
OTHER RECOMMENDED READING

- Boykin, L.L. (1957). Let's Eliminate the Confusion: What is Evaluation? *Educational Administration and Supervision*, 43 (2): 115-121.
- Debus, M. (1995). *Methodological Review: A Handbook for Excellence in Focus Group Research*. Washington, DC: Academy for Educational Development.
- Denzin, N.K., and Lincoln, Y.S. (eds.). (1994). *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage.
- Erlandson, D.A., Harris, E.L., Skipper, B.L., and Allen, D. (1993). *Doing Naturalist Inquiry: A Guide to Methods*. Newbury Park, CA: Sage.
- Fox, S. (2000). An Effective School Evaluation and Training Program. In *The Cultural Context of Educational Evaluation: The Role of Minority Evaluation Professionals*, NSF 01-43. Arlington, VA: National Science Foundation, Directorate for Education and Human Resources.
- Frierson, H.T. (2000). The Need for the Participation of Minority Professionals in Educational Evaluation. In *The Cultural Context of Educational Evaluation: The Role of Minority Evaluation Professionals*, NSF 01-43. Arlington, VA: National Science Foundation, Directorate for Education and Human Resources.
- Greenbaum, T.L. (1993). *The Handbook of Focus Group Research*. New York: Lexington Books.
- Hart, D. (1994). *Authentic Assessment: A Handbook for Educators*. Menlo Park, CA: Addison-Wesley.
- Herman, J.L., and Winters, L. (1992). *Tracking Your School's Success: A Guide to Sensible Evaluation*. Newbury Park, CA: Corwin Press.
- Hood, S. (forthcoming). Nobody Knows My Name: In Praise of African American Evaluators Who Were Responsive. In *Responsive Evaluation: Roots and Wings*, edited by J. Greene and T. Abma. New Directions for Program Evaluation. San Francisco, CA: Jossey-Bass.
- Hood, S. (2000). A New Look at an Old Question. In *The Cultural Context of Educational Evaluation: The Role of Minority Evaluation Professionals*, NSF 01-43. Arlington, VA: National Science Foundation, Directorate for Education and Human Resources.

-
-
- Hughes, G. (2000). Evaluation of Educational Achievement of Underrepresented Minorities: Assessing Correlates of Student Academic Achievement. In *The Cultural Context of Educational Evaluation: The Role of Minority Evaluation Professionals*, NSF 01-43. Arlington, VA: National Science Foundation, Directorate for Education and Human Resources.
- Hymes, D.L., Chafin, A.E., and Gondor, R. (1991). *The Changing Face of Testing and Assessment: Problems and Solutions*. Arlington, VA: American Association of School Administrators.
- Krueger, R.A. (1988). *Focus Groups: A Practical Guide for Applied Research*. Newbury Park, CA: Sage.
- LeCompte, M.D., Millroy, W.L., and Preissle, J. (eds.). (1992). *The Handbook of Qualitative Research in Education*. San Diego, CA: Academic Press.
- Merton, R.K., Fiske, M., and Kendall, P.L. (1990). *The Focused Interview: A Manual of Problems and Procedures*, 2nd Ed. New York: The Free Press.
- Miles, M.B., and Huberman, A.M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks, CA: Sage.
- Morgan, D.L. (ed.). (1993). *Successful Focus Groups: Advancing the State of the Art*. Newbury Park, CA: Sage.
- Morse, J.M. (ed.). (1994). *Critical Issues in Qualitative Research Methods*. Thousand Oaks, CA: Sage.
- National Science Foundation. (2001). *The Cultural Context of Educational Evaluations: The Role of Minority Evaluation Professionals*. Workshop Proceedings. June 1-2, 2000.
- Perrone, V. (ed.). (1991). *Expanding Student Assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Reich, R.B. (1991). *The Work of Nations*. New York: Alfred A. Knopf.
- Rodriguez, C. (2000). Assessing Underrepresented Science and Mathematics Students: Issues and Myths. In *The Cultural Context of Educational Evaluation: The Role of Minority Evaluation Professionals*, NSF 01-43. Arlington, VA: National Science Foundation, Directorate for Education and Human Resources.
- Sanders, J.R. (2000). *Evaluating School Programs*. Second Ed. Thousand Oaks, CA: Corwin Press.
- Schatzman, L., and Strauss, A.L. (1973). *Field Research*. Englewood Cliffs, NJ: Prentice-Hall.

-
- Seidman, I.E. (1991). *Interviewing as Qualitative Research: A Guide for Researchers in Education and Social Sciences*. New York: Teachers College Press.
- Smith, M.L. (1986). The Whole is Greater: Combining Qualitative and Quantitative Approaches in Evaluation Studies. In *Naturalistic Evaluation*, edited by Dave Williams. New Directions for Program Evaluation, Vol. 30. San Francisco, CA: Jossey-Bass, Inc.
- Stake, R. (1972). *Program Evaluation, Particularly Responsive Evaluation*. ERIC Document ED 075-187. [Available online.] <http://www.wmich.edu/evalctr/pubs/ops/>
- Stewart, D.W., and Shamdasani, P.N. (1990). *Focus Groups: Theory and Practice*. Newbury Park, CA: Sage.
- The Joint Committee on Standards for Educational Evaluation. (1994). *The Program Evaluation Standards*. Second Ed. Thousand Oaks, CA: Sage Publications.
- U.S. General Accounting Office (GAO). (1990). *Case Study Evaluations*. Paper 10.1.9. Washington, DC: GAO.
- Weiss, R.S. (1994). *Learning From Strangers: The Art and Method of Qualitative Interview Studies*. New York: Free Press.
- Wiggins, G. (1989). A True Test: Toward More Authentic and Equitable Assessment. *Phi Delta Kappan*, May, 703-704.
- Wiggins, G. (1989). Teaching to the (Authentic) Test. *Educational Leadership*, 46, 45.
- Yin, R.K. (1989). *Case Study Research: Design and Method*. Newbury Park, CA: Sage.

GLOSSARY

Accuracy: The extent to which an evaluation is truthful or valid in what it says about a program, project, or material.

Achievement: Performance as determined by some type of assessment or testing.

Affective: Consists of emotions, feelings, and attitudes.

Anonymity (provision for): Evaluator action to ensure that the identity of subjects cannot be ascertained during the course of a study, in study reports, or in any other way.

Assessment: Often used as a synonym for evaluation. The term is sometimes recommended for restriction to processes that are focused on quantitative and/or testing approaches.

Attitude: A person's opinion about another person, thing, or state.

Attrition: Loss of subjects from the defined sample during the course of data collection.

Audience(s): Consumers of the evaluation; those who will or should read or hear of the evaluation, either during or at the end of the evaluation process. Includes those persons who will be guided by the evaluation in making decisions and all others who have a stake in the evaluation (see stakeholders).

Authentic assessment: Alternative to traditional testing that focuses on student skill in carrying out real-world tasks.

Background: Information that describes the project, including its goals, objectives, context, and stakeholders.

Baseline: Facts about the condition or performance of subjects prior to treatment or intervention.

Behavioral objectives: Measurable changes in behavior that are targeted by a project.

Bias: A point of view that inhibits objectivity.

Case study: An intensive, detailed description and analysis of a single project, program, or instructional material in the context of its environment.

Categorical scale: A scale that distinguishes among individuals by putting them into a limited number of groups or categories.

Checklist approach: The principal instrument for practical evaluation, especially for investigating the thoroughness of implementation.

-
-
- Client:** The person or group or agency that commissioned the evaluation.
- Coding:** To translate a given set of data or items into descriptive or analytic categories to be used for data labeling and retrieval.
- Cohort:** A term used to designate one group among many in a study. For example, “the first cohort” may be the first group to have participated in a training program.
- Component:** A physically or temporally discrete part of a whole. It is any segment that can be combined with others to make a whole.
- Conceptual scheme:** A set of concepts that generate hypotheses and simplify description, through the classification and categorization of phenomena, and the identification of relationships among them.
- Conclusions (of an evaluation):** Final judgments and recommendations.
- Content analysis:** A process using a parsimonious classification system to determine the characteristics of a body of material or practices.
- Context (of an evaluation):** The combination of factors accompanying the study that may have influenced its results, including geographic location, timing, political and social climate, economic conditions, and other relevant professional activities in progress at the same time.
- Continuous scale:** A scale containing a large, perhaps infinite, number of intervals. Units on a continuous scale do not have a minimum size but rather can be broken down into smaller and smaller parts. For example, grade point average (GPA) is measured on a continuous scale, a student can have a GPA of 3, 3.5, 3.51, etc. (See categorical scale.)
- Criterion, criteria:** A criterion (variable) is whatever is used to measure a successful or unsuccessful outcome, e.g., grade point average.
- Criterion-referenced test:** Test whose scores are interpreted by referral to well-defined domains of content or behaviors, rather than by referral to the performance of some comparable group of people.
- Cross-case analysis:** Grouping data from different persons to common questions or analyzing different perspectives on issues under study.
- Cross-sectional study:** A cross-section is a random sample of a population, and a cross-sectional study examines this sample at one point in time. Successive cross-sectional studies can be used as a substitute for a longitudinal study. For example, examining today’s first year students and today’s graduating seniors may enable the evaluator to infer that the college experience has

produced or can be expected to accompany the difference between them. The cross-sectional study substitutes today's seniors for a population that cannot be studied until 4 years later.

Data display: A compact form of organizing the available information (for example, graphs, charts, matrices).

Data reduction: Process of selecting, focusing, simplifying, abstracting, and transforming data collected into written field notes or transcriptions.

Delivery system: The link between the product or service and the immediate consumer (the recipient population).

Descriptive data: Information and findings expressed in words, unlike statistical data, which are expressed in numbers.

Design: The process of stipulating the investigatory procedures to be followed in doing a specific evaluation.

Dissemination: The process of communicating information to specific audiences for the purpose of extending knowledge and, in some cases, with a view to modifying policies and practices.

Document: Any written or recorded material not specifically prepared for the evaluation.

Effectiveness: Refers to the worth of a project in achieving formative or summative objectives. "Success" is its rough equivalent.

Elite interviewers: Well-qualified and especially trained persons who can successfully interact with high-level interviewees and are knowledgeable about the issues included in the evaluation.

Ethnography: Descriptive anthropology. Ethnographic program evaluation methods often focus on a program's culture.

Executive summary: A nontechnical summary statement designed to provide a quick overview of the full-length report on which it is based.

External evaluation: Evaluation conducted by an evaluator outside the organization within which the project is housed.

Field notes: Observer's detailed description of what has been observed.

Focus group: A group selected for its relevance to an evaluation that is engaged by a trained facilitator in a series of discussions designed for sharing insights, ideas, and observations on a topic of concern to the evaluation.

Formative evaluation: Evaluation designed and used to improve an intervention, especially when it is still being developed.

Goal: A broad-based description of an intended outcome.

Hypothesis testing: The standard model of the classical approach to scientific research in which a hypothesis is formulated before the experiment to test its truth.

Impact evaluation: An evaluation focused on outcomes or payoff of a project.

Implementation evaluation: Assessing program delivery (a subset of formative evaluation).

Indepth interview: A guided conversation between a skilled interviewer and an interviewee that seeks to maximize opportunities for the expression of a respondent's feelings and ideas through the use of open-ended questions and a loosely structured interview guide.

Informed consent: Agreement by the participants in an evaluation to the use, in specified ways for stated purposes, of their names and/or confidential information they supplied.

Instrument: An assessment device (test, questionnaire, protocol, etc.) adopted, adapted, or constructed for the purpose of the evaluation.

Internal evaluator: A staff member or unit from the organization within which the project is housed.

Inter-rater reliability: A measure of the extent to which different raters score an event or response in the same way.

Intervention: Project feature or innovation subject to evaluation.

Intra-case analysis: Writing a case study for each person or unit studied.

Key informant: Person with background, knowledge, or special skills relevant to topics examined by the evaluation.

Longitudinal study: An investigation or study in which a particular individual or group of individuals is followed over a substantial period of time to discover changes that may be attributable to the influence of the treatment, or to maturation, or the environment. (See also cross-sectional study.)

Matrix: An arrangement of rows and columns used to display multi-dimensional information.

Measurement: Determination of the magnitude of a quantity.

Meta-evaluation: Evaluation of the merit of the evaluation itself.

Mixed-method evaluation: An evaluation for which the design includes the use of both quantitative and qualitative methods for data collection and data analysis.

Moderator: Focus group leader; often called a facilitator.

Nonparticipant observer: A person whose role is clearly defined to project participants and project personnel as an outside observer or onlooker.

Norm-referenced tests: Tests that measure the relative performance of the individual or group by comparison with the performance of other individuals or groups taking the same test.

Objective: A specific description of an intended outcome.

Observation: The process of direct sensory inspection involving trained observers.

Ordered data: Nonnumeric data in ordered categories (for example, students' performance categorized as excellent, good, adequate, and poor).

Outcome: Post-treatment or post-intervention effects.

Paradigm: A general conception, model, or "worldview" that may be influential in shaping the development of a discipline or subdiscipline (for example, "the classical, positivist social science paradigm in evaluation").

Participants: Those individuals who are directly involved in a project.

Participant observer: An evaluator who participates in the project (as participant or staff) in order to gain a fuller understanding of the setting and issues.

Performance evaluation: A method of assessing what skills students or other project participants have acquired by examining how they accomplish complex tasks or the quality of the products they have created (e.g., poetry, artwork).

Population: All persons in a particular group.

Prompt: Reminder used by interviewers to obtain complete answers.

Purposive sampling: Creating samples by selecting information-rich cases from which one can learn a great deal about issues of central importance to the purpose of the evaluation.

Qualitative evaluation: The approach to evaluation that is primarily descriptive and interpretative.

Quantitative evaluation: The approach to evaluation involving the use of numerical measurement and data analysis based on statistical methods.

Random sampling: Drawing a number of items of any sort from a larger group or population so that every individual item has a specified probability of being chosen.

Recommendations: Suggestions for specific actions derived from evidence-based conclusions.

Sample: A part of a population.

Secondary data analysis: A reanalysis of data using the same or other appropriate procedures to verify the accuracy of the results of the initial analysis or for answering different questions.

Self-administered instrument: A questionnaire or report completed by a study participant without the assistance of an interviewer.

Stakeholder: One who has credibility, power, or other capital invested in a project and thus can be held to be to some degree at risk with it.

Standardized tests: Tests that have standardized instructions for administration, use, scoring, and interpretation with standard printed forms and content. They are usually norm-referenced tests but can also be criterion referenced.

Strategy: A systematic plan of action to reach predefined goals.

Structured interview: An interview in which the interviewer asks questions from a detailed guide that contains the questions to be asked and the specific areas for probing.

Summary: A short restatement of the main points of a report.

Summative evaluation: Evaluation designed to present conclusions about the merit or worth of an intervention and recommendations about whether it should be retained, altered, or eliminated.

Transportable: An intervention that can be replicated in a different site.

Triangulation: In an evaluation, an attempt to get corroboration on a phenomenon or measurement by approaching it by several (three or more) independent routes. This effort provides confirmatory measurement.

Utility: The extent to which an evaluation produces and disseminates reports that inform relevant audiences and have beneficial impact on their work.

Utilization of (evaluations): Use and impact are terms used as substitutes for utilization. Sometimes seen as the equivalent of implementation, but this applies only to evaluations that contain recommendations.

Validity: The soundness of the inferences made from the results of a data-gathering process.

Verification: Revisiting the data as many times as necessary to cross-check or confirm the conclusions that were drawn.

Appendix A

Finding An Evaluator

There are many different sources for locating a project evaluator. The one that works best will depend on a number of factors including the home institution for the project, the nature of the project, and whether or not the principal investigator has some strong feeling about the type(s) of evaluation that are appropriate.

There are at least three avenues that can be pursued:

- If the project is being carried out at or near a college or university, a good starting point is likely to be at the college or university itself. Principal investigators can contact the department chairs from areas such as education, psychology, administration, or sociology and ask about the availability of staff skilled in project evaluation. In most cases, a few calls will yield several names.
- A second source for evaluation assistance comes from independent contractors. There are many highly trained personnel whose major income derives from providing evaluation services. Department chairs may well be cognizant of these individuals and requests to chairs for help might include suggestions for individuals they have worked with outside of the college or university. In addition, independent consultants can be identified from the phone book, from vendor lists kept by procurement offices in state departments of education and in local school systems, and even from resource databases kept by some private foundations, such as the Kellogg Foundation in Michigan.
- Finally, suggestions for evaluators can be obtained from calls to other researchers or perusal of research and evaluation reports. Western Michigan University also has a list of evaluators in their web site at www.wmich.edu/evalatr. A strong personal recommendation and a discussion of an evaluator's strengths and weaknesses from someone who has worked with a specific evaluator is very useful when starting a new evaluation effort.

Although it may take a chain of telephone calls to get the list started, most principal investigators will ultimately find that they have several different sources of evaluation support from which to select. The critical task then becomes negotiating time, content, and, of course, money.



The National Science Foundation promotes and advances scientific progress in the United States by competitively awarding grants and cooperative agreements for research and education in the sciences, mathematics, and engineering.

To get the latest information about program deadlines, to download copies of NSF publications, and to access abstracts of awards, visit the NSF Web Site at:

<http://www.nsf.gov>

- **Location:** 4201 Wilson Blvd.
Arlington, VA 22230
 - **For General Information (NSF Information Center):** (703) 292-5111
 - **TDD (for the hearing-impaired):** (703) 292-5090
 - **To Order Publications or Forms:**
- Send an e-mail to: pubs@nsf.gov
- or telephone: (301) 947-2722
- **To Locate NSF Employees:** (703) 292-5111

NATIONAL SCIENCE FOUNDATION
ARLINGTON, VA 22230

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE \$300

RETURN THIS COVER SHEET TO ROOM P35 IF YOU DO NOT WISH TO RECEIVE THIS MATERIAL , OR IF CHANGE OF ADDRESS IS NEEDED , INDICATE CHANGE INCLUDING ZIP CODE ON THE LABEL (DO NOT REMOVE LABEL).

**PRESORTED STANDARD
POSTAGE & FEES PAID
National Science Foundation
Permit No. G-69**

NSF 02-057