Term
Accessibility Definition: The extent to which the structural and organizational arrangements facilitate participation in the program.
Accountability Definition: The responsibility of program staff to provide evidence to stakeholders and sponsors that a program is effective and in conformity with its coverage, service, legal, and fiscal requirements.
Activities Definition: Things you do-activities you plan to conduct in your program.
After-Only Designs Definition: One-shot studies; evaluation designs involving only measures taken after the program has been completed.
Analysis Definition: A systematic approach to problem solving. Complex problems are made simpler by separating them into more understandable elements. This involves the identification of purposes and facts, the statement of defensible assumptions, and the formulation of conclusions.
Analysis of Covariance Definition: A method for analyzing the differences in the means of two or more groups of cases while taking account of variation in

Term
one interval-ratio variable.
Analysis of Variance Definition: A method for analyzing the differences in the means of two or more groups of cases.
Anchors Definition: Anchors are items that serve as reference points from which other items in the series or other points in the scale are judged or compared.
Anonymity (provision for) Definition: 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.
Applied Research Definition: Research designed for the purpose of producing results that may be applied to real world situations.
Association Definition: General term for the relationship among variables.
Asymmetric Measure of Association Definition: A measure of association that makes a distinction between independent and dependent variables.

Term
Attitude Surveys Definition: Data collection techniques designed to collect standard information from a large number of subjects concerning their attitudes or feelings. These typically refer to questionnaires or interviews.
Attribute Definition: A characteristic that describes a person, thing, or event.
Attribution Definition 1: The assertion that certain events or conditions were, to some extent, caused or influenced by other events or conditions. This means a reasonable connection can be made between a specific outcome and the actions and outputs of a government policy, program, or initiative. Definition 2: See Anonymity.
Attrition Definition: The loss of subjects during the course of a study. This may be a threat to the validity of conclusions if participants of study and comparison/control groups drop out at different rates or for different reasons.
Audit Definition: The systematic examination of records and the investigation of other evidence to determine the propriety, compliance, and adequacy of programs, systems, and operations. The auditing process may include tools and techniques available from such diverse areas as engineering, economics, statistics, and accounting.
Backfill Techniques

Term
Definition: Techniques used in cumulative case studies to collect information needed if the study is to be usable for aggregation; these techniques include, for example, obtaining missing information from the authors on how instances studied were identified and selected.
Baseline Data Definition: Initial information on a program or program components collected prior to receipt of services or participation activities. Baseline data are often gathered through intake interviews and observations and are used later for comparing measures that determine changes in a program.
Batch Definition: A group of cases for which no assumptions are made about how the cases are selected. A batch may be a population, a probability sample, or a nonprobability sample, but the data are analyzed as if the origin of the data is not known.
Before-After Designs Definition: The elementary quasi-experimental design known as the before-after design involves the measurement of "outcome" indicators (e.g., arrest rates, attitudes) prior to implementation of the treatment, and subsequent re-measurement after implementation. Any change in the measure is attributed to the treatment. This design provides a significant improvement over the one-shot study because it measures change in the factor(s) to be impacted. However, this design suffers from threats of history - the possibility that some alternate factor (besides the treatment) has actually caused the change.
Bell-Shaped Curve

Term
Definition: A distribution with roughly the shape of a bell; often used in reference to the normal distribution but others, such as the t distribution, are also bell-shaped.
Benchmarking Definition: Measuring progress toward a goal at intervals prior to the anticipated attainment of the goal.
Benchmarks Definition: Measures of progress toward a goal, taken at intervals prior to the program's completion or the anticipated attainment of the final goal.
Benefits Definition: Net program outcomes, usually translated into monetary terms. Benefits may include both direct and indirect costs.
Best Practices Definition: New ideas or lessons learned about effective program activities that have been developed and implemented in the field and have been shown to produce positive outcomes.
Between-Group Variances Definition: Indications of how the mean and variances of each group differ from the other groups.
Bias

Term
Definition: The extent to which a measurement, sampling, or analytic method systematically underestimates or overestimates the true value of an attribute. For example, words, sentence structure, attitudes, and mannerisms may unfairly influence a respondent's answer to a question. Bias in questionnaire data can stem from a variety of other factors, including choice of words, sentence structure, and the sequence of questions.
Biased Sample Definition: A sample that is not representative of the population to which generalizations are to be made.
Binary Variables Definition: A variable that identifies the presence or absence of a trait, characteristic, opinion, etc.; a "yes/no" variable.
Bivariate Analysis Definition: An analysis of the relationship between two variables.
Bivariate Data Definition: Information about two variables.
Black Box Evaluation Definition: Evaluation of program outcomes without the benefit of an articulated program theory to provide insight into what is presumed to be causing those outcomes and why.
Cascading Design

Term
Preferred Term: Emergent Design
Case Study Definition: A method for learning about a complex instance, based on a comprehensive understanding of that instance, obtained by extensive description and analysis of the instance, taken as a whole and in its context.
Categorical Measure Definition: A measure that places data into a limited numbers of groups or categories.
Causal Analysis Definition: A method for analyzing the possible causal associations among a set of variables.
Causal Association Definition: A relationship between two variables in which a change in one brings about a change in the other.
Causal Model Definition: A model or portrayal of the theorized causal relationships between concepts or variables.
Causal Relationship Definition: The relationship of cause and effect. The cause is the act or event that produces the effect. The cause is necessary to

Term
produce the effect.
Closed-Ended Questions Definition: A question that limits responses to predetermined categories.
Closed Question Definition: A question with more than one possible answer from which one or more answers must be selected.
Cluster Sample Definition: A probability sample for which groups or geographic areas comprising groups were randomly selected.
Clustering Definition: Identifying similar characteristics and grouping samples with similar characteristics together.
Codebook Definition: A document which lists the variables in a dataset, possible values for each variable, and the definitions of codes that have been assigned to these values.
Coding Definition: The process of converting information obtained on a subject or unit into coded values (typically numeric) for the purpose of data storage, management, and analysis.

Term
Coefficient Definition: A value expressing the degree to which some characteristic or relation is to be found in specified instances.
Collaborative Evaluation Preferred Term: Participatory Evaluation
Comparative Change Design Definition: The quasi-experimental design known as the comparative change design allows for the measurement of change in relevant outcome factors (using a pre- and post-test) and provides for comparison of this change between a treatment group and a non-random comparison group. Because comparison and treatment groups are not randomly selected, alternate explanations due to prior differences between groups continue to be a threat.
Comparative Post-test Design
Comparative Time-Series Design Definition: The quasi-experimental design known as the comparative time series tracks some outcome of interest for periods before and after program implementation for both the treatment group as well as a non-randomly selected comparison group. Because comparison and treatment groups are not randomly selected, alternate explanations due to prior differences between groups continue to be a threat.
Comparison Group

Term
Definition: A group of individuals whose characteristics are similar to those of a program's participants. These individuals may not receive any services, or they may receive a different set of services, activities, or products; in no instance do they receive the same services as those being evaluated. As part of the evaluation process, the experimental group (those receiving program services) and the comparison group are assessed to determine which types of services, activities, or products provided by the program produced the expected changes.
Composite Measure Definition: A measure constructed using several alternate measures of the same phenomenon.
Concept Definition: An abstract or symbolic tag that attempts to capture the essence of reality. The "concept" is later converted into variables to be measured.
Conditional Distribution Definition: The distribution of one or more variables given that one or more other variables have specified values. For example, a distribution of height might be made conditional on gender and age, allowing you to find the distribution of height for men aged 18-22.
Confidence Interval Definition: An estimate of a population parameter that consists of a range of values bounded by statistics called upper and lower confidence limits, within which the value of the parameter is expected to be located.
Confidence Level

Term
Definition: The level of certainty to which an estimate can be trusted. The degree of certainty is expressed as the chance that a true value will be included within a specified range, called a confidence interval.
Confidence Limits Definition: Two statistics that form the upper and lower bounds of a confidence interval.
Confidentiality Definition: Secrecy. In research this involves not revealing the identity of research subjects, or factors which may lead to the identification of individual research subjects.
Confidentiality Form Definition: A written form that assures evaluation participants that information they provide will not be openly disclosed nor associated with them by name. Since an evaluation may entail exchanging or gathering privileged or sensitive information about residents or other individuals, a confidentiality form ensures that the participants' privacy will be maintained.
Confounding Definition: An inability to distinguish the separate impacts of two or more individual variables on a single outcome.
Consensus Building Outcome Definition: The production of a common understanding among participants about issues and programs.
Constraint

Term
Definition: A limitation of any kind to be considered in planning, programming, scheduling, implementing, or evaluating programs.
Construct Definition: A concept that describes and includes a number of characteristics or attributes. The concepts are often unobservable ideas or abstractions.
Construct Validity Definition: The extent to which a measurement method accurately represents a construct and produces an observation distinct from that produced by a measure of another construct.
Consultant Definition: An individual who provides expert or professional advice or services, often in a paid capacity.
Contamination Definition: The tainting of members of the comparison or control group with elements from the program. Contamination threatens the validity of the study because the group is no longer untreated for purposes of comparison.
Content Analysis Definition: A set of procedures for collecting and organizing nonstructured information into a standardized format that allows one to make inferences about the characteristics and meaning of written and otherwise recorded material.

Term
Content Validity Definition: The ability of the items in a measuring instrument or test to adequately measure or represent the content of the property that the investigator wishes to measure.
Context (of an evaluation) Definition: The combination of the factors accompanying the study that may have influenced its results. These factors include the geographic location of the study, its timing, the political and social climate in the region at that time, the other relevant professional activities that were in progress, and any existing pertinent economic conditions.
Continuous Variable Definition: A quantitative variable with an infinite number of attributes.
Contract Definition: A written or oral agreement between the evaluator and client that is enforceable by law. It is a mutual understanding of expectations and responsibilities for both parties.
Control Group Definition: A group whose characteristics are similar to those of the program but who do not receive the program services, products, or activities being evaluated. Participants are randomly assigned to either the experimental group (those receiving program services) or the control group. A control group is used to assess the effect of program activities on participants who are receiving the services, products, or activities being evaluated. The same information is collected for people in the control group and those in the

Term
experimental group.
Control Variable Definition: A variable that is held constant or whose impact is removed in order to analyze the relationship between other variables without interference, or within subgroups of the control variable.
Convenience Sample Definition: A sample for which cases are selected only on the basis of feasibility or ease of data collection. This type of sample is rarely useful in evaluation and is usually hazardous.
Correlation Preferred Term: Association
Correlation coefficient Definition: A numerical value that identifies the strength of relationship between variables.
Cost-Benefit Definition: A criterion for comparing programs and alternatives when benefits can be valued in dollars. Cost-benefit is the ratio of dollar value of benefit divided by cost. It allows comparison between programs and alternative methods.
Cost-Benefit Analysis

Term
Definition: An analysis that compares present values of all benefits less those of related costs when benefits can be valued in dollars the same way as costs. A cost-benefit analysis is performed in order to select the alternative that maximizes the benefits of a program.
Cost-Effectiveness Definition: A criterion for comparing alternatives when benefits or outputs cannot be valued in dollars. This relates costs of programs to performance by measuring outcomes in nonmonetary form. It is useful in comparing methods of attaining an explicit objective on the basis of least cost or greatest effectiveness for a given level of cost.
Costs Definition: Inputs, both direct and indirect, required to produce an intervention.
Covariation Definition: The degree to which two measures vary together.
Coverage Definition: The extent to which a program reaches its intended target population.
Cross-Sectional Data Definition: Observations collected on subjects or events at a single point in time.
Cues

Term
Definition: The alternative responses to questions that increase or decrease in intensity in an ordered fashion. The interviewee is asked to select one answer to the question.
Cultural Competency Definition: A set of academic and interpersonal skills that allow individuals to increase their understanding and appreciation of cultural differences and similarities within, among, and between groups.
Data Definition: Documented information or evidence of any kind.
Data Analysis Definition: The process of systematically applying statistical and logical techniques to describe, summarize, and compare data.
Data Collection Instrument Definition: A form or set of forms used to collect information for an evaluation. Forms may include interview instruments, intake forms, case logs, and attendance records. They may be developed specifically for an evaluation or modified from existing instruments.
Data Collection Plan Definition: A written document describing the specific procedures to be used to gather the evaluation information or data. The document describes who collects the information, when and where it is collected, and how it is obtained.

Term
Database Definition: A collection of information that has been systematically organized for easy access and analysis. Databases typically are computerized.
Demographic Question Definition: A question used in compiling vital background and social statistics.
Dependent Variable Definition: A variable that may, it is believed, be predicted by or caused by one or more other variables called independent variables.
Descriptive Statistic Definition: A statistic used to describe a set of samples upon which observations were made.
Design Definition: The overall plan for a particular evaluation. The design describes how program performance will be measured and includes performance indicators.
Dichotomous Variable Definition: A variable with only two possible values.
Direct Benefit

Term
Definition: Result that is closely related with the program by cause and effect.
Direct Costs Definition: Resources that must be committed to implement a program.
Direct Impact Definition: An effect of a program that addresses a stated goal or objective of that program.
Discounting Definition: The treatment of time in valuing costs and benefits, that is, the adjustment of costs and benefits to their present values, requiring a choice of discount rate and time.
Discrete Variable Definition: A quantitative variable with a finite number of attributes.
Dispersion Definition: The extent of variation among cases.
Distribution of a Variable Definition: Variation of characteristics across cases.

Term
Distributional Effects Definition: Effects of programs that result in a redistribution of resources in the general population.
Document Review Definition: A technique of data collection involving the examination of existing records or documents.
Dummy Variables Definition: A dichotomous variable, typically used in regression analysis, which indicates the existence (and lack of existence) of a characteristic or group of characteristics in a case.
Effect Size Definition: The size of the relationship between two variables (particularly between program variables and outcomes).
Effectiveness Definition: Ability to achieve stated goals or objectives, judged in terms of both output and impact.
Efficiency Definition: The degree to which outputs are achieved in terms of productivity and input (resources allocated). Efficiency is a measure of performance in terms of which management may set objectives and plan schedules and for which staff members may be held accountable.

Term
Efficiency Assessment Definition: An evaluative study that answers questions about program costs in comparison to either the monetary value of their benefits or their effectiveness in terms of the changes they bring about in the social conditions they address.
Emergent Design Definition: An implementation plan in which the specification of every step depends upon the results of previous steps.
Empirical Definition: Relying upon or derived from observation or experiment.
Empirical Research Definition: Research that uses data drawn from observation or experience.
Empirical Validity Definition: Empirical evidence that an instrument measures what it has been designed to measure.
Empowerment Evaluation Definition: An approach to gathering, analyzing, and using data about a program and its outcome that actively involves key stakeholders in the community in all aspects of the evaluation process, and that promotes evaluation as a strategy for empowering communities to engage in system changes.

Term
Estimation Error Definition: The amount by which an estimate differs from a true value. This error includes the error from all sources (for example, sampling error and measurement error).
Evaluability Assessment Definition: Negotiation and investigation undertaken jointly by the evaluator, the evaluation sponsor, and possibly other stakeholders to determine if a program meets the preconditions for evaluation and, if so, how the evaluation should be designed to ensure maximum utility.
Evaluation Definition: Evaluation has several distinguishing characteristics relating to focus, methodology, and function. Evaluation (1) assesses the effectiveness of an ongoing program in achieving its objectives, (2) relies on the standards of project design to distinguish a program's effects from those of other forces, and (3) aims at program improvement through a modification of current operations.
Evaluation Plan Definition: A written document describing the overall approach or design that will be used to guide an evaluation. It includes what will be done, how it will be done, who will do it, when it will done, and why the evaluation is being conducted.
Evaluation Practice Definition: A practice or set of practices that consist mainly of management information and data incorporated into regular program management information systems to allow managers to monitor and assess the progress being made in each program toward its

Term
goals and objectives.
Evaluation Team Definition: The individuals, such as the evaluation consultant and staff, who participate in planning and conducting the evaluation. Team members assist in developing the evaluation design, developing data collection instruments, collecting data, analyzing data, and writing the report.
Ex-post Facto Design Definition: A research design in which all group selection, pretest data, and posttest data are collected after completion of the treatment. The evaluator is thus not involved in the selection or placement of individuals into comparison or control groups. All evaluation decisions are made retrospectively.
Executive Report Definition: An abbreviated report that has been tailored specifically to addresses the concerns and questions of a person whose function is to administer a program or project.
Executive Summary Definition: A non-technical summary statement designed to provide a quick overview of the full-length report on which it is based.
Experimental Data Definition: Data produced by an experimental or quasi-experimental design.

Term
Experimental Design Definition: A research design in which the researcher has control over the selection of participants in the study, and these participants are randomly assigned to treatment and control groups.
Experimental Group Definition: A group of individuals participating in the program activities or receiving the program services being evaluated or studied. Experimental groups are usually compared to a control or comparison group.
Experimental Mortality Definition: The loss of subjects from an experiment due to such factors as illness, lack of interest, or refusal to participate. This loss may effect the comparability of results between the experimental and control groups.
External Evaluation Definition: Collection, analysis, and interpretation of data conducted by an individual or organization outside of the organization being evaluated.
External Evaluator Preferred Term: Outside Evaluator
External Validity

Term
Definition: The extent to which a finding applies (or can be generalized) to persons, objects, settings, or times other than those that were the subject of study.
External Validity Threats Definition: Factors that may reduce the transferability of a program's findings to other groups or jurisdictions.
Externalities Definition: Effects of a program that impose costs on persons or groups who are not targets.
Feasibility Study Definition: A study of the applicability or practicability of a proposed action or plan.
Field Notes Definition: A written record of observations, interactions, conversations, situational details, and thoughts during the study period.
Field Test Definition: The study of a program, project, or instructional material in settings like those where it is to be used. Field tests may range from preliminary investigations to full-scale summative studies.
Flow Chart Definition: A graphic presentation using symbols to show the step-by-step sequence of operations, activities, or procedures. Used in

Term
computer system analysis, activity analysis, and in general program sequence representations.
Focus Group Definition: A group of people convened for the purpose of obtaining perceptions or opinions, suggesting ideas, or recommending actions. A focus group is a method of collecting information for the evaluation process.
Focused Interview Definition: An interview organized around several predetermined questions or topics but providing some flexibility in the sequencing of the questions and without a predetermined set of response categories or specific data elements to be obtained.
Forced-Choice Question Definition: A question that requires respondents to choose between available options. Options such as "other" or "none of the above" are not available alternatives.
Forecasting Definition: Estimating the likelihood of an event taking place in the future, based on available data from the past.
Formative Evaluation Definition: Evaluative activities undertaken to furnish information that will guide program improvement.
Frequency Distribution

Term
Definition: A distribution of the count of cases corresponding to the attributes of an observed variable.
Function Definition: A group of related activities and/or projects for which an organizational unit is responsible. This is also the principal purpose a program is intended to serve.
Generalizability Definition: The extent to which the findings of a study can be applied to other populations, settings, or times.
Goal Definition: A desired state of affairs that outlines the ultimate purpose of a program. This is the end toward which project or program efforts are directed.
Halo Effect Definition: The halo effect refers to the tendency to rate a person's skills and talents in many areas based upon an evaluation of a single factor. It creates bias by an observer's tendency to rate, perhaps unintentionally, certain objects or persons in a manner that reflects what was anticipated.
Hawthorne Effect Definition: When people know that they are being observed and someone is interested in them, their production increases. Researchers must be aware that the mere fact that they are making changes may cause changes in the behaviour of people. It is wise to ensure that your measuring does not affect the process you are trying to measure.

Term
History Effect Definition: This threat to internal validity refers to specific events, other than the program, which may have taken place during the course of study which may have produced the results.
Hypothesis Definition: A specific statement regarding the relationship between two variables. In evaluation research, this typically involves a prediction that the program or treatment will cause a specified outcome. Hypotheses are confirmed or denied based on empirical analysis.
Immediate Outcome Definition: The changes in program participants' knowledge, attitudes, and behavior that occur at certain times during program activities.
Impact Definition: The ultimate effect of the program on the problem or condition that the program or activity was supposed to do something about.
Impact Evaluation Definition: A form of outcome evaluation that assesses the net effect of a program by comparing program outcomes with an estimate of what would have happened in the absence of a program. This form of evaluation is employed when external factors are known to influence the program's outcomes, in order to isolate the program's contribution to achievements or its objectives.

Term
Impact Theory Definition: The beliefs, assumptions, and expectations inherent in a program about the nature of the change brought about by program action and how it results in the intended improvement in social conditions. Program impact theory is causal theory: It describes a cause-and-effect sequence in which certain program activities are the instigating causes and certain social beliefs are the effects they eventually produce.
Implementation Definition: Development of a program. The process of putting all program functions and activities into place.
Implementation Evaluation Preferred Term: Process Evaluation
Implementation Failure Definition: The program does not adequately perform the activities specified in the program design that are assumed to be necessary for bringing about the intended social improvements. It includes situations in which no service, not enough service, or the wrong service is delivered, or the service varies excessively across the target population.
Implementation Strategy Definition: The plan for development of a program and procedure for ensuring the fulfillment of intended functions or services.
Implemented

Term
Definition: Developed or put into place.
In-Person Interviewing Definition: Face-to-face interviewing. The interviewer meets personally with the respondent to conduct the interview.
Independent Evaluation Definition: An evaluation in which the evaluator has the primary responsibility for developing the evaluation plan, conducting the evaluation, and disseminating the results.
Independent Variable Definition: A variable that may, it is believed, predict or cause fluctuation in a dependent variable.
Index Definition: A set of related measures combined to characterize a more abstract concept.
Index of Dispersion Definition: A measure of spread; a statistic used especially with nominal variables.
Indicator Definition: A measure that consists of ordered categories arranged in ascending or descending order of desirability.
Indirect Benefit

Term
Definition: Results that are related to a program, but not its intended objectives or goals.
Indirect Costs Definition: The costs associated with impacts or consequences of a program.
Indirect Impact Definition: An effect of a program that is not associated with one of its stated objectives.
Inferential Statistic Definition: A statistic used to describe a population using information from observations on only a probability sample of cases from the population.
Information System Definition: An organized collection, storage, and presentation system of data and other knowledge for decision making, progress reporting, and for planning and evaluation of programs. It can be either manual or computerized, or a combination of both.
Informed Consent Definition: A written agreement by the program participants to voluntarily participate in an evaluation or study after having been advised of the purpose of the study, the type of the information being collected, and how information will be used.
Innovation Analysis

Term
Definition: Framework for answering a series of questions throughout the life cycle of an innovation-from pilot testing to broad-scale application in order to assess program successes, obstacles, and lessons learned.
Input Definition: Organizational units, people, dollars, and other resources actually devoted to the particular program or activity.
Instrument Definition: A tool used to collect and organize information.
Instrumental Outcome Definition: A measure or measures of phenomena directly related to program goals and objectives.
Instrumentation Bias Definition: Bias introduced in a study by a change in the measurement instrument during the course of the study.
Intermediate Outcome Definition: Results or outcomes of program activities that must occur prior to the final outcome in order to produce the final outcome.
Internal Consistency Definition: The extent to which all items in a scale or test measure the same concept.

Term
Internal Evaluation Definition: Evaluation conducted by a staff member or unit from within the organization being studied.
Internal Resource Definition: An agency's or organization's resources, including staff skills and experience and any information already available through current program activities.
Internal Validity Definition: The extent to which the causes of an effect are established by an inquiry.
Internal Validity Threat Definition: Factors other than program participation that may affect the results or findings.
Interquartile Range Definition: A measure of spread; a statistic used with ordinal, interval, and ratio variables.
Interrater Reliability Definition: The extent to which two different researchers obtain the same result when using the same instrument to measure a concept.
Interrupted Times Series Design

Term
Definition: The interrupted time series design involves repeated measurement of an indicator (e.g., reported crime) over time, encompassing periods both prior to and after implementation of a program. The goal of such an analysis is to assess whether the treatment (or program) has "interrupted" or changed a pattern established prior to the program's implementation. However, the impact of alternate historical events may threaten the interpretation of the findings.
Interval Estimate Definition: General term for an estimate of a population parameter that is a range of numerical values.
Interval Measure Definition: A quantitative measure with equal intervals between categories, but with no absolute zero.
Interval Scale Definition: A measurement scale that measures quantitative differences between values of a variable, with equal distances between the values.
Interval Variable Definition: A quantitative variable that attributes of which are ordered and for which the numerical differences between adjacent attributes are interpreted as equal.
Intervening Variable

Term
Definition: A variable that causally links other variables to each other. In a causal model, this intermediate variable must be influenced by one variable in order for a subsequent variable to be influenced.
Interviews Definition: Interviews involve face-to-face situations or telephone contacts in which the researcher orally solicits responses.
Judgment Sample Definition: A sample selected by using discretionary criteria rather than criteria based on the laws of probability.
Judgmental Forecasting Definition: Judgmental Forecasting attempts to elicit and synthesize informed judgments and are often based on arguments from insight.
Kaldor-Hicks Criterion Definition: A criterion of equity which states that one social state is better than another if there is a net gain in efficiency and if those that gain can compensate the losers.
Kendall's Tau Definition: A measure of association used to correlate two ordinal scales.
Kenneth Arrow's Impossibility Theorem

Term
Definition: A theorem to demonstrate that it is impossible to aggregate individual preferences through majority voting without violating one or more of five reasonable conditions of democratic decision-making.
Known-group Validation Definition: A procedure for validating an instrument which involves testing on a group for which the results are already known.
Kurtosis Definition: A term used to describe a curve indicating that it is more peaked than the normal curve.
Lambda Definition: A measure of association; a statistic used with nominal variables.
Lessons Learned Definition: Knowledge derived from the implementation and evaluation of a program that can be used to identify strengths and weaknesses of program design and implementation. This information is likely to be helpful in modifying and improving program functions in the future.
Level of Measurement Definition: Refers to the four levels of variables and their empirical attributes - nominal, ordinal, interval, and ratio.
Level of Significance

Term
Definition: The probability that observed or greater differences occur by chance.
Logic Model Definition: A diagram and text that describes and illustrates the logical (causal) relationships among program elements and the problem to be solved, thus defining measurements of success.
Longitudinal Data Definition: Observations collected over a period of time; the sample (instances or cases) may or may not be the same each time but the population remains constant.
Longitudinal Study Definition: The study of the same group over a period of time. These generally are used in studies of change.
Management Definition: The guidance and control of action required to execute a program. Also, the individuals charged with the responsibility of conducting a program.
Management Information System Definition: An information collection and analysis system, usually computerized, that facilitates access to program and participant information. It is usually designed and used for administrative purposes.
Marginal Distribution

Term
Definition: The distribution of a single variable based upon an underlying distribution of two or more variables.
Matching Definition: A method utilized to create comparison groups, in which groups or individuals are matched to those in the treatment group based on characteristics felt to be relevant to program outcomes.
Matrix of Categories Definition: A method of displaying relationships among themes in analyzing case study data that shows whether changes in categories or degrees along one dimension are associated with changes in the categories of another dimension.
Maturation Effect Definition: A threat to the internal validity of an evaluation in which observed outcomes are a result of natural changes of the program participants over time rather than because of program impact.
Mean Definition: A measure of central tendency, the arithmetic average; a statistic used primarily with interval-ratio variables following symmetrical distributions.
Measurement Definition: A procedure for assigning a number to an object or an event.

Term
Measurement Error Definition: The difference between a measured value and a true value.
Measures of Association Definition: Statistics that indicate the strength and nature of a relationship between variables.
Median Definition: A measure of central tendency, the value of the case marking the midpoint of an ordered list of values of all cases; a statistic used primarily with ordinal variables and asymmetrically distributed interval-ratio variables.
Meta-analysis Definition: The systematic analysis of a set of existing evaluations of similar programs in order to draw general conclusions, develop support for hypotheses, and/or produce an estimate of overall program effects.
Methodology Definition: The way in which information is found or something is done. The methodology includes the methods, procedures, and techniques used to collect and analyze information.
Mission Definition: The part of a goal or endeavor assigned as a specific responsibility of a particular organizational unit. It includes the task,

Term
together with the purpose, which clearly indicates the action to be taken and the reasons.
Mode Definition: A measure of central tendency, the value of a variable that occurs most frequently; a statistic used primarily with nominal variables.
Monitoring Definition: An on-going process of reviewing a program's activities to determine whether set standards or requirements are being met.
Monitoring System Definition: An on-going system to collect data on a program's activities and outputs, designed to provide feedback on whether the program is fulfilling its functions, addressing the targeted population, and/or producing intended services.
Multivariate Analysis Definition: An analysis of the relationships between more than two variables.
Needs Assessment Definition: A systematic process for gathering information about current conditions within a group that underlie the need for an intervention.
Nominal Variable

Term
Definition: A quantitative variable whose attributes have no inherent order.
Non-experimental Data Definition: Data not produced by an experiment or quasi-experiment.
Non-probability Sample Definition: A sample not produced by a random process.
Non-respondent Definition: A person who fails to answer either a questionnaire or a question.
Non-response Bias Definition: The bias created by the failure of part of a sample to respond to a survey or answer a question. If those responding and those not responding have different characteristics, the responding cases may not be representative of the population from which they were sampled.
Nonequivalent Comparison Group Design Definition: Evaluation designs that use nonrandomized comparison groups to evaluate program effects.
Normal Distribution (Curve)

Term
Definition: A theoretical distribution that is closely approximated by many actual distribution of variables.
Normative Question Definition: A type of evaluation question requiring comparison between what is happening (the condition) to norms and expectations or standards for what should be happening (the criterion).
Null Hypothesis Definition: A hypothesis stating that two variables are not related. Research attempts to disprove the null hypothesis by finding evidence of a relationship.
Objective Definition: Specific results or effects of a program's activities that must be achieved in pursuing the program's ultimate goals.
Observation Definition: A data collection strategy in which the activities of subjects are visually examined. The observer attempts to keep his/her presence from interfering in or influencing any behaviors.
Observational Techniques Definition: Data collection strategies which use observation of subjects as a means to collect data. These techniques generally involve attempts by the observer to not alter or change the behavior being observed.
One-group Designs

Term
Definition: Research designs which study a single program with no comparison or control group.
One-shot Case Study Definition: The one-shot case study involves the measurement of an identified "outcome" after a treatment or program has been implemented. However, there are no measures taken or available for comparison (i.e., status before the program, or outcome of a comparison or control group). Without a comparison measure, there is no means for inferring that the "outcome" was actually influenced by the treatment or program.
Open-ended Interview Definition: An interview in which, after an initial or lead question, subsequent questions are determined by topics brought up by the person being interviewed; the concerns discussed, their sequence, and specific information obtained are not predetermined and the discussion is unconstrained, able to move in unexpected directions.
Open-ended Question
Operational Definition Definition: Detailed description of how a concept or variable will be measured and how values will be assigned.
Operational Plan Definition: A tactical statement of when and what critical milestones must be passed to attain objectives programmed for a specific

Term
period.
Operationalize Definition: To define a concept in a way that can be measured. In evaluation research, to translate program inputs, outputs, objectives, and goals into specific measurable variables.
Opportunity Costs Definition: The value of opportunities forgone because of an intervention program.
Ordinal Scale Data Definition: Data classified into exhaustive, mutually exclusive, and ordered or ranked categories.
Ordinal Variable Definition: A quantitative variable whose attributes are ordered but for which the numerical differences between adjacent attributes are not necessarily interpreted as equal.
Outcome Definition: Changes or benefits resulting from activities and outputs. Short-term outcomes produce changes in learning, knowledge, attitude, skills or understanding. Intermediate outcomes generate changes in behavior, practice or decisions. Long-term outcomes produce changes in condition.
Outcome Evaluation

Term
Definition: This form of evaluation assesses the extent to which a program achieves its outcome-oriented objectives. It focuses on outputs and outcomes (including unintended effects) to judge program effectiveness but may also assess program process to understand how outcomes are produced.
Outlier Definition: Instances that are aberrant or do not fit with other instances: instances that, compared to other members of a population, are at the extremes on relevant dimensions.
Output Definition: Product or service delivery/implementation targets you aim to produce.
Outside Evaluator Definition: An evaluator not affiliated with the agency prior to the program evaluation.
Panel Data Definition: A special form of longitudinal data in which observations are collected on the same sample of respondents over a period of time.
Panel Interviewing Definition: Conducting repeated interviews with the same group of respondents over time.

Term
Parameter Definition: A number that describes a population.
Participant Observation Definition: A research method involving direct participation of the researcher in the events being studied. The researcher may either reveal or hide the true reason for involvement.
Participatory Evaluation Definition: An evaluation organized as a team project in which the evaluator and representatives of one or more stakeholder groups work collaboratively in developing the evaluation plan, conducting the evaluation, or disseminating and using the results.
Pearson Product-Moment Correlation Coefficient Definition: A measure of association; a statistic used with interval-ratio variables.
Peer Review Definition: An assessment of a product conducted by a person or persons of similar expertise to the author.
Performance Evaluation Definition: An evaluation that compares actual performance with that planned in terms of both resource utilization and production. It is used by management to redirect program efforts and resources and to redesign the program structure.

Term
Performance Measure Definition: Ways to objectively measure the degree of success a program has had in achieving its stated objectives, goals, and planned program activities.
Performance Measurement Definition: The ongoing monitoring and reporting of program accomplishments, particularly progress toward pre-established goals. It is typically conducted by program or agency management. Performance measures may address the type or level of program activities conducted (process), the direct products and services delivered by a program (outputs), or the results of those products and services (outcomes).
Pilot Definition: A pretest or trial run of a program, evaluation instrument, or sampling procedure for the purpose of correcting any problems before it is implemented or used on a larger scale.
Pilot Test Definition: Preliminary test or study of the program or evaluation activities to try out procedures and make any needed changes or adjustments.
Planning Definition: The process of anticipating future occurrences and problems, exploring their probable impact, and detailing policies, goals, objectives, and strategies to solve the problems. This often includes preparing options documents, considering alternatives, and

Term
issuing final plans.
Point Biserial Coefficient Definition: A measure of association between an interval-ratio variable and a nominal variable with two attributes.
Point Estimate Definition: An estimate of a population parameter that is a single numerical value.
Policy Analysis Definition: An analysis used to help managers understand the extent of the problem or need that exists and to set realistic goals and objectives in response to such problem or need. It may be used to compare actual program activities with the program's legally established purposes in order to ensure legal compliance.
Population Definition: The total number of individuals or objects being analyzed or evaluated.
Posttest Definition: A test or measurement taken after services or activities have ended. It is compared with the results of a pretest to show evidence of the effects or changes resulting from the services or activities being evaluated.
Precision

Term
Definition: The exactness of a question's wording or the amount of random error in an estimate.
Pretest Definition: A test or measurement taken before services or activities begin. It is compared with the results of a posttest to show evidence of the effects of the services or activities being evaluated. A pretest can be used to obtain baseline data.
Primary Data Definition: Data collected by the researcher specifically for the research project.
Probability Distribution Definition: A distribution of a variable that expresses the probability that particular attributes or ranges of attributes will be, or have been observed.
Probability Sample Definition: A group of cases selected from a population by a random process. Every member of the population has a known, nonzero probability of being selected.
Probability Sampling Definition: A method for drawing a sample from a population such that all possible samples have a known and specified probability of being drawn.
Probe

Term
Definition: To examine a subject in an interview in depth, using several questions.
Process Definition: The programmed, sequenced set of things actually done to carry out a program mission.
Process Evaluation Definition: This form of evaluation assesses the extent to which a program is operating as it was intended. It typically assesses program activities' conformance to statutory and regulatory requirements, program design, and professional standards or customer expectations.
Productivity Definition: The relationship between production of an output and one, some, or all of the resource inputs used in accomplishing the assigned task. It is measured as a ratio of output per unit of input over time. It is a measure of efficiency and is usually considered as output per person-hour.
Program Definition: Any activity, project, function, or policy that has an identifiable purpose or set of objectives.
Program Activities Definition: Activities, services, or functions carried out by the program (i.e., what the program does).

Term
Program Analysis Definition: The analysis of options in relation to goals and objectives, strategies, procedures, and resources by comparing alternatives for proposed and ongoing programs. It embraces the processes involved in program planning and program evaluation.
Program Effectiveness Evaluation Definition: The application of scientific research methods to estimate how much observed results, intended or not, are caused by program activities. Effect is linked to cause by design and analyses that compare observed results with estimates of what might have been observed in the absence of the program.
Program Evaluation Definition: Individual systematic studies conducted periodically or on an ad hoc basis to assess how well a program is working. They are often conducted by experts external to the program, either inside or outside the agency, as well as by program managers.
Program Failure Definition: A program shortcoming in which the outcome criteria are not affected by participation of the subjects in the program (i.e., the program does not accomplish its objective).
Program Implementation Objective Definition: What is planned to be done in the program, components, or services.
Program Justification

Term
Definition: The narrative and related analyses and statistical presentations supporting a program budget request. It includes: (1) definitions of program objectives, including a rationale for how the proposed program is expected to help solve the problem and the magnitude of the need, (2) plans for achieving the objectives , and (3) the derivation of the requested appropriation in terms of outputs or workloads showing productivity trends and the distribution of funds among organizational units.
Program Model Definition: A flowchart or model which identifies the objectives and goals of a program, as well as their relationship to program activities intended to achieve these outcomes.
Program Theory Definition: The set of assumptions about the manner in which the program relates to the social benefits it is expected to produce and the strategy and tactics the program has adopted to achieve its goals and objectives. Within program theory we can distinguish impact theory, relating to the nature of the change in social conditions brought about by program action, and process theory, which depicts the program's organizational plan and service utilization plan.
Purposive Sample Definition: Instances appropriately selected to answer different evaluation questions, on various systematic bases, such as best or worst practices; a judgmental sample. If conducted systematically, can be widely useful in evaluation.
Qualitative Analysis Definition: An analysis that ascertains the nature of the attributes, behavior, or opinions of the entity being measured.
Qualitative Data

Term
Definition: Information that is difficult to measure, count, or express in numerical terms.
Qualitative Research Definition: Research involving detailed, verbal descriptions of characteristics, cases, and settings. Qualitative research typically uses observation, interviewing, and document review to collect data.
Quality Control Definition: A procedure for keeping quality of inputs or outputs to specifications.
Quantify Definition: To attach numbers to an observation.
Quantitative Analysis Definition: An analysis that ascertains the magnitude, amount, or size, for example, of the attributes, behavior, or opinions of the entity being measured.
Quantitative Data Definition: Information that can be expressed in numerical terms, counted, or compared on a scale.
Quantitative Research

Term
Definition: Research that examines phenomenon through the numerical representation of observations and statistical analysis.
Quasi-experimental Design Definition: A research design with some, but not all, of the characteristics of an experimental design. While comparison groups are available and maximum controls are used to minimize threats to validity, random selection is typically not possible or practical.
Questionnaire Definition: Research instrument that consists of statistically useful questions.
Quota Sampling Definition: A nonprobability stratified sampling procedure in which units are selected for the sample to adhere to certain proportions of characteristics desired.
Random Assignment Definition: The assignment of individuals in the pool of all potential participants to either the experimental (treatment) group or the control group in such a manner that their assignment to a group is determined entirely by chance.
Random Comparison Group Design Definition: In this research design, the comparison group is randomly selected from the population of interest, even though the treatment group is not selected randomly.
Random Sampling

Term
Definition: A procedure for sampling from a population that gives each unit in the population a known probability of being selected into the sample.
Randomized Comparative Change Design Definition: In the experimental design known as the randomized comparative change design a treatment and control group are randomly selected for study. Both groups are administered a pre-test. The treatment group is given the treatment, while the control group is not. Both groups are tested or measured after the treatment. The test results of the two groups are compared. The pretest allows a check on the randomization process, and allows for control of any differences found.
Randomized Comparative Post-test Design Definition: In the experimental design known as the randomized comparative post-test design a treatment and control group are randomly selected for study. The treatment group is given the treatment, while the control group is not. Both groups are tested or measured after the treatment. The test results of the two groups are compared.
Range Definition: A measure of spread which gives the distance between the lowest and the highest values in a distribution; a statistic used primarily with interval-ratio variables.
Ratio Measure Definition: A level of measurement which has all the attributes of nominal, ordinal, and interval measures, and is based on a "true zero" point. As a result, the difference between two values or cases may be expressed as a ratio.

Term
Regression Analysis Definition: A method for determining the association between a dependent variable and one or more independent variables.
Regression Coefficient Definition: An asymmetric measure of association; a statistic computed as part of a regression analysis.
Regression Effects Definition: The tendency of subjects, who are initially selected due to extreme scores, to have subsequent scores move inward toward the mean.
Regression Fallacy Preferred Term: Regression Effects
Regression to the Mean Preferred Term: Regression Effects
Reliability Definition: The extent to which a measurement instrument yields consistent, stable, and uniform results over repeated observations or measurements under the same conditions each time.
Reliability Assessment

Term
Definition: An effort required to demonstrate the repeatability of a measurement or how likely a question may be to get consistently similar results. It is different from verification (checking accuracy) or validity.
Replication Definition: The duplication of an experiment or program.
Representative Definition: Reflecting the characteristics or nature of the larger population to which one wants to generalize.
Representative Sample Definition: A sample that has approximately the same distribution of characteristics as the population from which it was drawn.
Research Design Definition: A plan of what data to gather, from whom, how and when to collect the data, and how to analyze the data obtained.
Resistant Statistic Definition: A statistic that is not much influenced by changes in a few observations.
Resources Definition: Assets available and anticipated for operations. They include people, equipment, facilities and other things used to plan,

Term
implement, and evaluate public programs whether or not paid for directly by public funds.
Response Rate Definition: The percentage of persons in a sample who respond to a survey.
Response Style Definition: The tendency of a respondent to answer in a specific way regardless of how a question is asked.
Response Variable Definition: A variable on which information is collected and which there is an interest because of its direct policy relevance.
Rolling Design Preferred Term: Emergent Design
Sample Definition: A subset of the population. Elements are selected intentionally as a representation of the population being studied.
Sample Design Definition: The sampling procedure used to produce any type of sample.
Sampling Distribution

Term
Definition: The distribution of a statistic.
Sampling Error Definition: The maximum expected difference between a probability sample value and the true value.
Scale Definition: An aggregate measure that assigns a value to a case based on a pattern obtained from a group of related measures.
Scientific Sample Definition: A group of cases selected from a population by a random process. Every member of the population has a known, nonzero probability of being selected.
Scoping Definition: Analyzing alternative ways for conducting an evaluation. It is clarifying the validity of issues, the complexity of the assignment, the users of final reports, and the selection of team members to meet the needs of an evaluation. Scoping ends when a major go/no-go decision is made about whether to do the evaluation.
Secondary Data Definition: Data that has been collected for another purpose, but may be reanalyzed in a subsequent study.
Selection Bias

Term
Definition: Potential biases introduced into a study by the selection of different types of people into treatment and comparison groups. As a result, the outcome differences may potentially be explained as a result of pre-existing differences between the groups, as opposed to the treatment itself.
Self-evaluation Definition: The evaluation of a program by those conducting the program.
Self-Reported Data Definition: Information that program participants generate themselves that is used to assess program processes or outcomes.
Significance Level Definition: The probability of rejecting a set of assumptions when they are in fact true.
Simple Random Sample Definition: A method for drawing a sample from a population such that all samples of a given size have equal probability of being drawn.
Spread Definition: General term for the extent of variation among cases.
Stakeholder

Term
Definition: An individual or organization with a direct or indirect investment in a project or program (e.g., program champion, community leader, etc.).
Standard Definition: A criterion for evaluating performance and results. It may be a quantity or quality of output to be produced, a rule of conduct to be observed, a model of operation to be adhered to, or a degree of progress toward a goal.
Standard Deviation Definition: A measure of the spread, the square root of the variance; a statistic used with interval-ratio variables.
Standard Instruments Definition: An assessment, inventory, questionnaire, or interview that has been tested with a large number of individuals and is designed to be administered to program participants in a consistent manner. Results of tests with program participants can be compared to reported results of the tests used with other populations.
Standardized Question Definition: A question that is designed to be asked or read and interpreted in the same way regardless of the number and variety of interviewers and respondents.
Statistic Definition: A number computed from data on one or more variables.

Term
Statistical Analysis Definition: Analyzing collected data for the purposes of summarizing information to make it more usable and/or making generalizations about a population based on a sample drawn from that population.
Statistical Control Definition: A statistical technique used to eliminate variance in dependent variables caused by extraneous sources. In evaluation research, statistical controls are often used to control for possible variation due to selection bias by adjusting data for program and control group on relevant characteristics.
Statistical Procedure Definition: A set of standards and rules based in statistical theory by which one can describe and evaluate what has occurred.
Statistical Regression Preferred Term: Regression Effects
Statistical Sample Definition: Synonymous with probability sample; a group of cases selected from a population by a random process in which every member of the population has a known, nonzero probability of being selected.
Statistical Significance

Term
Definition: The degree to which a value is greater or smaller than would be expected by chance. Typically, a relationship is considered statistically significant when the probability of obtaining that result by chance is less than 5% if there were, in fact, no relationship in the population.
Statistical Test Definition: Type of statistical procedure that is applied to data to determine whether the results are statistically significant (that is, the outcome is not likely to have resulted by chance alone.)
Statistical Weighting Definition: A technique used to assure representation of certain groups in the sample. Data for underrepresented cases are weighted to compensate for their small numbers, making the sample a better representation of the underlying population.
Strategic Evaluation Definition: An evaluation used by managers as an aid to decide which strategy a program should adopt in order to accomplish its goals and objectives at a minimum cost. In addition, strategy evaluation might include alternative specifications of the program design itself, manpower specifications, progress objectives, and budget allocations.
Strategic Plan Definition: The process of comprehensive, integrative program planning that considers, at a minimum, the future of current decisions, overall policy, organizational development, and links to operational plans.
Stratified Random Sampling

Term
Definition: A sampling procedure for which the population is first divided into strata or subgroups based on designated criteria and then the sample is drawn, either proportionately or disproportionately, from each subgroup.
Structured Interview Definition: An interview in which questions to be asked, their sequence, and detailed information to be gathered are all predetermined; used where maximum consistency across interviews and interviewees is needed.
Summative Evaluation Definition: 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.
Supplementary Variable Definition: A variable upon which information is collected because of its potential relationship to a response variable.
Survey Definition: The collection of information from a common group through interviews or the application of questionnaires to a representative sample of that group.
Systematic Sample Definition: A sample drawn by taking every nth case from a list, after starting with a randomly selected case among the first individuals.

Term
Target Definition: An objective (constraint or expected result) set by management to communicate program purpose to operating personnel (for example, maintaining a monthly output level).
Target Population Definition: The population, clients, or subjects intended to be identified and served by the program.
Test-retest Definition: Administration of the same test instrument twice to the same population for the purpose of assuring consistency of measurement.
Testing Bias Definition: Bias and foreknowledge introduced to participants as a result of a pretest. The experience of the first test may impact subsequent reactions to the treatment or to retesting.
Theory Failure Definition: The program is implemented as planned but its services do not produce the immediate effects on the participants that are expected, the ultimate social benefits as intended, or both.
Third-party Evaluator

Term
Preferred Term: Outside Evaluator
Time-series Data Preferred Term: Longitudinal Data
Time-series Designs Definition: Research designs that collect data over long time intervals - before, during, and after program implementation. This allows for the analysis of change in key factors over time.
Transformed Variable Definition: A variable for which the attribute values have been systematically changed for the sake of data analysis.
Treatment Group Definition: The subjects of the intervention being studied.
Treatment Variable Definition: An independent variable in program evaluation that is of particular interest because it corresponds to a program's intent to change some dependent variable.
Trend Definition: The change in a series of data over a period of years that remains after the data have been adjusted to remove seasonal

Term
and cyclical fluctuations.
Triangulation Definition: The combination of methodologies in the study of the same phenomenon or construct; a method of establishing the accuracy of information by comparing three or more types of independent points of view on data sources (for example, interviews, observation, and documentation; different times) bearing on the same findings.
Unit of Analysis Definition: The class of elemental units that constitute the population and the units selected for measurement; also, the class of elemental units to which the measurements are generalized.
Univariate Analysis Definition: An analysis of a single variable.
Unobtrusive Measures Definition: Any method of data collection in which the subjects are not aware that they are being studied.
Utility Definition: The extent to which an evaluation produces and disseminates reports that inform relevant audiences and have beneficial impact on their work.
Validity

Term
Definition: The extent to which a measurement instrument or test accurately measures what it is supposed to measure.
Validity Assessment Definition: The procedures necessary to demonstrate that a question or questions are measuring the concepts that they were designed to measure.
Variable
Variance Definition: A measure of the spread of the values in a distribution. The larger the variance, the larger the distance of the individual cases from the group mean.
Verification Definition: An effort to test the accuracy of the questionnaire response data. The concern is uniquely with data accuracy and deals with neither the reliability nor the validity of measures.
Z Scores Definition: Standard deviation units measuring the deviation from the mean relative to the standard deviation.