

Glossary

A

Accountability: Responsibility for effective and efficient performance of programs. Measures of accountability focus on (1) benefits accruing from the program as valued by customers and supporters (2) how resources are invested and the results attained.

Action Cards: Use of index cards on which participants record what they did-the “action”-and when they reached their goal; primarily used in self-assessment.

Aggregation: Assembling data into groups or classes for reports and other purposes.

Analysis Data: Purposeful ordering of data in a manner that facilitates objective interpretation with respect to a particular question, concern, problem, or objective.

Anonymity: An attempt to keep the participants unknown to the people who use the evaluation and, if possible, to the investigators themselves.

Archival Research: Research method that involves the use of data extracted from existing written or computer records.

Attribution of Effect: Concept that refers to the ability to conclude that an intervention *caused* an outcome.

Audience: That aggregate of participants who are actual or potential program clientele.

B

Benchmarks: Performance data used either as a baseline against which to compare future performance or as a marker of progress toward a goal.

Bias: Error of the estimate from the true proportion that exists in a population.

C

Case Study: An in-depth examination of a particular case—a program, group of participants, single individual, site, or location. Case studies rely on multiple sources of information and methods to provide as complete a picture as possible.

Cluster Evaluation: A type of evaluation that seeks to determine the impacts of a collection of related projects on society as a whole. Cluster evaluation looks across a group of projects to identify issues and problems that affect an entire area of a program. Designed and used by the W.K. Kellogg Foundation to determine the effectiveness of its grantsmaking.

Comparison Group: In quasi-experimental evaluation design, a group of evaluation participants that is not exposed to the intervention. This term usually implies that participants were *not* randomly assigned, but were similar to the intervention group members in many respects.

Confidence Interval: An estimated range of values derived from sample statistics with a given high probability of covering the true population value.

Confidentiality: An attempt to remove any elements that might indicate the subject's identity.

Context Evaluation: A type of evaluation that examines how the project functions within the economic, social and political environment of its community and project setting.

Contingency Table: A form of relational analysis that classifies observations by their values on two or more variables. Each cell in the table represents a unique combination of values across the variables, and each observation qualifies for one and only one cell in the table.

Control Group: In experimental evaluation design, a group of participants that is essentially similar to the intervention group but is not exposed to the intervention. Participants are designated to be part of either a control or intervention group through random assignment.

Convenience Sample: A nonrandom study sample selected on the basis of convenience and access

Context Evaluation: A type of evaluation that examines how the project functions within the economic, social and political environment of its community and project setting.

Cost-Benefit Analysis: Process to estimate the overall cost and benefit of a program or components within a program. Seeks to answer the question, “Is this program or product worth its costs?” Or, “Which of the options has the highest benefit/cost ratio?” This is only possible when all values can be converted into money terms.

Creative Expression: Use of art forms to represent people’s ideas and feelings through stories, drama, dance, music and art.

Criteria: Standards for judging; standards or norms selected as a basis for use in making quantitative or qualitative comparisons, judgments, or evaluations. Qualities or dimensions on which a program, product, or activity is to be judged.

Methods for developing criteria include:

- armchair approach—in this case the evaluator selects criteria off the top of his or her head. This may occur with little forethought and with no frame of reference.
- empirical approach—this approach is based on scientific evidence. The evaluator relies on research findings experience to assist in selecting and developing criteria.
- rational approach—this consists of a systematic analysis of the situation, or development of criteria based on the best information available. Some combination of the rational and empirical approaches is highly recommended.

Indicators and criteria can be used to evaluate the objective of a program and is demonstrated in the following example:

Objective:

The county 4-H organization is to develop a high quality 4-H program.

<u>Indicators</u>	<u>Criteria</u>
Visibility of 4-H.....	50% of sample knows when “Spotlight 4-H” is aired
Size of enrollment.....	20% of potential
Members per leader.....	averages 10 to 1
Racial make-up.....	parity (15% minority)
Strength of curriculum.....	80% of projects have 6 planned learning opportunities implemented
Target Audience (involvement)	

<u>Indicators</u>	<u>Criteria</u>
Number enrolled.....	Expect 400
Number completing.....	Expect 200
Age.....	40 and below
Income.....	Low
Race.	90% minority

D

Data: Information collected according to a methodology and through specific research methods and instruments. Types of data include:

- Hard Data—Data than can be quantified
- Soft Data—Qualitative or perceptual data
- Secondary Data—Collecting data from existing records

Data Analysis: The process of examining systematically collected information.

Decision Situation: A set of alternatives for consideration.

Descriptive Analysis: Data analysis that results in information that characterizes the sample, such as measures of central tendency (e.g., mean, median, mode) and measures of variability (e.g., range, standard deviation,

variance). In addition to describing the sample, such data may be used as input for relational analysis.

Development Evaluation: Evaluation in which the evaluator is part of a collaborative team that monitors what is happening in a program, both processes and outcomes, in an evolving, rapidly changing environment of constant feedback and change.

Diary and Journal: Recording of events over time revealing the personal perspective of the writer/recorder.

Document Analysis: Use of content analysis and other techniques to analyze and summarize printed material and existing information.

E

Effectiveness: Degree to which the program yields desired/desirable results.

Efficiency: Comparison of outcomes to costs.

Empowerment Evaluation: Empowerment evaluation is the use of evaluation concepts, techniques, and findings to foster improvement and self-determination. In empowerment evaluation, program participants maintain control of the evaluation process; outside evaluators work to build the evaluation capacity of participants and help them use evaluation findings to advocate for their program.

Evaluation: Systematic inquiry to inform decision-making and improve programs. Systematic implies that the evaluation is a thoughtful process of asking critical questions, collecting appropriate information, and then analyzing and interpreting the information for a specific use and purpose.

- Context Evaluation (situation) – Provides the rationale for determining objectives and setting priorities. It defines the relevant environment, describes the desired and actual conditions pertaining

to the environment and identifies unmet needs and unused opportunities.

- **Input Evaluation (plan of work)** – Provides information for determining how to use resources to meet program goals. It involves identifying and assessing (1) relevant capabilities for achieving program goals; (2) strategies for achieving program goals; (3) designs for implementing a strategy.

Experimental Design: A methodology for examining intervention outcomes that involves the random assignment of subjects to intervention and control conditions with a controlled manipulation delivered to subjects in the intervention group. The design enables the evaluator to conclude that the outcomes were *caused* by the intervention. See also **quasi-experimental design**.

Expert or Peer Review: Examination by a review committee, a panel of experts or peers.

Face-to-face Questionnaire: See **structured interview**.

Focus Group: Qualitative research method that involves structured discussion among individuals with shared characteristics.

Formative Evaluation: Evaluation conducted during the development and implementation of a program whose primary purpose is providing information for program improvement. Evaluation used to facilitate decisions as the program progresses. Its primary concern is program improvement.

Frequency Tables: The simplest way of summarizing data for a single variable is to determine the number of responses for each of the response categories. In doing so, we are simply taking large quantities of unmanageable data and summarizing them into a manageable, easy to understand form. This summarizing process is referred to as “frequencies,” i.e. counting and recording the numbers and kinds of responses to a specific question.

Group Assessment: Collecting evaluation information through the use of group processes such as nominal group technique, focus group, Delphi, brainstorming and community forums.

Impact: The ultimate social, economic, and/or environmental effects or consequences of the program. Impacts tend to be long-term achievements. They may be positive, negative or neutral.

Impact Evaluation: A type of evaluation that determines the net causal effects of the program beyond its immediate results. Impact evaluation often involves a comparison of what appeared after the program with what would have appeared without the program. (program action) provides feedback.

Implementation Evaluation: Evaluation activities that document the evolution of a project and provide indications of what happens within a project and why. Project directors use information to adjust current activities. Implementation evaluation requires close monitoring of program delivery.

- Product Evaluation (outcome) – Measures and interprets attainments during and at the end of the program. It's concerned with main effects, side effects, costs, and superiority.
- Summative Evaluation – End of program evaluation which is concerned with determining overall program effectiveness.
- Evidence – Any information which may contribute to the consideration of a particular issue and is based on objective criteria.

Indicator: Expression of what is/will be measured or described; evidence which signals achievements, what you wish to measure. Answers the question, "How will I know it?"

Individual-Oriented Intervention: An intervention that attempts to change the behavior of individuals by enhancing the knowledge, attitudes, skills, and beliefs of individuals.

Informed Consent: The written permission obtained from research participants (or their parents if participants are minors) giving their consent to participate in an evaluation after having been informed of the nature of the research.

Inputs: Resources that go into a program including staff time, materials, money, equipment, facilities, volunteer time.

Institutional Review Board: A group of researchers and others appointed by an institution to assess proposed data collection regarding potential harm that might be caused to study participants.

Instrument: Device that assists evaluators in collecting data in an organized fashion, such as a standardized survey or interview protocol.

Intermediate Outcome: Intervention outcome, such as changes in knowledge, attitudes, or beliefs, that occurs before, and is necessary for changes in, substance use and substance-related problems. See also **long-term outcome**.

Internal validity: Concept that refers to the ability to make inferences about whether the relationship between variables is causal in nature and, if it is, the direction of causality.

Intervention: A manipulation applied to a population in order to change behavior. In substance abuse prevention, intervention at the individual or policy level may be used to prevent or lower the rate of substance use or related problems. See also **policy intervention; individual-oriented intervention**.

Intervention Group: In experimental and quasi-experimental evaluation designs, the group of participants that is exposed to the intervention. See also **control group** and **comparison group**.

Interviews: Information collected by talking with and listening to people. Interviews range on a continuum from those which are tightly structured (as in a survey) to those that are free-flowing and conversational.

Interview Schedule: List of items or questions, together with specific instructions for the interview, to be used in gathering data by the interview method.

- Telephone Interview Schedule—the series of items to be used for gathering information from respondents by use of the telephone. Include specific instructions for the interviewer.

- Individual Interview Schedule—information collected by an interviewer from a subject on a one-to-one basis. The interview

may be structured (interview schedule) or unstructured (interviewer asks questions based on written or mental notes).

- Group Interview—Prospective respondents are assembled in a group. Each person in the group is asked to record his/her answer as questions are read to them.

Judging: It is the act of choosing among several decision alternatives, hence, the act of decision making.

Judgment: A natural operation involving comparison and discrimination. There are two important ingredients to judgment: (1) evidence which is a basis for knowledge and (2) insight.

Two types of judgment include:

- Objective—Emphasizing or expressing the nature of reality as it is apart from personal reflections or feelings, expressing or involving the use of facts without distortion by personal feelings or prejudices.
- Subjective—Relating to experience or knowledge as conditioned by personal mental characteristics or states.

Learning Experience: The interaction between the learner and conditions in the environment to which the individual reacts; the encounter through which learning takes place and educational objectives are attained.

Log: Recording of chronological entries which are usually brief and factual.

Long-Term Outcome: Intervention outcome corresponding to the prevalence of substance use and substance-related problems.

Macroscopic Level: An evaluator's point of reference focused on overall evaluation of a total system.

Maturation Effects: Changes in outcomes that are attributable to participants' growing older, wiser, stronger, more experienced, and the like, solely through the passage of time.

Mean: A measure of central tendency computed by summing over all the values of a variable and dividing by the number of cases (on average).

Means Comparison: A form of relational analysis that involves comparing the average values of two or more groups to see if they differ more than would be expected by chance.

Measure/Measurement: Representation of quantity or capacity. In the past, these terms carried a quantitative implication of precision and, in the field of education, were synonymous with testing and instrumentation. Today, the term “measure” is used broadly to include quantitative and qualitative information to understand the phenomena under investigation.

Measures of Central Tendency: Indices that describe the “typical” or “average” value—for example, the arithmetic average and the median.

Median: A measure of central tendency referring to the point exactly midway between the top and bottom halves of a distribution of values.

Method: A planned procedure, sequence of experiences, activities, or events designed to bring about a desired end.

Methodology: Procedure for collecting data.

Microscopic Level: An evaluator’s point of reference focused on detailed evaluation of specific elements in a total system.

Mixed Methods: The use of both qualitative and quantitative methods to study phenomena. These two sets of methods can be used simultaneously or at different stages of the same study.

Mode: A measure of central tendency referring to the value most often given by respondents.

Monitoring: Ongoing assessment of the extent to which a program is operating consistent with its design. Often means site visits by experts for compliance-focused reviews of program operations.

Nonrepresentative Sample: A segment of a larger body or population that does not mirror in composition characteristics of the larger body or population. See also **representative sample**.

Objective: An end or aim stated in support of a goal.

- **Measurable Objective**—a statement of program intent that can, with a degree of certainty be accurately described when attained. Measurability would include quantitative and/or qualitative descriptive units related to determining results, outcomes, or effectiveness of the objectives. To be measurable, the objective must have some change of being implemented (operationalized) and attained.
- **Operational Objective**—A statement of procedure to be undertaken; a step or unit of what is to be done.
- **Organizational Objective**—A statement of change to be accomplished in the development or maintenance of an organization; a statement of a goal, end or aim of a group, social entity, or functional structure of people.

Observation: Data collection method involving unobtrusive examination of behavior and/or occurrences, often in a natural setting, and characterized by no interaction between participants and observers.

Outcome Valuation: Evaluation that focuses research questions on assessing intervention effects on intended outcomes. See also **process evaluation**.

Outcome Monitoring: The regular or periodic reporting of program outcomes in ways that stakeholders can use to understand and judge results. Outcome monitoring exists as part of program design and provides frequent and public feedback or performance.

Outcomes: End results or effects of the program. Outcomes answer the questions, “So what?”, “What difference does the program make in people’s lives?” Outcomes may be intended and unintended; positive and negative. Outcomes fall along a continuum from short-term/immediate, to medium-term/intermediate, to final outcomes, often synonymous with impact.

Outputs: Activities, services, events, products, participating generated by a program.

Participant-Observation: Qualitative research method that requires simultaneous participation in and examination of activity in a natural setting. The identity of the evaluator as an evaluator is usually made known to others in the setting.

Participatory Evaluation: Evaluation in which the perspective of the evaluator carries no more weight than other stakeholders, including participants and the evaluation process and its results are relevant and useful to stakeholders for future actions. Participatory approaches attempt to be practical, useful and empowering to multiple stakeholders and actively engage all stakeholders in the evaluation process.

Performance Evaluation: The evaluation of a particular achievement, in the form of output or process.

Performance Measure: A particular value or characteristic used to measure/examine program quality; may be expressed in a qualitative or quantitative way.

Performance Targets: The expected result or level of achievement, often set as numeric levels of performance.

Personnel Evaluation: Evaluation of policies, plans and proposals for use by policy makers and/or communities trying to effect policy change.

Plan of Work: (written document) A written outline of strategy for one year or less for each problem or concern included in a program that sets forth in an integrated and coordinated manner the following elements: (1) educational operations and organizational objectives to be achieved; (2) learning experiences, activities, events, and/or situations to be undertaken, calendarized, and related to appropriate objectives; (3) evidence of accomplishment, and calendar for evaluation; (4) time to be devoted to each activity, event, and/or learning situation; (5) who will assume primary and support leadership responsibilities; and (6) coordinator, internal and external.

Policy Evaluation: Evaluation of policies, plans and proposals for use by policy makers and/or communities trying to effect policy change.

Policy Intervention: An intervention that attempts to change the behavior of individuals by changing the economic or regulatory environment around substance use.

Portfolio Review: A collection of materials, including samples of work, that encompass the breadth and scope of the program or activity being evaluated.

Problem Story: Narrative account of past, present, or future situations as a means of identifying perceptions. Using fictional characters externalizes the problem situation.

Process: A course of action, procedure or a series of steps leading toward an end.

Process Evaluation: Evaluation that focuses research questions on the nature of intervention implementation or its structure and operations. See also **outcome evaluation**.

Product Evaluation: The evaluation of functional artifacts.

Program: A series of planned events (or activities) with specific objectives. These planned events are designed to deliver educational information for the purpose or purposes expressed in the objective(s). Under the concept of program there are a number of “activities” conducted to reach the program objective(s). These “activities” are time-structured or sequenced to take advantage of conditions that facilitate learning or acceptance of the information being provided. Program is not considered to be synonymous with “activity.” Activities are the various components or planned events that contribute to the achievement of the program objectives.

Program Accomplishments: Intended and unintended results of program activity.

Program Development: The continuous series of processes which includes organizing, planning a program, preparing a plan-of-work and teaching plans, implementing the plans, evaluating, and reporting accomplishments.

Program Evaluation: the evaluation of a structured intervention to improve the well-being of people, groups, organizations and communities.

Qualitative Data: Generally, contextual information in evaluation studies, usually describing participants and interventions. The strength of qualitative data, which is often presented as text, is its ability to illuminate evaluation findings derived from quantitative methods. See also **quantitative data**.

Qualitative Methodology: Methods that examine phenomena in depth and detail without predetermined categories or hypotheses. Emphasis is on understanding the phenomena as it exists. Often connoted with naturalistic inquiry, inductive, social anthropological world view. Qualitative methods usually consist of three kinds of data collection: Observation, open-ended interviewing, and document review.

Quantitative Analysis: The use of statistical techniques to understand quantitative data and to identify relationships between and among variables.

Quantitative Data: Data in a numerical form that measures that in evaluation studies capture changes in targeted outcomes (e.g., substance use) and intervening variables (e.g., attitudes toward use). The strength of quantitative data is its use in testing hypotheses and determining the strength and direction of effects.

Quantitative Methodology: Methods that seek the *facts* or *causes* of phenomena which can be expressed numerically and analyzed statistically. Interest is in generalizability. Often connoted with a positivist, deductive, natural science world view. Quantitative methods consist of standardized, structured data collection including surveys, closed-ended interviews, tests.

Quasi-Experimental Design: A plan for examining intervention outcomes that involves an intervention group and may involve a comparison group and/or preintervention and/or postintervention tests. This design does not involve random assignment of participants to conditions.

Questionnaire: Research instrument that consists of written questions, each with a limited set of possible responses. See also **self-administered questionnaire; structured interview; telephone survey**.

Random assignment: The process through which members of a pool of eligible evaluation participants are assigned to either the intervention group or

a control group on a random basis, such as through the use of a table of random numbers.

Relational Analysis: Data analysis that reveals the relationship between variables considered important for the evaluation—for example, correlational analysis and regression analysis.

Reliability: The consistency of a measure over repeated use. A measure is said to be reliable if repeated measurements produce the same result. The consistency of responses obtained from similar subjects at different times.

Reporting: Presentation, formal or informal, or evaluation data of other information to communicate processes, roles and results.

Representative Sample: A segment of a larger body or population that mirrors in composition the characteristics of the larger body or population. See also **non-representative sample**.

Research: Studious inquiry or critical and exhaustive investigation having as its aim new and generalizable knowledge

Sample: A segment of a larger body or population.

Sample Attrition: Unplanned reduction in the size of the study sample because of participants' dropping out of the evaluation—for example, because of relocation.

Sampling: Taking a small representative collection from some larger population about which we wish information. The sample is examined and the facts about it learned.

- Random Sampling—Selecting a sample in such a way that each item or person in the population being studied will have an equal chance to be selected for the study.
- Stratified Random Sampling—Dividing the population to be studied into subgroups and sampling each subgroup randomly. This procedure insures that each subgroup will be adequately represented in the final sample.

Self-Administered Questionnaire: A questionnaire that is completed by the respondent without any assistance or clarification from the evaluator.

Self-Evaluation: Self-assessment of program processes and/or outcomes by those conducting or involved in the program.

Self-Selection: An occurrence in which individuals themselves choose to participate in a program or become a member of a sample without the control of the evaluator.

Semi-Structured Interview: Qualitative data collection method that involves an interviewer and specific questions with unlimited response options. See also **structured interview; unstructured interview**.

Simulation: Use of models or mock-ups to solicit perceptions and reactions.

Snowball Sample: A nonrandom sample that is composed according to the referrals of initial sample members such that sample members not only share certain common characteristics, but are likely to be familiar with one another. Also referred to as chain sample.

Stakeholder Evaluation: Evaluation in which stakeholders participate in the design, conduct, and/or interpretation of the evaluation.

Standard Deviation: A unit of measure of variability or dispersion characterizing the tendency for observations to depart from central tendency. Standard deviation and its square, variance, reflect how accurately the central tendency measures (such as the mean) would describe a randomly selected observation.

Statistical Significance: A term referring to the strength of a particular relationship between variables. A relationship is said to be statistically significant when it occurs so frequently in the data that the relationship's existence is probably not attributable to chance.

Structured Interview: Quantitative data collection method that involves an interviewer, specific questions, and limited sets of possible responses to each question. Sometimes referred to as face-to-face questionnaire. See also **semi-structured interview; unstructured interview**.

Summative Evaluation: Evaluation conducted after completion of a program (or a phase of the program) to determine program effectiveness and worth.

Systematic Observation: Practices which have been recommended for an area observed and recorded in some systematic manner.

Telephone Survey: A structured interview conducted over the telephone.

Test: Use of established standards to assess knowledge, skill or performance such as a pen-and-pencil or skills test.

Testimonial: A statement made by a person indicating personal responses and reactions.

Theory-Based Evaluation: Evaluation that begins with identifying the underlying theory about how a program works and uses this theory to build in points for data collection to explain why and how effects occur.

Time-Series Analysis: A form of data analysis that involves examination of data derived from repeated assessments across time.

Unobtrusive Measures: Gathering information without the knowledge of the people in the setting; for example, examination of record books to identify areas of greatest activity; unobtrusive observations of playground interactions to record aggressive behaviors.

Unstructured Interview: Qualitative data collection method that involves an interviewer and given questions. Not all given questions may be asked, however, and additional substantive questions (not necessarily questions for clarification purposes) may be posed by the interviewer.

Utilization Focused Evaluation: A type of evaluation that focuses its design and implementation on use by the intended audience. The evaluator, rather than acting as an independent judge, becomes a facilitator of evaluative decision-making by intended users.

Validity: The extent to which a measure actually captures the concept of interest. The degree to which a data gathering instrument measures the objective it is supposed to measure.

Variable: Factor or characteristic of the intervention, participant, and/or context that may influence or be related to the possibility of achieving intermediate and long-term outcomes.

- **Independent Variables**—Independent variables are the conditions or characteristics the evaluator manipulates in an attempt to ascertain their influence upon some observed phenomena. For example, does the sex of an individual influence perceptions of the usefulness of a workshop? In this example, sex is an independent variable. Another example would be: Does the education level of participants influence the amount of knowledge gained? Educational level is the independent variable and knowledge gained is the dependent variable.
- **Dependent variables**—Dependent variables are conditions or characteristics that appear, disappear, or change as the evaluator introduces, removes, or changes the independent variables. Dependent variables (outcome variable) may be considered the properties or characteristics which are thought to be changed by manipulation or associated with change in another variable. Examples of dependent variables are: Test scores, attitude scores, opinions, ratings, yield, etc. Dependent variables describe outcomes.

Variable Scales: Evaluation studies are usually concerned with four types of variable scales: Nominal, ordinal interval and ratio.

- **Nominal Scales**--Nominal scales are useful in “telling apart.” They allow descriptions to be assigned to classes or categories. These categories are qualitative rather than quantitative. Examples include: Sex (male-female), nationality, and education level. Nominal scales are non-orderable. Nominal scales consist of two or more categories distinguishing the presence or absence of a characteristic or several categories of a characteristic. The categories are exhaustive and mutually exclusive: Male/female.
- **Ordinal Scales**—Ordinal scales permit ordering the variables into categories of “more than” or “less than.” This gives the relative

position of objects or individuals with respect to some attribute. Ordinal measures have no absolute values and the difference between adjacent values may not be equal. These data are called “rank” data, i.e. the items are ranked. For example, you might rank volunteer leaders from “best” to “worst”, or you might have participants in a workshop rank the value of various experiences from “most useful” to “least useful”. Ordinal scales are used in 4-H when “ranking” public speakers from top to bottom or when “ranking” 4-H Record Books.

- **Interval Scales**—In addition to ordering, interval data suggests that the distances between the categories are defined in terms of fixed and equal units. An interval scale not only allows rank ordering of the categories of a variable, but also specifies the relative distance between each pair of categories. The distance between each contiguous pair of categories is the same as that between any other contiguous pair. Likert scales are usually considered interval data.

- **Ratio Scales**—This variable scale has an absolute zero, equal intervals, and is the highest level of measurement. Examples are: Income, age. A ratio scale allows you to multiply or divide each of the values by a certain number without changing the properties of the scale.