding behavior into words, labels, or images results in better retention*

CDC Framework

- *Engage stakeholders (program staff and management, program participants users of evaluation)*
- *Describe the program (create logic model)*
- *Clarify evaluation design*
- *Collect credible evidence*
- *Justify conclusions (analyze data and compare to standards)*
- *Apply (use and share findings)*

Social Learning Theory

Albert Bandura's theory of social learning (Bandura, 1977, 1993) is the basis for numerous health programs (Card, 1999; Niego et al., 1998). Social learning theory states that behavior is heavily influenced by a person's knowledge and perceptions about expected results. Social learning theory supports the belief that people will act on what they believe will achieve desired results.

Witnessing what another person experiences can enhance one's self-efficacy, especially when the model is successful and is seen as similar to oneself. Quitting smoking seems possible to those who know someone who has already kicked the habit, especially if that person is a close friend, for example. Social learning theory is often used to explain the influence of television on aggressive behavior in children. Children learn to be aggressive by watching violence committed by characters with whom they identify or whom they admire. This effect is strongest when the characters are rewarded for their actions.

Many training programs are based on the theory because it recognizes cognitive and behavioral approaches to knowledge acquisition. Health programs incorporate social learning theory through a variety of means:

- Behavioral skills training
- Discussion
- Models
- Peer counseling
- Role-playing
- Video presentations

Centers for Disease Control Framework

The mission of CDC is to promote health and quality of life by preventing and controlling disease, injury, and disability. The federal agency published an evaluation framework for public health programs that consists of six core steps: engage stakeholders, describe the program (create a logic model), clarify evaluation design, collect credible evidence, justify conclusions, and apply findings. According to the authors, the framework is meant as a guide for professionals (1999, p. 11):

It is a practical, nonprescriptive tool, designed to summarize and organize essential elements of program evaluation. Adhering to the steps and standards of this framework will allow an understanding of each program's context and will improve how program evaluations are conceived and conducted. Furthermore, the framework encourages an approach to evaluation that is integrated with routine program operations. The emphasis is on practical, ongoing evaluation strategies that involve all program stakeholders, not just evaluation experts.

Defining the program is instrumental in guiding the evaluation. Defining a program should engage the stakeholders in describing the program's target audience, context, goals, strategies, and components. Focusing the evaluation design means considering and stating the purpose and methods of the evaluation. In addition to outlining and providing direction for the steps of the evaluation, the CDC framework includes standards for making the evaluation process and findings useful. See Milstein and Wetterhall (2000) for more information.

Methodology

There is a growing movement toward and, in some instances, an embracing of evaluation to better explain how health programs work and to improve them. Qualitative and quantitative data collection methods alike provide information about participants' knowledge, attitudes, and behavior. Experimental and field studies dominate much of health program research, along with studies that use various listening techniques, such as interviews and surveys. Table 4.1 summarizes the evaluation techniques found in the literature reviewed for this chapter.

Table 4.1. Methodology and Data

Technique	Type of Data Collected
Case study	<ul style="list-style-type: none"> Stakeholders' opinions and actions Includes listening strategies and observation
Experiment, field study	<ul style="list-style-type: none"> Effects on intervention, participants versus control population Analysis of medical records, attendance, program material
Listening strategies (interview, survey, letters, diaries)	<ul style="list-style-type: none"> Self-reported accounts of participants' attitudes, knowledge, behavior Often includes follow-up after program completion
Tracking	<ul style="list-style-type: none"> Actual accounts of behavior (contraceptive sales, purchase of cigarettes, participation) Policy changes (anti-smoking legislation, for example)

Case Study Data Collection

- Document analysis
- Interviews
- Literature review
- Observation
- Surveys

Case Studies

The case study is one means of representing multiple viewpoints and perspectives from a diverse group of people (including people targeted to change and those trying to effect change) and the nuances of an issue or change. In the study cited below, the researcher focused on understanding the process of social change.

Steven Ridini (1998) reports case studies from two Massachusetts communities that identified stakeholders' attitudes, knowledge, and behavior after the state had urged the inclusion of STD/HIV/AIDS prevention instruction in public schools and comprehensive sexual-health education in secondary schools. Data collection and creation of the case studies took about 3 years. Ridini gathered information from documents, interviews, and videotaped recordings of public meetings. The document analysis examined school and town newspaper articles, minutes from town meetings, grant proposals, and letters from community members. Ridini (1998) also collected demographic information. His analysis provided information on the history of the issue, and that information was used to develop interview protocols. Interviews, which lasted up to 3 hours, were conducted with more than 60 community members to provide information about differing perspectives

Let's Go with the Times

Let's Go with the Times is a Tanzanian radio soap opera about family planning. The following data were collected to assess the influence of the radio show:

- **Demographic and health surveys**
- **Listeners' letters**—Listeners were encouraged to write to the radio station about their connections, thoughts, and feelings toward the radio drama's characters. Most listeners identified with positive role models, and same-sex models were most salient.
- **Ministry of Health data**—Researchers asked new adopters of family-planning practices whether their change in behavior was influenced by the radio drama. 20% reported adopting family-planning strategies because of the show. The proportion of married women using contraception increased.
- **Personal interviews and surveys**—A Tanzanian research and education center gathered baseline data from households in the treatment and control areas. Interviewers conducted annual surveys in both areas and collected data on knowledge, attitudes, and actions regarding family planning, and exposure to and perceptions of the radio program. The perceived ideal marriage age for women rose.
- **Script content analysis**—A native speaker of Swahili analyzed the educational content of 98% of the radio drama scripts for the first 2 years.

on the decision-making process. The titles of the two chapters in Ridini's book illustrate what he found in the two communities: "Alpha Case Study: Students and Adults Working Together" and "Beta Case Study: If at First You Don't Succeed, Try, Try Again." Ridini's goal was to better understand the community decision-making process involved in discussing, framing, and implementing a sexual-health curriculum.

Experiments and Field Studies

Experiments and field studies are common in health program evaluation. The experimental or quasi-experimental design uses treatment and control groups and generally includes data collection at various intervals: pre-program (baseline) measurements and program, post-program, and follow-up measurements. Experimental designs incorporate document analysis, interviews, surveys, and other methods of obtaining information on participants' thoughts, feelings, and actions.

There was initial skepticism about using experimental design to evaluate public-health interventions, particularly for programs that target teen pregnancy and HIV/AIDS prevention. Service providers were reluctant to divert scarce program funds to evaluation because to do so might limit the population served in order to create a control group. Moreover, there was trepidation about the consequences of revealing negative results that might be perceived as showing programs as ineffective (Niego et al., 1998). Skepticism concerning evaluation is not exclusive to the health field. Program managers and other staff members may share some of the reluctance noted. In the past decade, however, evaluators and health program leaders

have begun to develop and improve instruments for program assessment (Niego et al., 1998). In addition, substantial positive effects have been observed and are supported by data collected from randomized, controlled experiments (Card, 1999).

An example of one experiment illustrates how evaluators measure attitude and knowledge changes attributable to a public-health intervention. To assess the effectiveness of *Let's Go with the Times*, a Tanzanian radio soap opera about family planning, evaluators studied knowledge, attitudes, and behavior among members of a treatment group consisting of listeners in a narrow broadcast area (Rogers et al., 1999). The listeners constituted the case subjects who were, in effect, a pilot group for the radio program. After a 2-year experimental phase, the broadcast area was broadened to include a comparison, or control, area, whose listeners were new to the program. In the data-gathering phases, evaluators surveyed listeners in both regions through interviews, listener feedback, demographic and health surveys, and information obtained from the Ministry of Health. Evaluators also analyzed the content of the first 2-years' scripts. One difficulty the evaluators noted—an issue in program evaluation—was accounting for confounding by external variables among the treatment and control groups. For example, although the control group was not in the pilot broadcast area, that group was exposed to all the other facets of Tanzania's national family-planning campaign—other radio programs among them. Notwithstanding, the evaluators concluded that differences between the treatment and control groups during the experimental years could be attributed to *Let's Go with the Times*.

Table 4.2 provides examples of teen pregnancy and STD/HIV/AIDS prevention program evaluations through experiments and field studies.

Table 4.2. Prevention Program Experiment and Field Study Models (from Niego et al., 1998)

Evaluation Design	Program
Experiment	<ul style="list-style-type: none"> • <i>Teen Talk</i> Six community agencies and 2 school grades compared the Teen Talk program with their regular sexual-health-education programs. Data were collected before the program began, upon its completion, and a year after it ended. Interview questions were designed to elicit information about participant behavior, including sexual activity and contraceptive use. • <i>Health care program for first-time adolescent mothers and their infants</i> Participants were assigned randomly to the treatment or control group (for infant health care only) at an urban teaching hospital. The 18-month-long field study included data collection on teenage mothers and their infants. Data came from medical documents, school records, immunization schedules, and interviews. At a final interview, participants were asked whether they had a subsequent pregnancy and whether they had returned to school.
Field study	<ul style="list-style-type: none"> • <i>Human Sexuality, values-based curriculum for 7th- and 8th-grade students</i> Three self-report surveys were administered in nine schools, using a control group. Surveys were administered before the program, at its end, and 3–4 months later. The questionnaire included more than 100 items about family background, sexual values, attitudes, knowledge, behavior, and communication with adults. • <i>Adolescents Living Safely</i> The field study included four interviews of males at a community-based agency serving gay youth. Interviews were conducted before the program, 3 months later (at program end), 6 months after the program's start, and 12-months after its conclusion. Interviews focused on sexual behavior.

Listening Techniques

Assessments of health programs rely heavily on listening techniques to acquire information on participants' knowledge, attitudes, and behavior. Data are often collected before, during, immediately after, and following a health program.

Surveys

Surveys and questionnaires are used extensively. A survey in American Legacy's evaluation of a youth empowerment tobacco prevention initiative included prompts such as

Listening: Does It Make a Difference?

Media campaigns allow governments and social change organizations to disseminate health information and promote healthful practices. Academic evaluation teams have examined letters from listeners to identify outcomes of educational radio dramas. The investigators reported that the programs were highly influential in affecting listeners' beliefs and behaviors at the individual and community level (listeners said they took action as result of the program). Changed behavior included modeling behavior of characters, setting up associations in the community to foster respect for and self-efficacy among women, and creating education programs for young people.

“This group allows me to have a say in planning events or activities” and “This group can reduce the amount of tobacco use in our community.” A Likert scale ranging from “strongly disagree” to “strongly agree” followed each prompt.

An assessment of Taiwan’s national anti-smoking campaign highlights the use of survey data and interviews (Hsieh, 1993). The health belief model guided the research. The evaluation team sought information on the link between increased knowledge and subsequent attitude and behavior changes. National Taiwan University conducted a national survey of smoking behavior. The researchers selected 25 areas of the country and identified 6 clusters of 15 households from each area. Adults in each household were interviewed. The survey consisted of 10 multiple-choice questions, focused primarily the health effects of smoking:

- What are the major health symptoms of heavy smokers?
- What is the addictive ingredient in cigarettes?
- How does smoking harm pregnant women?
- What is the harmful effect of passive smoking?

The researchers asked about involvement in smoking and awareness of the campaign. Participants were asked to use a Likert scale to rate their attitudes on smoking. The most effective (recalled and noted as influential) campaign messages were warning labels on cigarette packages and advertising to the youngest age group. Education was seen as a positive means of increasing public awareness of harmful effects, although smokers often ignored messages about the dangers of smoking.

Research Findings

The literature reviewed for this chapter revealed that personal characteristics are important in determining program outcomes. As noted in the two previous chapters, individual characteristics will affect attitude and behavior change, so they must be considered in the evaluative process. Perry (1999), for example, offers information on the sociodemographic, personal, and behavioral factors that shape youth behavior. Program components, of course, also are vital. One evaluator explores the complexities of program development and evaluation (Card, 1999, p. 267):

Individual Variables Influence Behavior

- Age
- Alcohol, drug use
- Attitudes
- Developmental stage
- Socioeconomic status
- Self-efficacy
- Self-esteem
- Sexual identity
- Social norms
- Values

There is an inherent weakness to addressing the “what works” question. ... [T]he most likely (indeed, perhaps the only) conclusion of any analysis will be one of apparent inconclusiveness and inconsistency. Some programs using Approach A work; others do not. Some programs using Approach B work; others do not. The inconclusiveness is built into the way the analysis is framed.

The overwhelming message is that tailoring the program to the situation and to participants is vital to successful results.

Program Effectiveness

Target Specific Behaviors, Teach Skills

Behavioral skills training and reinforcement are effective means of preventing high-risk sexual activity (NIH, 1997). The common characteristics of effective youth pregnancy and STD/HIV/AIDS prevention programs include a clear focus on providing basic information on ways to avoid unprotected intercourse, addressing social pressures, providing positive models, and offering structured opportunities to emulate models (role-playing activities). Skills training has also proven to be effective in school-based smoking-prevention programs (CDC, 1994).

There is no consensus, however, about which specific interventions are most effective. Although it is possible to replicate effective programs (or program elements), attention to the target audience and context is vital for success. A program that is successful in one area might not be effective in another setting. Table 4.3, summarized from Card (1999), highlights interventions and outcomes of pregnancy and HIV prevention programs.

Effective School-Based Programs

CDC (1994) work shows that effective anti-tobacco programs include all of the following elements:

- *Enforcement of school tobacco use policies*
- *Instruction of students about the short- and long-term negative physiological and social consequences of tobacco use, social influences on tobacco use, peer norms, and refusal skills*
- *Education in tobacco-use-prevention strategies in all grades (especially during middle school with reinforcement in high school)*
- *Program-specific training for teachers*
- *Family involvement in support of school programs*
- *Support for cessation efforts among students and all school staff who use tobacco*
- *Regular program assessment*

Table 4.3. Teen Pregnancy and HIV Prevention Programs

<p>Teen Talk Theoretical framework: Social learning theory, health belief model</p>	
<p>Core program components</p> <ul style="list-style-type: none"> • Collaborative school–community health center • Large-group, fact-based lectures • Small-group discussions, role-playing, games about factual lectures and decision making 	<p>Outcomes</p> <ul style="list-style-type: none"> • Sexual experience and gender affected results • Especially effective for males; sexual activity delayed for males, increased contraceptive use among sexually active • Comparison group (female virgins) more likely to use contraceptives when sexual activity began
<p>Adolescents Living Safely: HIV/AIDS awareness, attitudes, actions for gay, lesbian, bisexual teens Theoretical framework: Social learning theory: Teens know safe sex practices but lack the skills to act safely</p>	
<p>Core program components</p> <ul style="list-style-type: none"> • Factual HIV content through participant-centered activities and prevention videos • Counseling and personalized case management, health care and risk assessment, small-group discussions, video presentations, art, self-acceptance, access to contraceptives • Social skills training, behavioral self-management, peer support • Recognizing and communicating own feelings • Observing, imitating, practicing behavior 	<p>Outcomes</p> <ul style="list-style-type: none"> • Reduction in high-risk behavior by participants at 3-month post-program interview • No significant effect on contraceptive use during intercourse or oral sex; use returned to baseline at 12-month follow-up • 3- and 6-month measurements showed results varied by ethnicity: Black and Hispanic males significantly reduced high-risk acts • High-risk activity of participants who engaged in sex-for-hire increased despite program
<p>Health care program for first-time adolescent mothers and their infants Theoretical framework: Based on challenges and needs of teen mothers, review of services provided to teen mothers in other countries, and providing information on baby health care</p>	
<p>Core program components</p> <ul style="list-style-type: none"> • Hospital-based mother-and-baby clinic • Flexible services • Routine infant health check-ups • Family-planning discussions, referrals to clinics • Parenting education through videotapes; discussion; skills training, including one-on-one discussion, modeling, and distribution of educational information 	<p>Outcomes</p> <ul style="list-style-type: none"> • More likely to attend clinic regularly • Increase in full baby immunizations • Reduced emergency room use for infant health care • Fewer repeat pregnancies • No effect on return to school

Table 4.3. Teen Pregnancy and HIV Prevention Programs *Continued***Human Sexuality—Values and Choices**

Theoretical framework: Fishbein's (in Niego et al., 1998) behavioral intention model: Action results from intention to act and is based on beliefs about outcomes, perceptions of others' opinions, motivation to comply with those opinions

Core program components

- 15 student lessons, 3 adult lessons, video presentations
- Role-playing in school classrooms
- Values-based promotion of abstinence and healthy social relationships (equality, honesty, respect, responsibility, promise-keeping, self-control, social justice)

Outcomes

- Increased knowledge about risks
- Increased support for postponing sex
- Decreased intention to have sex

Follow-up: 3–4 months after program

- Less likely to believe boys have stronger or uncontrollable sex drives
- More likely to accept possibility of STD or pregnancy

A list of key program components was derived from a review of effective programs for preventing teen pregnancy and sexually-transmitted-disease-prevention programs (Niego et al., 1998). Effective pregnancy and HIV prevention programs included at least half of the following:

- Adult involvement
- Case management
- Group discussion
- Lectures
- Peer counseling
- Public service announcements
- Role-playing
- Video presentations

Changing Behavior Through Attitudes and Modeling

Law and Singhal (1999) analyzed letters sent by listeners of *Happiness Lies in Small Things*, an educational radio drama broadcast in India. The educational themes of the drama included gender equality, family planning, and youth delinquency. Driven by social

Education and Health

Some researchers find that education that involves attitudes can be a positive means of increasing public awareness about smoking's harmful effects. People who have positive attitudes toward smoking are more likely to use tobacco (Hsieh, 1993). The assessment of a national anti-smoking campaign in Taiwan reported that smokers often ignore messages about tobacco's dangers. The implications are that effective programs change attitudes to impact smoking behavior. See "Guidelines for School Health Programs to Prevent Tobacco Use and Addiction" (CDC 1994) for specific recommendations on how programs should target knowledge, attitudes, and behavioral skills.

learning theory, efficacy was defined by a belief, a stated behavior, or as an expression of individual or collective will. During one broadcast year, the radio station received 130,000 letters (accounting for less than 1% of the total estimated audience). The researchers analyzed a random sample of the letters. Listeners' words, thoughts, feelings, and statements of behavior provided evaluators with data (Law and Singhal, p. 358):

Many letters expressed *affective* states ranging from disappointment and disgust with antisocial practices, to fervent entreaties for recognition and change, to a renewed sense of confidence or assurance at having learned new ideas. Strong, deep feelings had been stirred regarding both personal and social issues and the individual's capabilities in coping with these issues.

Researchers gained insight into knowledge gain and behavior through "private unbiased research journals" (in addition to the letters, the researchers completed a case study that included interviews with the program's producers, and they analyzed the content of more than 100 episodes of the radio show).

Chapter Summary

Theories and Methods

Health programs rely on a strong theoretical base.

- *Health belief model.* The health belief model asserts that people must understand and accept the value of changing their behavior before they are able to change. One must believe in the negative consequences of not acting in order to adopt a behavior.
- *Social learning theory.* Social learning theory states that behavior is influenced by what a person thinks and knows about expected results. Beliefs and knowledge are gained through observation of the results of others' actions. Behavior, therefore, is strongly influenced by the positive and negative consequences we observe and subsequently model.
- *Centers for Disease Control.* The CDC framework consists of six core steps: engage stakeholders, describe the program (create a logic model), clarify evaluation design, collect credible evidence, justify conclusions, and apply findings.

Experimental and field studies dominate much of health research, along with studies that use various listening techniques, such as interviews and surveys.

Key Findings

Successful interventions are as varied as their targeted outcomes and participants. Only since the 1970s has there been an effort to track the success of health programs that target behavior change in young people.

The overwhelming message is that tailoring interventions to individual participants and teaching behavioral skills is vital to successful programming. There is a thorough review of effective programs that target teen pregnancy and STD/HIV/AIDS prevention as well as other school-based programs, and the evaluations provide useful information on effective program elements. Youth programs that consider multiple components—school environment, family, and the community—appear to be most effective. The most effective programs incorporate multiple techniques.

Effective Youth Programs

- *Include multiple components*
- *Participant centered (tailored)*
- *Target specific behaviors*
- *Teach skills*

Summary of Studies

Table 4.4 is a summary of the in-depth descriptions of health interventions and individual program evaluations presented in this chapter.

Table 4.4. Summary of Studies: Health Programs

Change Sought	Effect	Intervention	Population	Method	Follow-Up	Reference
Family planning Attitude Knowledge Behavior	+	Mass media, radio: Entertainment–educational radio drama aimed at influencing listeners’ efficacy, values, behavior	Adults (northern India)	Analysis of listener mail	No	Law and Singhal, 1999
Behavior Condom use, num- ber of births, visits to clinic Attitudes	+ +	Mass media, radio (family planning and HIV prevention focus): Entertainment–educational radio drama aimed at influencing listeners’ knowledge and behavior	Adults (St. Lucia)	Pre-, post-intervention tests, focus groups	Yes	Vaughan et al., 2000
Knowledge, attitude, behavior	Mixed	Mass media, radio: Entertainment–educational radio drama aimed at influencing listeners’ knowledge, attitudes, behavior	Adults (Tanzania)	Experiment, interviews, surveys, document analysis	5 Annual sur- veys	Rogers et al., 1999
Pregnancy, HIV Decision-making process Knowledge Attitude	NA	Adoption of sexual health curriculum	Community	Case studies	No	Ridini, 1998
Knowledge of risks Attitudes: increased support for post- poning sex Behavior: decreased intention to have sex	+ + +	Human sexuality: 15 student lessons, 3 adult lessons, videos, role-playing; promotion of abstinence, healthy social relationships (equality, hon- esty, respect, responsibility, promise- keeping, self-control, social justice)	Grades 7–8	Field study, survey	3–4 Months	Niego et al., 1998

Table 4.4. Summary of Studies: Health Programs *Continued*

Change Sought	Effect	Intervention	Population	Method	Follow-Up	Reference
Behavior Attend clinic regularly Full baby immunizations Reduce emergency room use Fewer repeat pregnancies School return	+ + + + -	Health care program: Hospital-based teen-and-baby clinic; flexible, routine infant health check-ups; family-planning discussions; referrals to clinics; parenting education through videotapes, discussion, skills training, one-on-one discussions, modeling	First-time adolescent mothers and their infants	Experiment, document analysis, interviews	Yes	Niego et al., 1998
Knowledge, attitude, behavior	Mixed	Teen Talk: Large-group lecture, small-group discussions, role-playing, video programs, decision-making skills	Adolescents	Experiment, interviews	12 Months	Niego et al., 1998
Behavior, knowledge Reduce high-risk behavior Contraceptive use	+ Mixed	Adolescents Living Safely: Counseling; personalized case management; health care and risk assessment; small-group discussions; video programs; art; access to contraceptives; social skills training; behavioral self-management; peer support; observing, imitating, practicing behavior	Gay youth, males	Field study, interviews	3, 6, 12 Months	Niego et al., 1998
Reduce smoking Knowledge, attitudes Knowledge-behavior connection	+	Anti-smoking campaign: Label warnings; advertising, vending-machine ban; anti-smoking messages in mass media, school programs, other venues to warn of health effects	Adult males (Taiwan)	Document analysis, interview	Yes	Hsieh, 1993
Behavior Smoking reduction, healthy diet, exercise	+	Community-based heart disease prevention, including youth education element (school-based, family component, community-mass media, peer leaders, skills training)	Adults	Experiment: compare region with control group, annual survey over 6-year period	No	Perry, 1999

Table key: *Change Sought*, intended effect of program; *Effect*, positive (+), negative (-), or mixed study findings; *Intervention*, type of health program and techniques used; *Population*, study participants; *Method*, data collection technique; *Follow-Up*, if and when data collection took place after the intervention; *Reference*, author and date of study; *NA*, not applicable.

References

- Bandura, A. 1977. Social learning theory. Englewood Cliffs, N.J.: Prentice-Hall.
- Bandura, A. 1993. Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist* 28(2):117–148.
- Card, J. J. 1999. Teen pregnancy prevention: Do any programs work? *Annual Review of Public Health* 20:257–285.
- CDC (Centers for Disease Control and Prevention). 1994. Guidelines for school health programs to prevent tobacco use and addiction. *MMWR* 1994; 43(No. RR-2):1–15.
- CDC. 1999a. Best practices for comprehensive tobacco control programs. Atlanta, Ga.: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.
- CDC. 1999b. Framework for program evaluation in public health. *MMWR* 1999; 48(No. RR-11).
- Hsieh, C. R. 1993. Smoking, health knowledge and anti-smoking campaigns: An empirical study in Taiwan. Institute of Economics discussion paper (November):9335.
- Law, S., and A. Singhal. 1999. Efficacy in letter writing to an entertainment–education radio serial. *Gazette* 61(5) (repr.), Thousand Oaks, Calif.: Sage.
- Milstein, B., and S. Wetterhall. 2000. A framework featuring steps and standards for program evaluation. *Health Promotion Practice* 1(3):221–228. Thousand Oaks, Calif.: Sage.
- Niego, S., M. J. Park, and Card, J. J. 1998. Promising teen pregnancy and STD/HIV/AIDS prevention programs. Los Altos, Calif.: Sociometrics Corporation.
- NIH (National Institutes of Health). 1997. Interventions to prevent HIV risk behaviors. NIH consensus statement 1997 Feb 11–13, 15(2):1–41. Online, accessed Aug. 24, 2002: http://consensus.nih.gov/cons/104/104_statement.htm.
- Perry, C. 1999. Creating health behavior change: How to develop community-wide programs for youth. Thousand Oaks, Calif.: Sage.
- Ridini, S. P. 1998. Health and sexuality education in schools: The process of social change. Westport, Conn.: Bergin and Garvey.
- Rogers, R. M., P. W. Vaughan, R.M.A. Swalehe, N. Rao, P. Svenkerud, and S. Sood. 1999. Effects of an entertainment–education radio soap opera on family planning behavior in Tanzania. *Studies in Family Planning* 30(3):193–211.

- Salazar, B. K. 1991. Comparison of four behavioral theories. *AAOHN Journal* 39(30):129–135.
- Vaughan, P. W., A. Regis, and E. St Catherine. 2000. Effects of an entertainment–education radio soap opera on family planning and HIV prevention in St. Lucia. *International Family Planning Perspectives* December.

Resources

Centers for Disease Control and Prevention <http://www.cdc.gov/eval/>

The CDC Evaluation Working Group web site includes a recommended evaluation framework for use in public health programs and links to other resources.

EC/UNFPA Initiative for Reproductive Health in Asia

http://www.asia-initiative.org/m_and_e_2002.html

The Monitoring and Evaluation Manual of Sexual and Reproductive Health Interventions is available at this site. The manual provides information on evaluating sexual and reproductive-health projects or portions of larger programs.

University of Ottawa Epidemiology and Community Medicine Faculty of Medicine

www.uottawa.ca/academic/med/epid/what.htm

Shreeniwas, S., and R. K. Homan. 1995. *Knowledge and attitudes about condom use among college students in urban India*. Number 63, working paper series. Stanford, Calif.: Morrison Institute for Population and Resource Studies.

Chapter 5

Social Marketing

In This Chapter

Theoretical Frameworks

- Antecedents, moderators, consequences
- Cost–benefit analysis

Methodologies

- Case studies
- Media metrics
- Social return on investment

Research Findings

- Effective social marketing
- Fundraising and civic engagement
- Evaluation

Chapter Summary

- Theories and methods
- Key findings
- Summary of interventions

References

Resources

Social marketing, which applies traditional techniques of commercial marketing to areas of public benefit, is not intended to promote products or sell services but rather aims to change individual and group behavior to benefit society at large. Charitable giving, community involvement, nuclear disarmament, safer sex, responsible environmental practices, and even “random acts of kindness” are all promoted by social marketing.

“Focusing the Concept of Social Marketing” (Karim and Rangan, 1991) provides two examples of social marketing initiatives that represent the challenges that are often encountered by those seeking behavior change. The first involves a family-planning effort in rural Bangladesh. There was opposition to the effort. Farmers place economic value on children, who contribute to a farm’s success by working: The more children, the more productive the farm. Muslims in the area were opposed to family planning as unnatural and sacrilegious. The second example is closer to home. Anti-smoking efforts must compete with the massive advertising budgets of the tobacco companies, whose profits depend on selling a dangerous product. The research on the dangers of smoking and of second-hand smoke is universally accepted, but anti-smoking campaigns have been only marginally successful because they must compete with the glamorous images of smokers depicted on billboards and in print ads. Both examples illustrate the difficulty faced by social marketing efforts. Behavior change is sought in these cases through advertising, mass media, education, and community programs.

There is a good deal of overlap between the interventions discussed in the previous chapters and social marketing efforts. The target audience (potential adopters of change) must be convinced that the rewards of behavior change exceed the cost of altering behavior. The most evident challenge in social marketing includes this difficulty—bringing about change in people or groups who do not individually profit from or understand the value of behavior change (Karim and Rangan, 1991).

With creativity and perseverance, nonprofit organizations can measure their success in achieving their mission—by defining the mission to make it more quantifiable, by investing in research to show that specific methods work, or by developing concrete microlevel goals that imply success.

(Sawhill and Williamson, 2001, p. 107).

Assessments of social marketing generally include the use of theoretical frameworks based on traditional for-profit marketing, psychology, sociology, behavior modification, and education.

The literature reviewed for this section of the report focuses on donor studies and volunteering, and it includes “how to” articles that set out guidelines for taking action. Less work has been documented on conceptualizing assessment strategies and measuring outcomes. This chapter introduces social marketing interventions and it highlights some of the methods that have been used to explore and measure the effectiveness of social marketing. The object here is to provide readers with conceptual frameworks and evaluation techniques that will promote more effective programming.

Theoretical Frameworks

Social marketing aims to alter individual attitudes and behavior to benefit the larger, even global, community (Karim and Rangan, 1991).

Behavior modification theories are strong influences in initiative development and analysis. The theoretical models discussed here are additions to the frameworks presented in the previous chapters. The models are included here to describe the thinking that underlies social marketing concepts and practices.

Antecedents, Moderators, Consequences

According to the model of antecedents, moderators, and consequences (AMC), fundraising, volunteer recruitment, and other efforts to promote civic engagement are influenced by selected, controllable agents, such as an organization’s image, actions, and communication techniques (Figure 5.1). AMC also accounts for moderating factors, including donor variables, and for external variables such as the state of the economy, social norms, and government policies.

Figure 5.1 represents the conceptual framework of the factors that influence giving to charitable organizations, and it provides a basis marketers can use in adopting practices and strategies to encourage participant engagement by donating time, money, or even blood, for example.

The “Marketing” in Social Marketing

To a large extent, social marketing relies on for-profit—or consumer-marketing strategies. For example, the “four Ps” of the for-profit arena are at times used in social marketing:

- *Product: targeted behavior change*
- *Pricing: individual’s cost or effort of change*
- *Promotion: communication of social benefits of change*
- *Placement: where the message about change is disseminated*

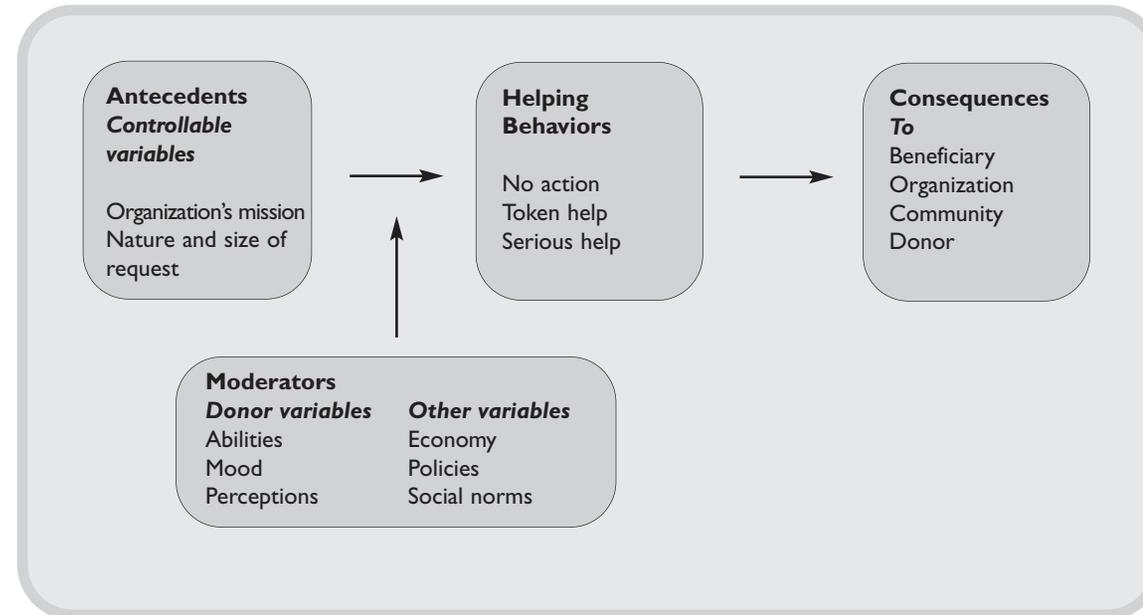


Figure 5.1. The AMC model for enhancing helping behavior, from Bendapudi et al. (1996, p. 38).

Cost–Benefit Analysis

Cost–benefit analysis is a traditional marketing strategy that weighs the perceived costs and benefits to the potential adopters of change (whether it be consumers or target audience). By considering how a person or group understands a situation or believes a specific change or action can improve quality of life, for example, social marketing practitioners can tailor the message to the audience. Rangan and colleagues (1996) provide a model for approaching social change by considering the value and feasibility of traditional consumer marketing strategies.

Figure 5.2 shows a cost–benefit model to explain the perceived costs of change to the targeted adopters of change. The upper-left-hand quadrant is for change that requires minimal individual effort (or cost) and offers a relatively high reward for behavior change. Promoting or instigating change of this kind is relatively easy and might only require

informing the audience about the issues. On the opposing side, the lower-right-hand quadrant shows the high cost or effort involved with change and the relatively intangible but potentially high societal gain.

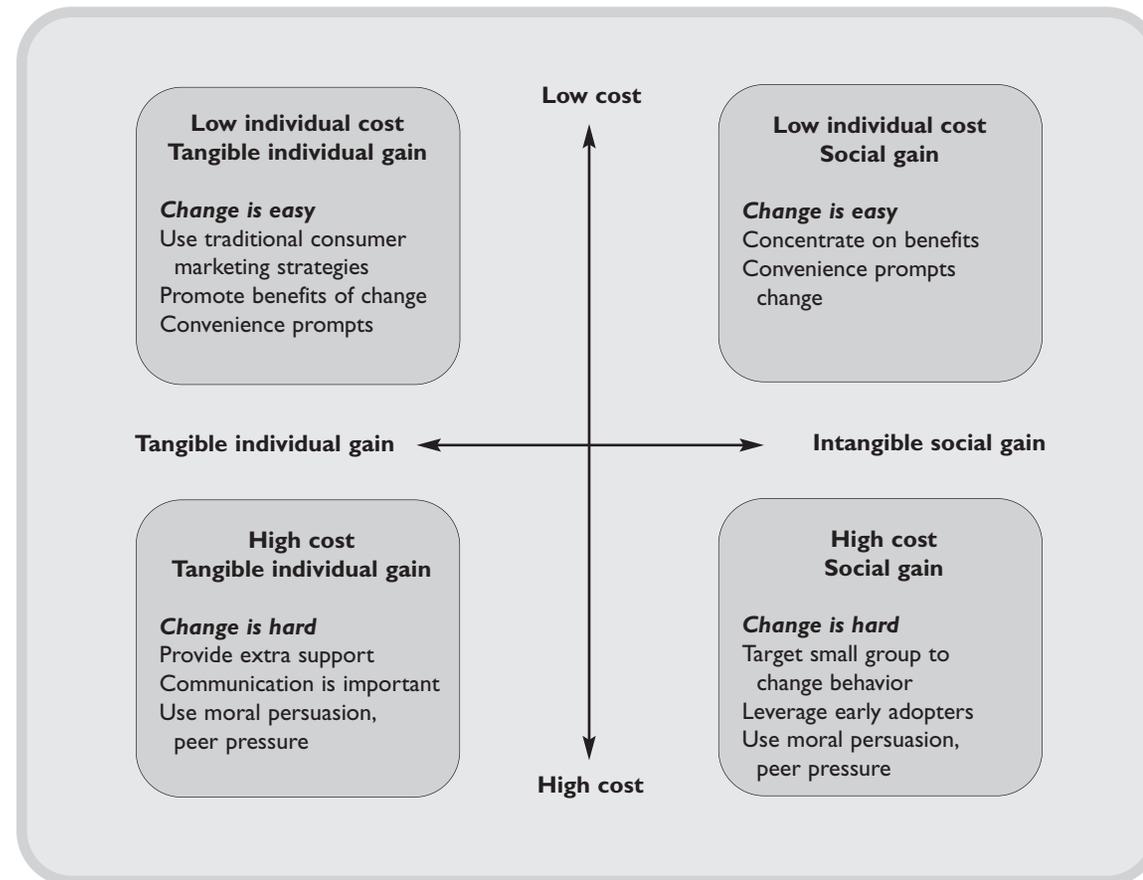


Figure 5.2. Cost–benefit analysis.

Cost–benefit analyses are useful in developing effective strategies for change. The kinds of changes represented by the lower-right-hand quadrant require significant effort by or large costs to the system or community. And even if they are successful, there might be no clear and immediate or even eventual gain. Rangan and colleagues (1996) suggest repositioning issues from the lower-right to the lower-left quadrant by focusing on any potential short-term or individual gain.

Cost–Benefit Analysis

In “Carrots, Sticks, and Promises: A Conceptual Framework for the Management of Public Health and Social Behaviors,” Rothschild (1999, pp. 25–30) describes the potential role and effectiveness of education, marketing, and law based on cost–benefit analysis:

Education refers to messages of any type that attempt to inform and/or persuade a target to behave voluntarily ... but [that] do not provide immediate reward or punishment. ... Education will be an appropriate tool when individual self-interest is strong and consistent with societal goals but the target is merely uninformed.

Marketing refers to attempts to manage behavior by offering reinforcing incentives and/or consequences in an environment that invites voluntary exchange. ...

Marketing will be appropriate when the level of self-interest is insufficiently consistent with societal goals to elicit behavior. Law involves the use of coercion to achieve behavior in a nonvoluntary manner. ... Law will be appropriate when the preexisting self-interest of the target cannot be overcome with additional rewards through exchange, when rewarding is inconsistent with societal gains, or when the rights of the target are believed to be irrelevant.

For example, in the Bangladeshi family-planning example above, social marketers discovered that women were the better population to target as potential adopters of family planning practices. The social marketers addressed the issue by providing contraceptives and information to women through local clinics and by creating a media campaign to educate men about communicating with their wives about family planning.

Rothschild (1999, p. 30) writes: “the more favorable the individual costs–benefit relationship ... the more likely education will be sufficient. The less favorable ... the more likely that law will be needed. The middle cases (mixed cost and benefits) would be most likely to use marketing solutions to improve the cost–benefit relationship.” Individual motivation, opportunity, and ability also influence behavior change. Clearly, cost–benefit analysis can be applied to a range of initiatives, including the promotion of family planning, buying local produce, or recycling.

Methodologies

How do organizations, researchers, and evaluators judge social marketing success? The Nature Conservancy experience (Chapter 1) shows that the challenges are formidable and pervasive. The Conservancy’s approach—interview professionals from other nonprofit organizations, reconsider the mission, state the targeted outcomes—illustrates the inherent difficulties for quantifying and assessing the work of social marketing.

The success of a social marketing effort depends on influencing the targeted audience to bring about substantial change. Data to gauge success can be collected in many of the ways described in previous chapters: case studies, document analysis, experiments, interviews, and surveys. Table 5.1 lists measurement techniques for evaluating social marketing initiatives and efforts.

Table 5.1. Social Marketing Assessment Data

Technique	Type of data collected
Case studies	<ul style="list-style-type: none"> • Information on intervention, targeted audience, change process • Tools include document analysis, focus groups, interviews, observation and surveys
Media metrics	<ul style="list-style-type: none"> • Include “built-in” measurements such as toll-free numbers • Interviews and surveys to determine awareness, knowledge of issues
Social return on investment (SROI)	<ul style="list-style-type: none"> • Quantifies the social–economic value of nonprofit work through financial calculations (true cost accounting and discounted-cash-flow models) • Qualitative descriptions of the intervention, organization, and people served through interviews and observation

Case Studies

Case studies are pervasive in the documentation of social marketing initiatives. They provide a way to illustrate an intervention’s effectiveness, direct the expansion of successful efforts, and disseminate lessons learned by similar organizations or communities. The case study can provide in-depth information about multiple groups, including those targeted to change their behavior and those who instigate change. Two interventions explored through case studies are described below.

Aravind Eye Hospital

A detailed case study, *The Aravind Eye Hospital, Madurai, India: In Service for Sight*, conducted by Professor V. K. Rangan (1993) at Harvard Business School, traces the operations, condition, strategy, staff training, procedures, and effectiveness of eye surgeries at the hospital in India. The case study states that, in Asia, blindness rates are about six times higher than they are in industrialized countries (1.5% of the population as opposed to a rate that ranges from 0.15% to 0.25% in developed nations). More specifically, there is a large backlog of patients in India who need cataract treatment (cataracts are the main cause of blindness in India).

The case study uncovers the hospital founder’s mission and strategy through personal interviews, observations, and medical records to provide an intensive qualitative study.

The reader gains an understanding of the challenges faced by the organization and of the inner workings and social benefits of the eye hospital and its affiliated networks. The conclusion states the founder's challenges (Rangan, 1993, p. 15):

My goal is to spread the Aravind model to every nook and corner of India, Asia, Africa; wherever there is blindness, we want to offer hope. Tell me, what is this concept of franchising? Can't we do what McDonald's and Burger King have done in the United States?

Internal evaluation by the hospital staff took the form of personal communication with patients. When patients had received a recommendation to return to the hospital for surgery but did not, a hospital research team visited villages and interviewed people at home. The explanations given—lack of funds for food and transportation, fear of surgery—were addressed by the organization and its supporters. The hospital called on respected people in the village to communicate the value and safety of the cataract surgery, and they organized bus transport to ease the burden of transportation and provide a support group for villagers who could travel together. The Aravind case study illustrates organizational effectiveness, discusses expansion models, and disseminates lessons learned. Another example is provided below.

Boston Fights Drugs

The case study *Boston Fights Drugs* (Lawrence and Rangan, 1987) describes the process by which a research team selected a core data collection method—the focus group—to formulate a plan to reach and influence urban youth. The research team met in focus groups with adolescents and it reviewed documents and conducted interviews with policymakers and drug-prevention-program directors. Graduate students approached the research according to the following protocol:

- Develop an understanding of drug use and of prevention programs through document analysis and interviews.
- Develop a drug use model to create a research question and define a protocol for convening focus groups.
- Screen participants for focus groups.

- Meet with adolescents to discuss their attitudes about advertising, drug awareness, and information sources on drug use and prevention.

The research yielded formative information that was used in the design of an advertising campaign aimed at Boston's young people. The focus groups also provided detailed information about effective prevention methods (Lawrence and Rangan, 1987):

We had a policeman come into school and talk to us about drugs. And he listened to our ideas and talked to us for a while (p. 19).

“Take me out to the ballgame” [an ad showing a young woman using drugs and later being wheeled into a hospital] was scary. Like that could happen (p. 20).

The case study emphasized the importance of modeling, story telling, and creating realistic portrayals of drug use for those who want to communicate with teenagers.

Media Metrics

Social marketing makes extensive use of media campaigns. Public service announcements (PSAs) and other radio and television advertising and educational programming are used to inform and influence attitudes and behavior. The effectiveness is governed by the medium's reach: In some places and for some situations radio will reach a larger audience than will television advertising; in other cases, full-page newspaper ads will work better. Effectiveness also depends on tailoring the message to the audience. One way to target an audience is to use information from focus groups and interviews. Part of evaluation therefore is measuring both the reach and the effectiveness of the message to communicate with and influence the targeted audience as planned.

Some PSAs incorporate data collection tools within the message (Rangan and Kramer, 1997). Television campaigns with a toll-free (or “1-800”) telephone number provide evaluators with one means of assessing the use and level of interest in the message, for example. Monitoring the use of the telephone number can give demographic information about the reach of the ad and identify the need for more information. For example, an 80% monthly increase in calls to an information number for runaway kids was linked to

Ad Council Survey

The Ad Council, a nonprofit organization that works with organizations to promote social change, assesses effectiveness through various means, one of which is the use of survey data collected by external marketing consultants. A survey of 1000 adults provided the Ad Council with the following information about an ad that promoted sober driving (Rangan and Kramer, 1997): Among those who saw the advertisement, 80% felt more strongly about preventing people from driving drunk. After exposure to the ads, 79% of respondents personally acted to prevent drunk driving. The primary action taken was to discuss the problem with a friend (54%). About 40% of respondents said they stopped a friend from drinking and driving. 25% said they stopped drinking and driving themselves.

Questions Asked by SROI's Creators

- How do we measure the success?
- For each dollar invested, what is the benefit to individuals and society?
- How can both investor and investee be assured that each dollar is maximizing its value?
- How can we calculate the social return on our investments?

(SROI Methodology, 2000, p. 6)

Ad Council PSAs. A media campaign to recruit teachers resulted in a 400% increase in inquiries in one year. What this kind of data may not provide, however, is information about subsequent behavior or about specific individual reactions to the information presented.

Social Return on Investment

The Roberts Enterprise Development Fund created the model of social return on investment, or SROI, to measure the effectiveness of nonprofit organizations, specifically those dedicated to employing and training disadvantaged people. Since the late 1990s, the organization has worked to refine the concept and create metrics to calculate SROI. More information about SROI, including how to calculate returns, is available online from the fund's web site: http://www.redf.org/pub_sroi.htm (accessed Aug. 24, 2002).

An explanation of social economic value is best understood through the description provided in *SROI Methodology* (2000, ch. 1, p. 12):

Social value can be found in a wide variety of activities from anti-racism efforts, community organizing, environmental protection and arts support efforts, to a family moving from welfare to work. The psychological impact on an individual whose family has moved from welfare to work may be significant but hard to monetize. ... Measures of economic value are standardized and support the basis for most financial activity in the world. In the social value arena there are factors that are beyond measurement, yet clearly are of value and worth affirming. In between these two poles of value creation lies socio-economic value. Socio-economic value meas-

urement builds on the foundation of economic value measurement by quantifying and monetizing certain elements of social value, and incorporating those monetized values with the measures of economic value created.

SROI is based on a discounted-cash-flow model and calculation of program's economic effects. The financial analysis tools used include true cost-accounting analysis, discounting and cost of capital, discounted cash flow, and net present value. *SROI Methodology* includes an in-depth discussion of the method's calculations (ch. 2, pp. 16–18). In summary, true cost accounting methods are used. Nonprofit organizations track financial effectiveness by separating nonprofit enterprise revenue from business revenue and social operating expenses from business expenses. Economic value is based on the free cash flow the initiative it is expected to generate over its lifetime (as noted in the report, value is based on cash, rather than income, because only cash can be re-invested). In SROI analysis, different rates are used to discount social purpose cash value and enterprise value:

Discounted cash flow is the process by which one discounts the cash flow projections based on an appropriate discount rate. ... [I]t is the value of future expected cash receipts and expenditures at a given common date. [Businesses] use this approach to perform valuations of potential investment opportunities. It converts a business's forecasts into an estimate of its value. Detailed, multi-year forecasts of free (available) cash flows are discounted at an appropriate discount rate to arrive at an estimated present value.

The point of the financial models is to value the economic return a program provides to society. The creators of SROI emphasize that the calculations explained above exist in a broader context that considers other issues and information.

SROI reports also include qualitative data, social-impact information, and general business data. For example, data from employee surveys are included in the *Einstein's Café SROI Report (2000)*. Information was collected in personal interviews with employees at the time they were hired (baseline information) and 6 months later. The data gathered included information about income, housing stability, public assistance, social services, and criminal justice involvement.

Youth Industry SROI Report (2000)

As a result of the social savings and new income per employee generated, it is projected that Einstein's Café will provide society with total social savings and new taxes of more than \$7.7 million going forward (in current dollars). Einstein's profits will fully cover its social operating expenses and capital requirements while returning \$563,582 to Youth Industry, its parent nonprofit.

It is difficult to define success or measure the effectiveness of social marketing. Consider, for example, how difficult it is to calculate the value of learning to read, or to establish a success rate for targeted outcomes in terms of number of participants, or to measure the benefit of worldwide biodiversity. A few of the shortcomings and challenges of evaluation include those identified in the literature by Gallagher and Weinberg (1991), Rangan and colleagues (1996), and Rangan and Kramer (1997):

- Difficulty or challenge in determining and defining success
- Difficulty in tracking audience
- Lack of good performance metrics
- Multiple stakeholders and managing opinion of investors

Essential Components

- **Opinion leaders:**
Credible individual or group of experts committed to social benefits of change
- **Multiple credible sources:** *More than one reliable place to verify the message and appeal to various people and groups*
- **Intensive personal support:** *Opportunity for intensive personal support to members within the community*
- **Communication:**
Advertising, promotions

Research Findings**Effective Social Marketing**

According to various experts in the field (Bendapudi et al., 1996; Karim and Rangan, 1991), the components of effective social marketing include having an opinion leader, having information come from multiple credible sources, the possibility of intensive personal support, and communication (a means of spreading the word). Each of the four components was evident in literature reviewed for this chapter.

Opinion Leaders

The Aravind Eye Hospital epitomizes the value of a strong opinion leader in effecting change. The founder, who established the eye hospital in his retirement, not only expanded and provided services to everyone he could, but he was able to attract his family and other doctors from higher paying positions to engage in his effort to provide hope and treatments for visual impairments and blindness. His project also appealed to respected people in the villages to urge their neighbors to go to the hospital for treatment.

Credible Message Providers

The importance of having credible sources of information emerged in the *Boston Fights Drugs* case study (Lawrence and Rangan, 1987). The adolescents who were asked what would work to keep kids off drugs encouraged more open discussion between peers; stories from young people in similar circumstances; examples of how young people are hurt, punished, or die as a result of drug use; and specific education about the harmful impact of drug use.

Personal Support

Personal support is evident in many forms. In family-planning efforts, for example, the women may receive educational literature, home visits, and contraceptives as well as peer support forums.

Communication

An excerpt from the Ad Council's web site concerning research on PSAs shows that PSAs can change attitudes and behavior (http://www.adcouncil.org/research/impact_of_psas/, accessed Aug. 24, 2002):

Increased research funding and assistance has allowed us to uncover more accurate information about the effect of our messages. The results conclusively show that public service announcements are an effective means of communication. ... Even if the message is used alone or is unwelcome and intrusive, the PSAs increase awareness, reinforce positive beliefs, intensify personal concern and move people to action. Our PSAs benefit from regular exposure—the longer our ads run, the greater the effect they have on the audience.

Advertising or promotion is an essential component of effective social marketing work. Communication may be through media or word of mouth.

Fundraising and Civic Engagement

Bendaputi and colleagues (1996) focused on the AMC model to explain and influence donors' decision-making process. Significant differences exist between those people who are willing and those who are not willing to make charitable donations. There is a strong link between self-efficacy and helping behavior. In addition, mood also influences giving patterns: People who are happy are likely to be encouraged by the rewards of taking action; unhappy people might more likely be swayed by the punishments or negative attributes of not acting.

The decision-making process can be divided into four distinct steps: perception of need, motivation, behavior (or action), and consequences.

Giving Time and Money Intention and Behavior

- *Attitudes and normative influences affect behavioral intentions. Information and normative requests are more effective in bringing about change than either is alone.*
- *Emotional appeal evokes greater willingness to help.*
- *Labeling potential donors elicits greater intentions to help and more helping behavior.*
- *Labeling also increases positive perceptions of the organization.*
- *Messages delivered from familiar and credible sources are more likely to be influential.*

Perception of Need

Understanding that a need exists can result from exposure to facts, information, and the disparity between beneficiaries' reality and a desired state of well-being. When need is perceived as an external element rather than as a self-induced state, people are more apt to donate because of their desire to reduce inequity between themselves and their beneficiaries. According to Bendaputi and co-workers (1996, p. 37), credibility is an essential factor in effective promotional strategies:

The image of a charity thus may be the single most critical element of its promotional program because it may determine whether the first step of the helping decision process—perception of need—is initiated. ... [M]essages from charities that are perceived as familiar or well known, efficient, and effective result in greater perceptions of need and greater helping behavior.

Motivation

Motivation is the next step in the donor's decision-making process. A donation with an egoistic motivation is self-interested: By giving money I am helping myself in some way. Altruistic motivation is rooted in the desire to go beyond oneself to benefit an organization or person in need: My donation is intended for your betterment.

Behavior

Once donors are motivated, a cost–benefit analysis determines the degree of helping behavior, which can range from no action at all, to a small contribution, to a large one. Social comparison and strategic altruism can influence action.

Consequences

The “foot in the door” is one approach for escalating from a small request (perhaps one hour volunteering or a nominal financial donation) to larger involvement (a monthly commitment or an annual pledge). Another approach takes an opposing strategy—begin with a large request but follow up with solicitation for a smaller, more manageable response. According to the AMC model, this tactic, called “door in the face,” is most effective for soliciting new donations.

Evaluation

The development of the SROI illustrates a commitment to understanding and measuring work aimed to benefit society at large. In addition to SROI, the notion of integrated assessment tools (such as toll-free telephone numbers) provides data for the evaluation of marketing work.

SROI has been used extensively by the Roberts Enterprise Development Fund to assess social marketing efforts and quantify social benefits that accrue to various community-development initiatives. SROI also addresses a strong criticism of social marketing posed by Rothschild (1999, p. 27), who discusses exchange: “The fundamental nature of exchange must be considered ... much of what has been called social marketing in the past has neglected the exchange ... marketing occurs when there is an attempt to transfer value from one entity’s assortment to another’s for the purpose of enhancing the assortment of the first party.”

Chapter Summary

Theories and Methods

Social marketing aims to benefit society at large. Practices and assessments are based on traditional for-profit models and on psychology, sociology, behavior modification, and education theory. Social marketing taps the behavior development and modification models described in the previous chapters. The models noted below describe the concepts that make social marketing distinct from those disciplines.

Cost–Benefit Analysis

Cost–benefit analysis proposes that programmers view behavior change from the perspective of the targeted population and develop strategies based on the perceived individual costs to and benefits for the targeted population.

Antecedents, Moderators, Consequences

According to the AMC model, giving behaviors (fundraising, volunteer recruitment, civic engagement) are influenced by controllable agents, such as an organization’s image, its actions, and communication techniques, as well as by moderating factors, such as donor variables.

Social Return on Investment

SROI is based on a discounted-cash-flow model that seeks to quantify the socioeconomic value of nonprofit work. SROI reports include qualitative descriptions of organizations and targeted program outcomes.

Social Marketing Breadth

Social marketing seeks to change individual and community behavior for the primary benefit of society. Attitudes and behavior are targeted as a means of bringing about positive social change. There are numerous examples of areas for social marketing:

- *AIDS prevention*
- *Blindness prevention*
- *Charitable fundraising*
- *Family planning*
- *Nuclear disarmament*
- *Pollution prevention*
- *Recycling efforts*
- *Reduction of chloro-fluorocarbons*
- *Smoking cessation*

Key Findings

Effective social marketing includes having an opinion leader, having information come from multiple credible sources, the possibility of intensive personal support, and communication.

Social marketing assessments show individual variables, the nature of the message, and the message provider influence charitable giving and volunteering. Specifically, the size of a request affects the number of donors and the amount given. Labeling potential donors elicits greater intentions to help and more helping behaviors. Familiar and credible sources are most likely to influence the decision to give.

Summary of Interventions

Table 5.2 highlights assessments of social marketing interventions as described in social marketing-related articles and business school case studies. The table is sorted by field.

Table 5.2. Summary of Social Marketing Interventions

Behavior Change	Findings	Intervention	Population	Methodology	Follow-Up	Reference
Employment Jobs for youth Create positive effects in lives Professional skills, training, experience Social benefits, savings	\$7.7 million savings, new tax revenue	Employment in Einstein's Café, a San Francisco restaurant	Homeless young people (age 18–24)	SROI report, personal interviews (income, housing stability, public assistance, social services, criminal justice involvement)	6-Month follow-up; annual reports	Einstein's Café, 2000
Environment Inform outcome measurement	Define mission, create microgoals	NA	NA	Interviews with nonprofit-organization senior managers	NA	Sawhill and Williamson, 2001
Health Reduce unintended pregnancies		6 States pooled Title X family-planning funds for a media campaign: print, radio, educational pamphlets	Primarily young women, 18–24, below 200% of poverty level	Case study	Unpublished data	Weinreich, 1999
Reduce youth drug use	Peer discussion, education on dangers of drug use effective; messages from adults ignored or promote rebellion	NA	Urban youth	Case study: document analysis, interviews, focus groups	Yes	Lawrence and Rangan, 1987
Reduce blindness, meet eye care needs of population	Questions about expansion and growth	Hospital staff outreach to rural areas	India	Case study: interviews, observation	NA	Rangan, 1993

Table Key: *Behavior Change*; intended effect of intervention; *Findings*; conclusions, perhaps in the form of questions developed from evaluation; *Intervention*; program type; *Population*; Targeted adopters of change; *Methods*; Data collection technique; *Follow-Up*; If and when data collection took place after the intervention ended; *Reference*; Author and date of study; NA, not applicable.

References

- Bendapudi, N., S. N. Singh, and V. Bendapudi. 1996. Enhancing helping behavior: An integrative framework for promoting planning. *Journal of Marketing* 60:33–49.
- Einstein's Café SROI report. 2000. Roberts Enterprise Development Fund. Online, accessed Aug. 24, 2002: http://www.redf.org/pub_sroi.htm#top.
- Gallagher, K., and C. B. Weinberg. 1991. Coping with success: New challenges for non-profit marketing. *Sloan Management Review* 33(1):27–42.
- Karim, S., and V. K. Rangan. 1991. Focusing the concept of social marketing. Note 9-591-047. Boston: Harvard Business School Publishing.
- Lawrence, J., and V. K. Rangan. 1987. Boston fights drugs (A): Designing communications. Case study 9-588-031. Boston: Harvard Business School Publishing.
- Rangan, V. K. 1993. The Aravind Eye Hospital, Madurai, India: In service for sight. Case study 9-593-098. Boston: Harvard Business School Publishing.
- Rangan, V. K., S. Karim, and S. K. Sandberg. 1996. Do better at doing good. *Harvard Business Review* (repr. 96308 May–June).
- Rangan, V. K., and J. D. Kramer. 1997. The Advertising Council. Case study 9-593-059. Boston: Harvard Business School Publishing.
- Rothschild, M. L. 1999. Carrots, sticks, and promises: A conceptual framework for the management of public health and social behaviors. *Journal of Marketing* 63:24–37.
- Sawhill, J., and D. Williamson. 2001. Measuring what matters in nonprofits. *McKinsey Quarterly* 2:98–107. Online, accessed Aug. 24, 2002: http://www.mckinseyquarterly.com/article_page.asp?ar=1053&L2=33&L3=95.
- SROI methodology: Analyzing the value of social purpose enterprise within a social return on investment framework. 2000. Roberts Enterprise Development Fund. Online, accessed Aug. 24, 2002: http://redf.org/pub_sroi.htm#methodology.
- Weinreich, N. K. 1999. Hands-on social marketing: A step-by-step guide. “The don't kid yourself' campaign case study.” Online excerpt, accessed Aug. 30, 2002: <http://www.social-marketing.com/dky.html>.

Resources**Agency for Healthcare Research and Quality**

http://www.ahrq.gov/chip/content/outreach_enrollment/outreach7.htm

Basic elements of social marketing: evaluation strategies, mainly as applied to state health insurance programs for children.

Weinreich Communications *<http://www.social-marketing.com/>*

Social Marketing Research and Evaluation defines social marketing and social marketing assessment. The links section provides an extensive list of online resources including evaluation guides and sources for statistics of interest to social marketers.

Andreasen, A. R. 1995. Marketing social change. San Francisco: Jossey-Bass.

Chapter 6

Lessons Learned

In This Chapter

Effective Behavior

Change Strategies

- Target behaviors
- Know your audience: tailor interventions
- Address affective characteristics

Evaluation Highlights

- Create and communicate goals
- Use multiple methods
- Systematize and disseminate

Chapter Summary

This chapter reviews and emphasizes the effective behavior change strategies and evaluation methods discussed in prior chapters. Effective practices in single or multiple fields provide support for better understanding promising approaches to changing behavior. The findings stem from the literature reviewed for this report and from the authors' experience working with nonprofit organizations. The findings are significant to the Coevolution Institute and likeminded organizations.

This final chapter of *Measuring Success* is presented in two core sections: effective behavior changes strategies and evaluation highlights. This distinction, however, is merely for organizational purposes. Clearly, evaluation is beneficial at all stages of program development and growth. In fact, as noted in the museum and health chapters, evaluation can be, and we would argue should be, integrated from the start.

Integrating evaluation from program conception can take the form of reviewing literature specific to the field, interviewing experts, and surveying potential visitors or participants. Participatory efforts and facilitating in-house work has been a growing trend in the evaluation field to support an ongoing commitment to assessment efforts.

Table 6.1 summarizes the highlights from Chapters 2 through 5.

Evaluation is not a scorecard; it is a partner in planning for and achieving success.

There is always something more to learn.

(Laurie Adams,
Coevolution Institute)

Table 6.1. Summary of Theory and Method Highlights by Field

Field	Theoretical Base and Intervention Characteristic	Evaluation Methods
Environmental education	<ul style="list-style-type: none"> • Theoretical shift away from the traditional knowledge–attitude–behavior (KAB) model toward emphasis on behavioral theories emphasizing intention and skills training • Weak link between theory and practice • Positive influence on knowledge gain through hands-on activities, pre- and post-program activities • Mixed results on the impact of attitude modification and subsequent behavior • Intrinsic motivation, intention, knowledge of action strategies, and prompts are likely to lead to adoption of targeted behaviors 	<ul style="list-style-type: none"> • Traditional emphasis on quantitative tools changing to incorporate multiple-method approach • Dominated by one-time individual program studies • Challenge of valid tools (measuring behavior modification) • Few longitudinal studies • Reliance on self-reported data • No systematic guidelines • Few longitudinal studies
Museums	<ul style="list-style-type: none"> • Learning influenced by multiple variables: individual, interaction with exhibit, exhibit, and social interaction • Time spent with an exhibit promotes knowledge gain • Interactive exhibits increase exhibit attracting and holding power, encourage social interaction, increase knowledge gain • Visitor variables—age, education, agenda—influence visit outcomes • Mixed results on gender, prior knowledge, attitudes on learning • Families tend to spend 1 to 1.5 hours per visit, follow predictable behavior patterns 	<ul style="list-style-type: none"> • Qualitative and quantitative base • Dominated by timing and tracking studies and observation; includes measuring the amount of time visitors spend with an exhibit (holding power), monitoring movements and where and for how long they stop (attracting power), and visitor interaction (both with the exhibit and with one another) • Few longitudinal studies • Challenge of measuring long-term outcomes
Health programs	<ul style="list-style-type: none"> • Strong link between theory and practice: Social learning theory states that people learn behavior by observing the results of others' actions. Behavior, therefore, is strongly influenced by the positive and negative consequences we observe and subsequently model. The health belief model asserts that people must believe the value of changing their behavior to change. One must believe in the negative consequences of not changing to be motivated to act. • Specific individual behaviors targeted and taught • Tailoring to participants is vital 	<ul style="list-style-type: none"> • Experiments and field studies are pervasive • Case studies, interviews, document analysis also widely used • Follow-up data collection and longitudinal studies common • Adoption of systematic processes (e.g., CDC model)

Table 6.1. Summary of Theory and Method Highlights by Field *Continued*

Field	Theoretical Base and Intervention Characteristic	Evaluation Methods
Social marketing	<ul style="list-style-type: none"> • Attention to cost–benefit analysis; targeted, cost-effective strategies • Consumer marketing terminology (product, targeted behavior change; pricing, individual cost or effort of change; promotion, communication of social benefits of change; placement, where the message about change is disseminated) • Program strategies include advertising, informational entertainment, education, and media campaigns 	<ul style="list-style-type: none"> • Case studies are common and serve to document efforts as well as disseminate findings • Social return on investment (SROI) framework quantifies value of programs with respect to social savings and gains and includes qualitative descriptions and analyses • Built-in metrics such as the inclusion of toll-free telephone numbers

Effective Behavior Change Strategies

Effective behavior change practices include targeting specific actions, tailoring interventions, building self-efficacy, and using prompts or tools that trigger or make behavior change easier. Interventions that develop skills through hands-on activities, prompt issue recognition and understanding, and encourage action steps are most likely to lead to responsible environmental behavior—or to any targeted behavior.

Target Behaviors

Behavior analysts claim it is usually more cost-effective to target behaviors directly than to expect behavior change as a result of increased issue awareness or general information. As acknowledged by social learning theory and in the health belief model (see Chapter 4), when the results of performing targeted behaviors are seen as positive and the consequences of ignoring them as destructive (to oneself, to others, or to the environment), people are motivated to act. Furthermore, the model of responsible environmental behavior, which is based on the results from a meta-analysis of empirical studies, expresses the role of cognitive skills (action knowledge, competence) in leading to behavior change.

Change Behavior

- *Target behaviors*
- *“Audience-centered” approach (know your audience, participants, or visitors, use multiple strategies to reach individuals in diverse groups)*
- *Target affective and cognitive factors: knowledge of issues, skills and confidence to act, belief that actions are worthwhile*

A clear emphasis on specific skills and behavior, before now largely ignored in environmental education, has proven effective in domains such as health and social marketing. Such interventions demonstrate the specific targeted behavior through modeling or factual information and sometimes provide opportunities for participants to practice their behavioral skills through role-playing. The core implications for educators focused on fostering environmentally responsible citizens are building learning experiences that provide participants with concrete skills, information related to responsible decision making, and content on environmental issues.

A wide range of literature related to the impact of duration was not reviewed for this report. However, based on at least one environmental education program evaluation, positive findings on field trip novelty reduction and follow-up curriculum for museum visits, *and* the structure of health and social marketing interventions (multiple-day, long-term), another implication from this review is that programs should strive to reinforce their message. For environmental educators, reinforcement can take the form of integrated curricula for schools, partnerships with community groups, repeat visits, longer programs, and follow-up contact. Assessment is important for informing decisions about adapting practices and for determining which techniques are effective with specific audiences.

Know Your Audience: Tailor Interventions

Tailoring a program to its audience seems critical, given the cross-field findings that individual characteristics and agendas influence learning and experiences in the same setting. Therefore, using multiple strategies could have the most impact in promoting behavior change among a diverse group (the term “diverse” is used to mean differences—including age, ethnicity, geography, or prior knowledge, for example).

Museum studies support the use of labels, objects, and interactive elements in exhibits more than less-interactive (no textual references, sound) exhibits in promoting attracting and holding power. Likewise, successful programs to promote healthier living among young people (anti-tobacco use and pregnancy prevention) combine multiple methods—one strategy by itself is not as effective as reinforcing and multiple approaches.

Environmental education programs, for example, could be most influential in changing individual behavior if they consider participants' individual characteristics: agendas, culture, day-to-day habits, home community, prior knowledge, and experience. Interventions that address and focus on participants as much as on behavior and the setting seem more likely to increase the salience of responsible action messages.

Address Affective Characteristics

Finally, people need to believe in their own abilities to engage in action. From the social learning theory, the health belief model, and the model of responsible environmental behavior, we see that attitudes and feelings are significant in behavior adoption. In fact, some evaluators believe that intention to act can be used as a proxy for future behavior.

The practical implications are that educators can work to instill positive attitudes toward taking specific actions. People are likely to plan to take action if they believe in the value of taking action, know how to act, and believe it is important for them to do so, and as indicated above, have had these intentions reinforced by successful opportunities to practice that action or skill.

Evaluation Highlights

Create and Communicate Goals

The creation of specific targeted outcomes assists in defining program mission and in constructing valid measurement tools. It is important to identify outcomes that are measurable and to consider how they contribute to an organization's larger goal or mission.

Health and social marketing programs are especially focused on assessing effectiveness in relation to the adoption of specific practices and increased knowledge. In addition, the collection of baseline and follow-up data provides valuable information on program impact. While difficult, attention to follow-up studies is paramount to capturing information on participants' knowledge gain and actions subsequent to the intervention.

Use Multiple Methods

Using qualitative and quantitative methods provides more depth and greater understanding of complex issues. Multiple-method evaluation and authentic assessment or “built-in” metrics are promising strategies that are gaining increasing application across fields. Using mixed methods together provides a deeper understanding of a program and its outcomes, and it offers a credible way to triangulate results.

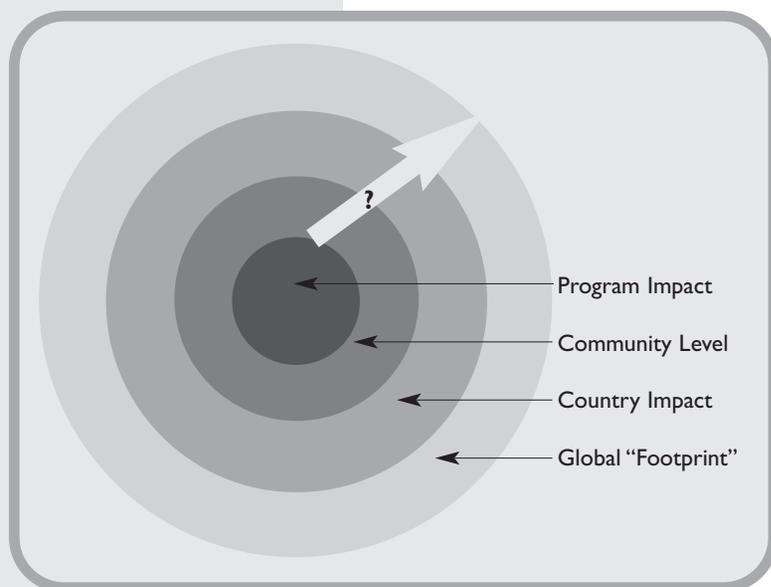


Figure 6.1. Thinking about change in a broader context.

Authentic and built-in assessments (structuring evaluation into a program) hold great promise. Largely unused in informal learning environments, authentic assessment involves projects, activities, or program components designed to enhance learning while providing program staff or external evaluators with information on participants’ knowledge and growth. Examples include projects that require application of knowledge and behavioral skills, such as using data and critical thinking skills to form arguments and engage in debate with other students, identifying local issues within a community and developing potential actions to address them, or demonstrating advocated skills (recycling, issue investigation, gardening projects). The challenges of using authentic assessment should not be underestimated—it takes time and consideration to structure valid tools and projects to identify and illuminate participant learning and potential adoption of new practices. Similarly, built-in program evaluation can include such tools as comment cards or logbooks set near museum exhibits and reflection assignments in which participants consider their program experiences.

Systematize and Disseminate

Emphasis on the interventions themselves and on evaluations of specific interventions, as is common practice across fields, ignores the larger, more complex context in which programs operate and in which we all live. The information most often presented in

Commit to Improvement and Measuring Success

- *Clarify and communicate goals*
- *Develop targeted and measurable outcomes related to goals*
- *Understand the program (based on audience, research findings specific to program area)*
- *Assess cost-effective strategies for collecting information useful and indicative of the program and program outcomes (online resources, external consultants)*

evaluation emphasizes the individual affecting society by changing behavior, one person or community at a time. Beyond the development and evaluation of individual programs and organizations, there is a broader context: What is the overarching goal, the overall effect, and how can this be measured across interventions?

Part of the answer to this question lies in systemizing evaluation strategies across the field. To better understand what works, there needs to be some standard guide—through indicators or tools or both—that resonates with the larger goals of practitioners. Systematization should demonstrate the overall successes and failures of a field at large. We need some basis (baseline) for understanding where we are going as a field. Funding agencies are working to shape evaluations across a field to provide better understanding of what works (the Centers for Disease Control model, for example).

Dissemination (or communication of strategies and findings) is crucial to achieving greater knowledge. It is the responsibility of program providers to share with others what is working, *and*—just as important—what is not working, within their organizations. Collaboration and information exchange between organizations is an important step in learning what is effective, as is having valid and useful evaluations across the field.

Systemization and dissemination works to bridge the gap between evaluation and practice, and is one especially important area for granting agencies. Funding groups can support and influence the practice and sharing of information across organizations. The work of United Way and the Roberts Enterprise Development Fund are models for how this might look: Each group researched and developed metrics to guide evaluation. The United Way's logic model and its emphasis on outcomes and the Roberts social return on investment model provide systematic frameworks to guide evaluation within the domain funded by the respective funding agencies. Both organizations provide support for funded recipients and communicate their findings.

Chapter Summary

Common challenges across fields suggest that several areas require further development and exploration. Anyone who designs or evaluates programs or interventions should consider the need to target specific behaviors as opposed to more general areas of influencing attitudes as a way to promote behavior change. This task is more feasible if another lesson learned from evaluation is adopted: Interventions that tailor their messages and operations to participants are likely to be more successful than are general, one-size-fits-all approaches.

The social science nature of evaluation and the focus on human behavior have made for a lack of systematic analysis, which is attributed at least in part to the necessary reliance on self-reported data. Tracking people's adoption of positive behavior or retention of what they have learned is easier in some cases than it is in others. Follow-up is inconsistent and longitudinal analyses are rare enough to be notable in some fields. Nevertheless, follow-up data collection is a useful and often worth the effort and cost.

What emerges from this assessment of the assessments, then, is that to do the job effectively, program designers, managers, funding organizations, and evaluators should keep a multifaceted approach in mind and consider evaluation throughout the stages of program development and growth. Furthermore, there needs to be some part of evaluation—through indicators or tools or both—that resonates with the larger goals of practitioners within a domain that demonstrates the overall success or failure of a field at large.

To accomplish the Coevolution Institute's mission, we must promote behavior that protects biodiversity. Our interest in evaluation stemmed from a desire to be as effective as possible. *Measuring Results* has shown the importance of designing programs that target a specific behavior within a particular group. It also has illustrated that effective assessment calls for systematic, multifaceted, longitudinal methodology built in to the program concept and budget. It is expensive to measure outcomes, and to do so requires solid financial support, but thorough assessment informs a program and the field. An effective blueprint for change, or, in the case of the Coevolution Institute's work, for the active protection of our planet's resources, deserves a strategy that is guided and assessed as effectively as it is executed.

CoE Leadership**Board of Directors**

- **Paul J. Growald**, Founder and Chair, Coevolution Institute; Former Chair, California League of Conservation Voters
- **Laurie Davies Adams**, Executive Director, Coevolution Institute; Former Director, Shaklee Corp.
- **Peter Carpenter**, Former President, Alza Development Corporation
- **William F. Hutton**, Esq., Professor, University of California, Hastings School of Law
- **Herbert E. Lembcke**, Vice President, Sony Development Corporation
- **Frederic J. Prager**, Managing Director, Prager, Sealy & Co.
- **Martin Rosen**, Former President and CEO, Trust for Public Land
- **Bradley A. Stirn**, Former President, U.S. Operations, Executive Search Division, TMP Worldwide

Science Advisors

- **Paul Ehrlich**, Ph.D., Stanford University
- **Thomas Eisner**, Ph.D., Cornell University
- **Adrian Forsyth**, Ph.D., Smithsonian Institution
- **Daniel Janzen**, Ph.D., University of Pennsylvania
- **Donald Kennedy**, Ph.D., Stanford University
- **Peter Raven**, Ph.D., Missouri Botanical Garden
- **E. O. Wilson**, Ph.D., Harvard University

National Advisors

- **Yvon Chouinard**, Founder and Former CEO, Patagonia, Inc.
- **Joan Ganz Cooney**, Founder, Children's Television Workshop (creator of Sesame Street)
- **William Conway**, Former General Director, Wildlife Conservation Society/Bronx Zoo
- **Anne Ehrlich**, Ph.D., Associate Director, Center for Conservation Biology, Stanford University
- **Richard Goldman**, Founder, Goldman Environmental Prize
- **Warren Hellman**, Principal, Hellman & Friedman
- **Ira Michael Heyman**, Ph.D., Former Secretary of the Smithsonian Institution; Former Chancellor, University of California

- **Scott McVay**, President, Chautauqua Institution
- **Will Rogers**, President, Trust for Public Land
- **Peter Seligmann**, Chair and CEO, Conservation International
- **Martha Stewart**, Chair and CEO, Martha Stewart Living Omnimedia
- **Jerry Tone**, Former National Board Chair, Trust for Public Land
- **Chuck Williams**, Founder and Vice Chair, Williams-Sonoma, Inc.

Measuring Results

Response Form

CoE is dedicated to self-assessment. We would like to hear from you after you have read *Measuring Results*. To access the *Measuring Results* Executive Summary document, contact info@coevolution.org.

Your Information*

Date: _____

Name: _____

Field: _____

Institution/Professional Title: _____

E-Mail: _____

Address: _____

City: _____

State, ZIP: _____

I would like to be notified of other projects of this kind.

*Your name and contact information will *not* be distributed.

We request it solely for the Coevolution Institute's in-house assessment of *Measuring Results*.

How did you learn about *Measuring Results*? Please select all that apply:

- Article CoE web site Colleague
 Conference Received copy from CoE Other: _____

Please rate the following:

	1 (Low)	2	3	4	5 (High)
Relevance of evaluation to your work	1	2	3	4	5
Your knowledge of evaluation	1	2	3	4	5
Your interest in evaluation	1	2	3	4	5

Please rate the following:

1 = strongly agree, 2 = agree, 3 = neutral (neither agree nor disagree), 4 = disagree, 5 = strongly disagree.

	1 (Strongly disagree)	2	3	4	5 (Strongly agree)
<i>Measuring Results</i> enhanced my understanding of evaluation.	1	2	3	4	5
Evaluation is relevant to my work	1	2	3	4	5
I want to know more about evaluation methods	1	2	3	4	5
I want to learn more about evaluation findings	1	2	3	4	5
I want to know more about CoE's work	1	2	3	4	5

Please check. I read the following:

- Whole document
 Ch. 1: Introduction Ch. 2: Environmental Education Ch. 3: Museums
 Ch. 4: Health Ch. 5: Social Marketing Ch. 6: Lessons Learned

Thank you for your time and response. Please mail or fax this form to:

The Coevolution Institute

423 Washington Street, 4th Floor
San Francisco, CA 94111
415-362-1137 tel
415-362-3070 fax