

EVALUATION TERMINOLOGY

TIP SHEET

Archival Data: Data from existing written or computer records; data that is collected, stored and updated periodically by certain public agencies, such as schools, social services and public health agencies. This data can be cross-referenced in various combinations to identify individuals, groups and geographic areas.

Attrition: Loss of subjects from the defined sample during the course of the program; unplanned reduction in the size of the study sample because of participants dropping out of the program, for example because of relocation.

Baseline Data: The initial information collected about the condition or performance of subjects prior to the implementation of an intervention, against which progress can be compared at strategic points during and at completion of the program.

Benchmark: Reference point or standard against which performance or achievements can be assessed. A benchmark may refer to the performance that has been achieved in the recent past by other comparable organizations.

Bias: A point of view that inhibits objectivity.

Case studies: An intensive, detailed description and analysis of a single participant, intervention site or program in the context of a particular situation or condition. A case study can be the “story” of one person’s experience with the program. A case study does not attempt to represent every participant’s exact experience.

Coding: To translate a given set of data or items into descriptive or analytic categories to be used for data labeling and retrieval.

Continuous Quality Improvement (CQI): The systematic assessment and feedback of information about planning, implementation and outcomes; may include client satisfaction and utility of the program for participants. Also known as “formative evaluation.”

CSAP's Core Measures: A compendium of data collection instruments that measure those underlying conditions – risks, assets, attitudes and behaviors of different populations – related to the prevention and/or reduction of substance abuse. (CSAP is the Center for Substance Abuse Prevention.)

Data Driven: A process whereby decisions are informed by and tested against systematically gathered and analyzed information.

Data Collection Instruments (evaluation instruments): Specially designed data collection tools (e.g., questionnaires, surveys, tests, observation guides) to obtain measurably reliable responses from individuals or groups pertaining to their attitudes, abilities, beliefs or behaviors. The tools used to collect information.

Feasibility: Refers to practical and logistic concerns related to administering the evaluation instrument. (e.g. Is it reasonably suited to your time constraints, staffing, and/or budgetary resources?)

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

Formative evaluation: Information collected for a specific period of time, often during the start-up or pilot phase of a project, to refine and improve implementation and to assist in solving unanticipated problems. This can include monitoring efforts to provide ongoing feedback about the quality of the program. Also known as “Continuous Quality Improvement.”

Goal: A measurable statement of desired longer-term, global impact of the prevention program that supports the vision and mission statements.

Immediate outcome: The initial change in a sequence of changes expected to occur as a result of implementation of a science-based program.

Impact evaluation: A type of outcome evaluation that focuses on the broad, long-term impacts or results of program activities; effects on the conditions described in baseline data.

Incidence: The number of new cases/individuals within a specified population who have initiated a behavior – in this case drug, alcohol, or tobacco use – during a specific period of time. This identifies new users to be compared to the number of new users historically, over comparable periods of time, usually expressed as a “rate.”

Indicator: A measure selected as a marker of whether the program was successful in achieving its desired results. Can be a substitute measure for a concept that is not directly observable or measurable, e.g. prejudice, self-efficacy, substance abuse. Identifying indicators to measure helps a program more clearly define its outcomes.

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.

Intermediate Outcome: The changes that are measured subsequent to immediate change, but prior to the final changes that are measured at program completion in a sequence of changes expected to occur in a science-based program,

Interviews: Guided conversations between a skilled interviewer and an interviewee that seek 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.

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

Key informant: Person with background, knowledge, or special skills relevant to topics examined by the evaluation; sometimes an informal leader or spokesperson in the targeted population.

Logic Model: A graphic depiction of how the components of a program link together and lead to the achievement of program goals/outcomes. A logic model is a succinct, logical series of statements linking the needs and resources of your community to strategies and activities that address the issues and define the expected results/outcomes.

Longitudinal study: A 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 influences of an intervention, maturation, or the environment.

Mean: The measure of central tendency – the sum of all the scores added and then divided by the total number of scores; the “average.” The mean is sensitive to extreme scores, which can cause the mean to be less representative of the typical average.

Median: A measure of central tendency referring to the point exactly midway between the top and bottom halves of a distribution of values; the number or score that divides a list of numbers in half.

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

Mode: A measure of central tendency referring to the value most often given by respondents; the number or score that occurs most often.

Objective: A specific, measurable description of an intended outcome.

Objectivity: The expectation that data collection, analysis and interpretation will adhere to standards that eliminate or minimize bias; objectivity insures that outcome or evaluation results are free from the influence of personal preferences or loyalties.

Observation: Data collection method involving unobtrusive examination of behaviors and/or occurrences, often in a natural setting, and characterized by no interaction between participants and observers. Data is obtained by simply watching what people do and documenting it in a systematic manner.

Outcomes: Changes in targeted attitudes, values, behaviors or conditions between baseline measurement and subsequent points of measurement. Changes can be immediate, intermediate or long-term; the results/effects expected by implementing the program’s strategies.

Outcome evaluation: Examines the results of a program’s efforts at various points in time during and after implementation of the program’s activities. It seeks to answer the question, “what difference did the program make?”

Pre/post tests: Standardized methods used to assess change in subjects’ knowledge and capacity to apply this knowledge to new situations. The tests are administered prior to implementation of the program and after completion of the program’s activities. Determines performance prior to and after the delivery of an activity or strategy.

Participatory evaluation: The process of engaging stakeholders in an evaluation effort. Stakeholders are those people most invested in the success of a program, including staff, board members, volunteers, other agencies, funding agencies and program participants. Getting input from the stakeholders at all stages of the evaluation effort – from deciding what questions to ask, to collecting data, to analyzing data and presenting results.

Prevalence: The number of all new and old cases of a disease or occurrences of an event during a specified time, in relation to the size of the population from which they are drawn. (Numbers of people using or abusing substances during a specified period.) Prevalence is usually expressed as a rate, such as the number of cases per 100,000 people.

Process evaluation: Addresses questions related to how a program is implemented. Compares what was supposed to happen with what actually happened. Includes such information as “dosage” and “frequency,” staffing, participation and other factors related to implementation. Documents what was actually done, how much, when, for whom, and by whom during the course of the project – inputs.

Qualitative data: Non-numerical data rich in detail and description that are usually presented in a textual or narrative format, such as data from case studies, focus groups, interviews or document reviews. Used with open-ended questions, this data has the ability to illuminate evaluation findings derived from quantitative methods.

Quantitative data: Numeric information, focusing on things that can be counted, scored and categorized; used with close-ended questions, where participants have a limited set of possible answers to a question. Quantitative data analysis utilizes statistical methods.

Questionnaire: Highly structured series of written questions that is administered to a large number of people; questions have a limited set of possible responses.

Random assignment: A process by which the people in a sample to be tested are chose at random from a larger population; a pool of eligible evaluation participants are selected on a random basis.

Reliability: The consistency of a measurement or measurement instrument over time. Consistent results over time with similar populations and under similar conditions confirm the reliability of the measure. (Can the test produce reliable results each time it is used and in different locations or with different populations?)

Representative sample: A segment or group taken from a larger body or population that mirrors in composition the characteristics of the larger body or population.

Sample: A segment of a larger body or population.

Standardized tests: Tests or surveys that have standard printed forms and content with standardized instructions for administration, use, scoring, and interpretation.

Subjectivity: Subjectivity exists when the phenomena of interest are described or interpreted in personal terms related to one's attitudes, beliefs or opinions.

Surveys: Uses structured questions from specially designed instruments to collect data about the feelings, attitudes and/or behaviors of individuals.

Targets: Defines who or what and where you expect to change as a result of your efforts.

Theory of change: A set of assumptions about how and why desired change is most likely to occur as a result of your program, based on past research or existing theories of behavior and development. Defines the evidenced-based strategies or approaches proven to address a particular problem. Forms the basis for logic model planning.

Threshold: Defines how much change you can realistically expect to see over a specific period of time, as a result of your program strategies.

Validity: The extent to which a measure of a particular construct/concept actually measures what it purports to measure; how well a test actually measures what it is supposed to measure.