

Glossary

- attrition** - The percentage of the individuals (or groups) originally randomized for whom outcome data could not be obtained.
- baseline data** - Initial information on program participants or other program aspects collected prior to receipt of services or program intervention.
- bias** - (refers to statistical bias). Inaccurate representation that produces systematic error in a research finding. Bias may result in overestimating or underestimating certain characteristics of the population. It may result from incomplete or invalid information and may be intentional or unintentional.
- comparison group** - Individuals whose characteristics (such as race/ethnicity, gender, and age) are similar to those of your program participants. As part of the evaluation process, the experimental (or treatment) group and the comparison group are assessed to determine which type of services, activities, or products provided by your program produced the expected changes.
- control group** - A group of individuals whose characteristics (such as race/ethnicity, gender, and age) are equivalent to those of your program participants, but do not receive the program (services, products, or activities) you are evaluating.
- cost effectiveness analysis** - A type of analysis that involves comparing the relative costs of operating a program with the extent to which the program met its goals and objectives. For example, a program to reduce cigarette smoking would estimate the dollars that had to be expended in order to convert each smoker into a nonsmoker.
- cost-benefit analysis** - A type of analysis that involves comparing the relative costs of operating a program to the benefits it generates. For example, a program to reduce cigarette smoking would focus on the difference between the dollars expended for converting smokers into nonsmokers with the dollar savings from reduced medical care for smoking-related disease, days lost from work, and the like.
- cultural relevance** - Demonstration that evaluation methods, procedures, and/or instruments are appropriate for the culture(s) to which they are applied.
- data** - Specific information or facts that are collected. Examples of data items might include age, date of entry into program, or reading level.
- data analysis** - The process of systematically applying statistical and logical techniques to describe, summarize, and compare data collected.
- data collection instruments** - Forms used to collect information for your evaluation. Forms may include interview instruments, intake forms, case logs, and attendance records.
- data collection plan** - A written document describing the specific procedures to be used to gather the evaluation information or data.
- design** - The overall plan and specification of the approach expected in a particular evaluation. The design describes how you plan to measure program components and how you plan to use the resulting measurements.
- effect size** - A measure of the strength of the relationship between two variables. In a nontechnical context, if a tutoring program asserts that it raises school performance by one letter grade, this grade increase is the claimed “effect size” of the program. In a technical context, a standardized effect size can be measured as the difference in mean outcome between the treatment and control groups, divided by the pooled standard deviation.

- evaluation** - A systematic method for collecting, analyzing, and using information to answer questions about a program. It helps to identify effective and ineffective services, practices, and approaches.
- evaluation plan** - Document describing the approach or design used to guide an evaluation. Includes what you plan to do, how & who will do it, when it will be done, and why the evaluation is being conducted.
- evaluator** - An individual trained and experienced in designing and conducting an evaluation that uses tested and accepted research methodologies.
- evidence-based practice** - Program practices that are developed/implemented using data from the published research literature. For example, in mentoring programs several studies have found an hour of volunteer training contributes to greater outcomes for youth. Based on this information, a mentoring program implements an hour of volunteer training.
- evidence-based program** – Programs that have been found to be effective based on rigorous evaluations, usually involved experimental or quasi-experimental designs.
- experimental group** - A group receiving the treatment or intervention being evaluated or studied. Experimental groups (or treatment groups) are usually compared to a control or comparison group.
- experimental design/randomized control trial (RCT)** - A type of scientific experiment most commonly healthcare or health technology evaluation. RCTs involve the random allocation of different interventions (treatments or conditions) to subjects. The most important advantage of proper randomization is that it eliminates selection bias in the assignment of treatments.
- focus group** - A group of 7-10 people convened for the purpose of obtaining perceptions or opinions, suggesting ideas, or recommending actions.
- formative evaluation** - A type of process evaluation that focuses on collecting data on program operations so that needed changes or modifications can be made to the program in its early stages. Formative evaluations are used to provide feedback to staff about the program components that are working and those that need to be changed.
- impact evaluation** – Assesses the changes that can be attributed to a particular intervention, such as a project, program or policy. In contrast to outcome monitoring, which examines whether targets have been achieved, impact evaluation is structured to answer the question: how would participants' well-being have changed if the intervention had not been undertaken? For example, an impact evaluation could show that a decrease in a community's overall infant mortality rate was the direct result of a program designed to provide early prenatal care.
- instrument** - A tool used to collect and organize information. Includes written instruments or measures, such as questionnaires, scales, and tests.
- intention-to-treat analysis** - This means that the study sought outcome data for all of the individuals randomly assigned to the intervention group, even those who did not complete the intervention, and used such data in calculating the intervention's effect.
- intervention** - The specific services, activities, or products developed and implemented to change or improve program participants' knowledge, attitudes, behaviors, or awareness.
- logic model** - See the definition for program model.
- measurable terms** - Specifying, through clear language, what it is you plan to do and how you plan to do it. Stating time periods for activities, "dosage" or frequency information (such as three 1-hour training sessions), and number of participants helps to make project activities measurable.
- methodology** - A methodology describes how something will be (or was) done. The methodology includes the methods, procedures, and techniques used to collect and analyze information.

- monitoring** - The process of reviewing a program or activity to determine whether set standards or requirements are being met. Unlike evaluation, monitoring compares a program to an ideal or exact state.
- objective** - A specific statement that explains how a program goal will be accomplished. For example, an objective of the goal to improve adult literacy could be to provide tutoring to participants on a weekly basis for 6 months. An objective is stated so that changes, in this case, an increase in a specific type of knowledge, can be measured and analyzed. Objectives are written using measurable terms and are time-limited.
- outcome** - Outcomes are a result of the program, services, or products you provide and refer to changes in knowledge, attitude, or behavior in participants.
- outcome evaluation/summative evaluation** - Addresses the question of what are the results. It is common to speak of short-term outcomes and long-term outcomes. For example, in an exercise program, a short-term outcome could be a change in knowledge about the health effects of exercise, or it could be a change in exercise behavior. A long-term outcome could be less likelihood of dying from heart disease.
- outcome objectives** - The changes in knowledge, attitudes, awareness, or behavior that you expect to occur as a result of implementing your program component, service, or activity. Also known as participant outcome objectives.
- outside evaluator** - An evaluator not affiliated with your agency prior to the program evaluation. Also known as a third-party evaluator.
- participant** - An individual, family, agency, neighborhood, community, or State, receiving or participating in services provided by your program. Also known as a client or target population group.
- pilot test** - Preliminary test or study of your program or evaluation activities to try out procedures and make any needed changes or adjustments. For example, an agency may pilot test new data collection instruments that were developed for the evaluation.
- post-test** - A test or measurement taken after a service or intervention takes place. It is compared with the results of a pretest to show evidence of the effects or changes as a result of the service or intervention being evaluated.
- pre-test** - A test or measurement taken before a service or intervention begins. It is compared with the results of a posttest to show evidence of the effects of the service or intervention being evaluated. A pretest can be used to obtain baseline data.
- process evaluation** - An evaluation that examines the extent to which a program is operating as intended by assessing ongoing program operations and whether the targeted population is being served. A process evaluation involves collecting data that describes program operations in detail, including the types and levels of services provided, the location of service delivery, staffing; sociodemographic characteristics of participants; the community in which services are provided, and the linkages with collaborating agencies. A process evaluation helps program staff identify needed interventions and/or change program components to improve service delivery. It is also called formative or implementation evaluation.
- program implementation objectives** - What you plan to do in your program, component, or service. For example, providing therapeutic child care for 15 children, giving them 2 hot meals per day, are referred to as program implementation objectives.
- program model/logic model** - A concise, coherent outline of what a program/intervention/treatment is designed to do. It articulates the target participants, intended short and long-term outcomes, the resources being invested, and the actions and strategies employed. It can be used both to plan an effective program and to design appropriate evaluation.
- qualitative data** - Information that is difficult to measure, count, or express in numerical terms. For example, a participant's impression about the fairness of a program rule/requirement is qualitative data.

quantitative data - Information that can be expressed in numerical terms, counted or compared on a scale. For example, improvement in a child's reading level as measured by a reading test.

quasi-experimental design (QED) - A research design in which subjects are assigned to “treatment” (that is, they receive the intervention being studied) and “comparison” groups through a process that is not random. Used where it is not feasible or desirable to conduct an experiment or randomized control trial, e.g. evaluating the impact of public policy changes, educational interventions or large scale health interventions. Various statistical techniques such as multiple regression and the use of propensity scores can be used to model and reduce the effects of confounding variables, thereby improving the accuracy of the results obtained from quasi-experiments.

randomized control trial (RCT) – See experimental design.

random assignment - The assignment of individuals in the pool of all potential participants to either the experimental (treatment) or control group in such a manner that their assignment to a group is determined entirely by chance.

reliability - Extent to which a measurement (such as an instrument or a data collection procedure) produces consistent results over repeated observations or administrations of the instrument under the same conditions each time. It is also important that reliability be maintained across data collectors; this is called interrater reliability.

sample - A subset of participants selected from the total study population. Samples can be random (selected by chance, such as every 6th individual on a waiting list) or nonrandom (selected purposefully, such as all 2-year olds in a Head Start program).

sample size - This is the total number of individuals (or groups) randomly assigned to the intervention and control groups.

statistical test - Type of statistical procedure, such as a t-test or Z-score, that is applied to data to determine whether your results are statistically significant (i.e., the outcome is not likely to have resulted by chance alone).

statistically significant - If a study finds that the difference in outcomes between the intervention and control groups is “statistically significant,” it means it is highly unlikely that the difference is the result of chance. For example, statistical significance at the .05 level (the most commonly-used level) means that there is only a 1 in 20 probability that the difference could have occurred by chance if the intervention's true effect is zero.

summative evaluation - A type of outcome evaluation that assesses the results or outcomes of a program. This type of evaluation is concerned with a program's overall effectiveness.

treatment group - Also called an experimental group, a treatment group is composed of a group of individuals receiving the services, products, or activities (interventions) that you are evaluating.

validity - The extent to which a measurement instrument or test accurately measures what it is supposed to measure. For example, a reading test is a valid measure of reading skills, but is not a valid measure of total language competency.

variables - Specific characteristics or attributes, such as behaviors, age, or test scores, that are expected to change or vary. For example, the level of adolescent drug use after being exposed to a drug prevention program is one variable that may be examined in an evaluation.

Sources - U.S. Department of Health & Human Svces, Administration for Children & Families (<http://goo.gl/yqqqs>); Coalition for Evidence-based Policy (<http://goo.gl/DKYy>); U.S. Department of Education, Institute for Education Sciences *What Works Clearinghouse* (<http://goo.gl/bclx>).