

Child Care & Early Education RESEARCH CONNECTIONS

<http://www.researchconnections.org>

Research Connections Research Glossary

The research glossary defines terms used in conducting social science and policy research, for example those describing methods, measurements, statistical procedures, and other aspects of research.

Accuracy

In survey research, accuracy refers to the match between a sample and the target population. It also indicates how close a value obtained from a survey instrument or assessment is to the actual (true) value.

Action Research

Action research conducted to solve problems, inform policy, or improve the way that issues are addressed and problems solved. There are two broad types of action research: participatory action research and practical action research.

Adjusted R-Squared

A measure of how well the independent, or predictor, variables predict the dependent, or outcome, variable. A higher adjusted R-square indicates a better model. Adjusted R-square is calculated based on the R-square, which denotes the percentage of variation in the dependent variable that can be explained by the independent variables. The adjusted R-squared adjusts the R-square for the sample size and the number of variables in the regression model. Therefore, the adjusted R-square is a better comparison between models with different numbers of variables and different sample sizes.

Administrative Data

Administrative data are used in support of the operations and service delivery of government departments and other organizations. Examples are information about individual children, families, and/or providers of early care and education and other family benefits and services. The data are collected and maintained primarily for administrative (not research) purposes.

Alpha Level

The probability that a statistical test will find significant differences between groups (or find significant predictors of the dependent variable), when in fact there are none. This is also referred to as the probability of making a Type I error or as the significance level of a statistical test. A lower alpha level is better than a higher alpha level, with all else equal.

Alternative Hypothesis

The experimental hypothesis stating that there is some real difference between two or more groups. It is the alternative to the null hypothesis, which states that there is no difference between groups.

Analysis of Covariance (ANCOVA)

Is an advanced form of ANOVA (analysis of variance). While ANOVA is used to test for differences in the means of two or more groups, ANCOVA removes the effect(s) of one or more continuous variable(s) before testing for the group differences. For example, in an analysis that

examines differences in child outcomes by type of child care, the analyst may want to remove the effects of parental education.

Analysis of Variance (ANOVA)

A statistical test that determines whether the means of two or more groups are significantly different.

Anonymity

An ethical safeguard against invasion of privacy whereby the researcher is unable to identify the respondents by their responses.

Association

The relationship between objects or variables. Two variables are positively associated when the values of one increase as the values of the other increases. They are negatively associated when the values of one decrease as the values of the other increase. Income and education are usually positively associated and student absenteeism is generally negatively associated with student achievement.

Attrition

Attrition occurs when participants drop out of a longitudinal study or panel study over time. If particular types of study participants drop out at a higher rate than other types of participants, attrition can introduce bias that can potentially threaten the internal and external validity of a longitudinal study and the internal validity of an experiment with treatment and control groups.

Average

A single value (mean, median, mode) representing the typical, normal, or middle value of a set of data.

Average Treatment Effect

The average treatment effect (ATE) measures the difference in the mean (average) outcome between the individuals or other units (e.g., classrooms, schools) assigned to the treatment and those assigned to the control. For example, in a study of the effects of a preschool reading intervention, the ATE would be the difference in average reading scores for children who received the intervention (treatment group) and the average reading scores for those who did not (control group).

Axiom

A statement widely accepted as truth.

Bar Chart/Graph

Bar charts are used by researchers to visually represent the frequencies or percentages with which different categories of a variable occur. They are most often used when describing and comparing the percentages of different groups with a specific characteristic. For example, the percentages of boys and girls who participate in team sports. However, they may also be used when describing averages such as the average time boys and girls spend per week participating in team sports. A bar is drawn for each of the groups along the horizontal axis and the height of the bar corresponds to the frequency or percentage with which the characteristic occurs (vertical axis).

Bayesian Statistics

Bayesian statistics is a general approach to estimating population parameters (characteristics) that uses both information about the prior distribution of the parameter of interest along with new evidence (likelihood function). In Bayesian statistics, the posterior probability distribution is the probability distribution once all information is taken into account.

Bell-Shaped Curve

A curve characteristic of a normal distribution, which is symmetrical about the mean and extends infinitely in both directions. The mean (average) is always in the center of the bell or normal curve. One half of the data points are to the left and one half are to the right of the mean.

Beta Level

The probability of making an error when comparing groups and stating that differences between the groups are the result of the chance variations when in reality the differences are the result of the experimental manipulation or intervention. Also referred to as the probability of making a Type II error.

Between-Group Variance

A measure of the difference between the means of various groups.

Between-Subject Design

Experimental design in which a different group of subjects are used for each level of the variable under study.

Bias

Influences that distort the results of a research study.

Bimodal Distribution

A distribution in which two scores or values are the most frequently occurring. Interpreting the average of a bimodal distribution is problematic because the data are not normally distributed. Identifying bimodal distributions is done by examining a frequency distribution or by looking at indices of skew or kurtosis, which are frequently available with statistical software packages.

Bootstrapping

A popular method for variance estimation in surveys. It consists of subsampling from the initial sample. Within each stratum in the sample, a simple random subsample is selected with replacement. This creates a finite number of new samples (or repetitions). The same parameter estimate is then calculated for each of the subsamples. The variance of the estimated parameter is then equal to the variance of the estimates from these subsamples.

Canonical Correlation Analysis

Canonical correlation analysis is used to examine the associations between multiple independent variables and two or more intercorrelated dependent (outcome) variables. That is, it is used in situations where multiple regression would be used, but where there are multiple dependent variables that are correlated with each other. For example, a researcher might be interested in the associations between children's race/ethnicity, gender and family SES and their performance of several measures of academic achievement (e.g., math, reading, science).

Case Study

An intensive investigation of the current and past behaviors and experiences of a single person, family, group, or organization.

Categorical Data

Variables with discrete, non-numeric or qualitative categories (e.g. gender or marital status). The categories can be given numerical codes, but they cannot be ranked, added, multiplied or measured against each other. Also referred to as nominal data.

Categorical Data Analysis

Categorical data classify responses or observations into discrete categories (e.g., respondents' highest level of education is often classified as less than high school, high school, college, and post-graduate). While there are many techniques for analyzing such data, 'categorical data analysis' usually refers to the analysis of one or more categorical dependent variables and the relationships to one or more predictor variables (e.g., logistic regression).

Causal Analysis

An analysis that seeks to establish the cause and effect relationships between variables.

Causal Inference with Interference

One assumption of randomized experiments is that a subject's response to treatment (intervention) depends only on the treatment to which the subject is assigned, not on the treatment assignments of other subjects. The ability to draw causal inferences from the findings of experimental studies rests on the assumption that there is no such interference. However, such interference is common in experimental studies where there is a high level of interaction between subjects, such as studies of different curricula where teachers in the same school are assigned to different treatment groups.

Ceiling

The highest limit of performance that can be assessed or measured by an instrument or process. Individuals who perform near to or above this upper limit are said to have reached the ceiling, and the assessment may not be providing a valid estimate of their performance levels.

Census

The collection of data from all members, instead of a sample, of the target population.

Central Limit Theorem

A mathematical theorem that is central to the use of statistics. It states that for a random sample of observations from any distribution with a finite mean and a finite variance, the mean of the observations will follow a normal distribution. This theorem is the main justification for the widespread use of statistical analyses based on the normal distribution.

Central Tendency

A measure that describes the "typical" or average characteristic; the three main measures of central tendency are mean, median and mode.

Chi Square

A statistic used when testing for associations between categorical, or non-numeric, variables. It

is also used as a goodness-of-fit test to determine whether data from a sample come from a population with a specific distribution.

Chi-Square Test

There are several different Chi-square tests in statistics. One of the more commonly used is the Chi-square test of independence. It is used to determine if there is a statistically significant association between two categorical variables. The frequency of each category for one variable is compared across the categories of the second variable, such as in a $n \times n$ cross tabulation. It is the null hypothesis for this test that there is no association between the two variables (i.e., the distributions of the two variables are independent of each other). The alternative hypothesis is that there is an association. For example, a Chi-square test could be used to examine whether parents' decision to delay their children's entry to kindergarten (delay vs. do not delay) is statistically significantly associated with their child's sex (male vs. female).

Cluster Analysis

Cluster analysis is a multivariate method used to classify a sample of subjects (or objects) in such a way that subjects in the same group (called a cluster) are more similar (e.g., in terms of their personal attributes, beliefs, preferences) to each other than to those in other groups (clusters).

Cluster Sampling

A type of sampling method where the population is divided into groups, called clusters. Cluster designs are often used to control costs. For example, researchers first randomly select clusters of potential respondents, and then respondents are selected at random from within the pre-identified clusters. The researcher randomly selects several counties or groups of counties and then draws a random sample of households from within the selected counties. Cluster sampling is often used in education and early childhood research. Researchers sample schools/programs and then students/children enrolled in the selected schools/programs. Clustered sampling designs necessitate the use of special variance estimation techniques.

Codebook

Information on the structure, content, and layout of a data set. The codebook typically provides background on the project, describes the study design, and gives detailed information on variable names and variable value codes. User's manuals and user's guides are examples of codebooks.

Codes

Values, typically numeric, that are assigned to different levels of variables to facilitate analysis of the variable. For example, codes such as strongly disagree=1, disagree=2, agree=3, and strongly agree=4 are often assigned.

Coding

The process of assigning values, typically numeric values, to the different levels of a variable. The process of assigning values to behaviors observed in parent-child interactions and assigning numeric values to responses to open-ended survey questions are examples of coding.

Coefficient of Determination

A coefficient, ranging between 0 and 1, that indicates the goodness of fit of a regression model.

Cognitive Interviewing

A research method used to pretest interview questions or items on a questionnaire. Cognitive interviews collect information on how respondents answer questions, their interpretation of the questions asked and their reasons for responding in a particular way. Additional verbal information is collected to evaluate whether respondents understand a question or series of questions, if the response categories are appropriate and if the question is measuring the construct it was designed to measure. The information gathered from the cognitive interview is used to make adjustments to questions before they are administered to the full sample.

Cohort

A group of people sharing a common demographic experience who are observed through time. For example, all the people born in the same year constitute a birth cohort. All the people married in the same year constitute a marriage cohort.

Comparability

The quality of two or more objects that can be evaluated for their similarity and differences.

Completion Rate

In survey research, this is the number of people who answered a survey divided by the number of people in the sample. It is sometimes used interchangeably with response rate.

Conditional Probability Models

Conditional probability models are a class of statistical models that are used to study the probability of an outcome given some prior event(s) or characteristic(s). For example, researchers might use such a model to examine the probability that children attend an early education program at age three given their prior child care history (attended child care program at age 2) and their mothers' employment (employed full-time versus part-time).

Confidence Interval

A range of estimated values that is the best guess as to the true population's value. Confidence intervals are usually calculated for the sample mean. In behavioral research, the acceptable level of confidence is usually 95%. Statistically, this means that if 100 random samples were drawn from a population and confidence intervals were calculated for the mean of each of the samples, 95 of the confidence intervals would contain the population's mean. For example, a 95% confidence interval for IQ of 95 to 105, indicates with 95% certainty that the actual average IQ in the population lies between 95 and 105.

Confidence Level

The percentage of times that a confidence interval will include the true population value. If the confidence level is .95 this means that if a researcher were to randomly sample a population 100 times, 95% of the time the estimated confidence interval for a value will contain the population's true value. In other words, the researcher can be 95% confident that the confidence interval contains the true population value.

Confidentiality

The protection of research subjects from being identified. A common standard in social science research is that records or information used for research should not allow participants to be identified and that researchers should not take any action that would affect the individual to whom the information pertains.

Confirmatory Factor Analysis

Confirmatory factor analysis (CFA) is a special form of factor analysis that is used to test whether the data fit a hypothesized measurement model. It tests how well the data fits a predetermined factor structure. For example, a researcher might use CFA to test whether the data from his/her sample fit the factor structure of an existing measure of parental stress.

Confounding Variable

A variable that is not of interest, but which distorts the results if the researcher does not control for it in the analysis. For example, if a researcher is interested in the effect of education on political views, the researcher must control for income. Income is a confounding variable because it affects political views and education is related to income.

Consistency

Answers to a set of questions are consistent if they do not contain any logical contradictions.

Constant

A value that stays the same for all the units of an analysis. For instance, in a research study that explores fathers' involvement in their children's lives, gender would be constant, as all subjects (units of analysis) are male.

Construct

A construct or a concept is a theoretical creation that cannot be directly observed, but can be measured using one or more indicators or tests. Examples of constructs in early care and education research include classroom quality, professional development, reading and math achievement, and socioemotional development.

Construct Validity

The degree to which a variable, test, questionnaire or instrument measures the theoretical concept that the researcher hopes to measure. For example, if a researcher is interested in the theoretical concept of "marital satisfaction," and the researcher uses a questionnaire to measure marital satisfaction, if the questionnaire has construct validity it is considered to be a good measure of marital satisfaction.

Contamination

In an experimental study, contamination occurs when individuals or groups of individuals receive certain features of an intervention intended for a different experimental group. For example, students in an elementary school assigned at random to receive one of three accelerated math curricula, may receive parts of a curriculum that are different from the one to which they were assigned. Or, students who are randomly assigned to a control group in an experimental study of the effects of a new approach to teaching math (treatment) may be

exposed to some of the features of the treatment. In educational research, random assignment is often done at the classroom or school level to avoid such contamination.

Content Analysis

A procedure for organizing narrative, qualitative data into themes and concepts. It is used to describe and draw meaning from different forms of written (e.g., books, magazines, newspapers) and oral (e.g., audiotapes, television and videos) communications. Content analysis is used in both quantitative and qualitative research.

Content Validity

Content validity, like face validity, refers to whether a given test or other measurement tool (e.g., classroom observation protocol) actually measures the construct that it claims to measure. Content validity requires the use of recognized experts to evaluate whether test items or items in an observation protocol actually assess the defined content.

Context Conditionality

Context conditionality refers to those situations where the effect(s) of some independent variable(s) on some dependent variable(s) depend upon or are moderated by a third (set of) independent variable(s). One way that researchers address this is by including interaction terms for the independent variables in their models.

Context Effects

The change in the dependent variable which is resulted from the influence of the research environment. This influence is external to the experiment itself.

Continuous Variable

A variable that, in theory, can take on any value within a range. The opposite of continuous is discrete or categorical, which can have only a particular set of values. For example, a person's height could be 5 feet 1 inch, 5 feet 1.1 inches, 5 feet 1.11 inches, and so one, and thus it is continuous. The type of child care setting where a child spends the greatest number of hours each week (e.g., center-based care, relative and nonrelative care) is a discrete or categorical variable.

Control

The processes of making research conditions uniform or constant, so as to isolate the effect of the experimental condition. When it is not possible to control research conditions, statistical controls often will be implemented in the analysis.

Control Group

In an experiment, the control group does not receive the intervention or treatment under investigation. This group may also be referred to as the comparison group.

Control Variable

A variable that is not of interest to the researcher, but which interferes with the statistical analysis. In statistical analyses, control variables are held constant or their impact is removed to better analyze the relationship between the outcome variable and other variables of interest. For example, if one wanted to examine the impact of education on political views, a researcher

would control income in the statistical analysis. This removes the impact of income on political views from the analysis.

Controlled Experiment

A form of scientific investigation in which one variable, termed the independent variable, is manipulated to reveal the effect on another variable, termed the dependent or responding variable, while all other variables in the system are held fixed.

Convenience Sampling

A sampling strategy that uses the most easily accessible people (or objects) to participate in a study. This is not a random sample, and the results cannot be generalized to individuals who did not participate in the research.

Cooperation Rate

In survey research, this is the percentage of persons who answer a survey or complete an interview out of all persons who were contacted and asked to complete the survey or interview.

Correlation

The degree to which two variables are associated. Variables are positively correlated if they both tend to increase at the same time. For example, height and weight are positively correlated because as height increases weight also tends to increases. Variables are negatively correlated if as one increases the other decreases. For example, number of police officers in a community and crime rates are negatively correlated because as the number of police officers increases the crime rate tends to decrease.

Correlation Coefficient

A measure of the degree to which two variables are related. A correlation coefficient is always between -1 and +1. If the correlation coefficient is between 0 and +1 then the variables are positively correlated. If the correlation coefficient is between 0 and -1 then the variables are negatively correlated.

Covariate Balance

In randomized experiments, randomization creates covariate balance between the treatment (or treatment and control) groups. In such experiments, the characteristics of participants in the different groups are approximately equal. In observational studies (studies using nonexperimental designs) the characteristics of participants in different groups will rarely be equal (covariate unbalance). Researchers use different methods to achieve covariate balance when using observational data to test for treatment or causal effects (for example, propensity scores).

Coverage

Coverage reflects the extent to which all elements on a sampling frame (list) are members of the population to be sampled, and the extent to which every element in that population appears on the frame (list) once and only once. For example, the extent to which a list of child care providers includes (covers) all the providers in a given location (state, community).

Critical Incident Technique

A qualitative research method that collects observations or participant reports of behaviors that have critical significance. The observations or reports are recorded and tracked over a period of time. Observations/reports are placed into categories, summarized, and counted and used to solve practical problems and develop broad theories of human behavior.

Cross-Sectional Data

Coverage reflects the extent to which all elements on a sampling frame (list) are members of the population to be sampled, and the extent to which every element in that population appears on the frame (list) once and only once. For example, the extent to which a list of child care providers includes (covers) all the providers in a given location (state, community).

Cross-Tabulation

A method to display the relationship between two categorical variables. A table is created with the values of one variable across the top and the values of the second variable down the side. The number of observations that correspond to each cell of the table are indicated in each of the table cells.

Crossover Design

A crossover design (also called switch over or change over design) is a type of experimental design in which each research subject receives a sequence of experimental treatments over a number of time periods. A major advantage of the crossover design is that each subject serves as his or her own control. They often require a smaller number of participants than more traditional Randomized Control Trial designs.

Curvilinear

A statistical relationship between two variables that is not linear when plotted on a graph, but rather forms a curve.

Data

Information collected through surveys, interviews, or observations. Statistics are produced from data, and data must be processed to be of practical use.

Data Analysis

The process by which data are organized to better understand patterns of behavior within the target population. Data analysis is an umbrella term that refers to many particular forms of analysis such as content analysis, cost-benefit analysis, network analysis, path analysis, regression analysis, etc.

Data Collection

The observation, measurement, and recording of information in a research study.

Data Imputation

A method used to fill in missing values (due to nonresponse) in surveys. The method is based on careful analysis of patterns of missing data. Types of data imputation include mean imputation, multiple imputation, hot deck and cold deck imputation. Data imputation is done to allow for statistical analysis of surveys that were only partially completed.

Data Reduction

Data reduction is the process of transforming numerical or alphabetical digital information into a corrected, ordered, and simplified form. The basic concept is the reduction of large amounts of data down to the meaningful parts.

Deduction

The process of reasoning from the more general to the more specific.

Deductive Method

A method of study that begins with a theory and the generation of a hypothesis that can be tested through the collection of data, and ultimately lead to the confirmation (or lack thereof) of the original theory.

Degrees of Freedom

The number of independent units of information in a sample used in the estimation of a parameter or calculation of a statistic. The degrees of freedom limits the number variables that can be included in a statistical model. Models with similar explanatory power, but more degrees of freedom are generally preferred because they offer a simpler explanation.

Delphi Survey Method

The Delphi survey method is used when a researcher is interested in collecting opinions from a group of experts but wishes to limit the undue influence of participants upon one another (the members of the group are unknown to each other and questionnaires are often distributed via email). Both qualitative and quantitative research methods can be used in a Delphi survey. The experts answer questions in two or more rounds. After each round, a summary of the experts' responses is prepared by a member of the research team with expertise in the area(s) of inquiry and the questions the group is asked to respond to are revised.

Dependent Variable

The outcome variable. A dependent variable is something that depends on other factors. Researchers often try to find out what causes changes in the dependent variable. For example, in a study of factors associated with children's scores on standardized tests, children's scores would be the dependent variable.

Descriptive Statistics

Basic statistics used to describe and summarize data. Descriptive statistics generally include measures of the average values of variables (mean, median, and mode) and measures of the dispersion of variables (variance, standard deviation, or range).

Dichotomous Variables

Variables that have only two categories, such as whether a child is enrolled in a preschool program or not.

Differential Attrition

Differential or selective attrition occurs when the rates of dropping out or leaving a study with several data collection waves (e.g., longitudinal study or experimental research) vary across different study groups. This is particularly troublesome when the characteristics of those who

drop out are systematically different from those who remain, and may introduce bias in the study findings.

Direct Effect

The effect of one variable on another variable, without any intervening variables.

Direct Observation

A method of gathering data primarily through close visual inspection of a natural setting. Direct observation does not involve actively engaging members of a setting in conversations or interviews. Rather, the direct observer strives to be unobtrusive and detached from the setting.

Disconfirming Evidence

A procedure whereby, during an open-ended interview, a researcher actively seeks accounts from other respondents that differs from the main or consensus accounts in critical ways

Discrete Variables

A variable that can assume only a finite number of values; it consists of separate, indivisible categories. The opposite of discrete is continuous. For example, in a specified time period, a child's biological parents either live in the same household as the child or they do not. In contrast, a child's height could be 3 feet 1 inch, 3 feet 1.1 inches, 3 feet 1.11 inches, and so on, thus it is continuous variable.

Discriminant Analysis

A grouping method that identifies characteristics that distinguish between groups. For example, a researcher could use discriminant analysis to determine which characteristics identify families that seek child care subsidies and which identify families that do not.

Dispersion

In statistics, dispersion refers to the spread of a variable's values. Techniques that are used to describe dispersion include range, variance, standard deviation, and skew.

Distribution

The frequency with which values of a variable occur in a sample or a population. To graph a distribution, first the values of the variables are listed across the bottom of the graph. The number of times the value occurs are listed up the side of the graph. A bar is drawn that corresponds to how many times each value occurred in the data. For example, a graph of the distribution of women's heights from a random sample of the population would be shaped like a bell. Most women's height are around 5'4" This value would occur most frequently, so it would have the highest bar. Heights that are close to 5'4", such as 5'3" and 5'5" would have slightly shorter bars. More extreme heights, such as 4'7" and 6'1" would have very short bars.

Double Barreled Question

A survey question whereby two separate ideas are erroneously presented together in one question.

Double Blind Experiment

A research design where both the experimenter and the subjects are unaware of which is the treatment group and which is the control.

Double-Barreled Question

A double-barreled question asks two questions in one. For example, "Please rate how strongly you agree or disagree with the following statement: 'My child is happy at his/her child care center and the price is right for the type of care'." In the example, respondents are asked to give a single response to two different questions, one about their child's happiness and one about the price of care.

Dummy Coding

A coding strategy where a categorical variable is turned into a series of dichotomous variables (dummy variables) each of which have a value of 0 or 1. Dummy variables are used in regression analysis to measure the effect of a categorical variable on the outcome when the original categorical variable has more than 2 values. For a given categorical variable, $k-1$ dummy variables are entered into the regression equation, where k is the number of values in the original categorical variable.

Dummy Variables

Categorical variables that are assigned a value of 0 or 1 for use in a statistical analyses (see Dummy Coding).

Duration Models

A group of statistical models used to measure the length of a status or process.

Ecological Fallacy

False conclusions made by assuming that one can infer something about an individual from data collected about groups.

Econometrics

A field of economics that applies mathematical statistics and the tools of statistical inference to the empirical measurement of relationships postulated by economic theory.

Effect Size

A measure of the strength of the effect of the predictor (or independent) variable on the outcome (or dependent) variable. It is a measure of the magnitude or size of the difference between two or more groups. There are many ways to calculate an effect size, but one of the more common is to express differences in outcomes between groups in standard deviation units. For example, a researcher finds a significant difference in the math scores of African American and Hispanic first graders. Hispanic scores are 5 points higher than those of African American children. The effect size for this 5 point difference would be equal to .33 standard deviation given that the overall standard deviation of the assessment is 15 (effect size = score difference/standard deviation of assessment).

Egocentric Social Network Analysis

Egocentric social network analysis (SNA) is a methodological tool used to understand the structure, function, and composition of an individual's patterns of interactions in multiple social settings. It is used to study how people's patterns of interaction shape their individual-level outcomes (e.g. health, education, employment opportunities).

Endogeneity

A threat to the assumption that the independent (exogenous) variable actually causes the dependent (or endogenous) variable. Endogeneity occurs when the dependent variable may actually be a cause of the independent variable. Sometimes this is referred to as reverse causality. For example, a researcher may note that states with the death penalty also have high murder rates. The researcher may conclude that the death penalty causes an increase in the murder rate; however, it could be that states that experience a high murder rate are more likely to institute the death penalty. Endogeneity is the opposite of exogeneity.

Epistemology

A way of understanding and explaining how we know what we know. Each research methodology is underpinned by an epistemology that serves as a guiding philosophy and provides a concrete process of research steps.

Error

The difference between the actual observed data value and the predicted or estimated data value. Predicted or estimated data values are calculated in statistical analyses, such as regression analysis.

Error Term

The part of a statistical equation that indicates what remains unexplained by the independent variables. The residuals in regression models.

Estimated Sampling Error

When researchers use a sample to predict the true but unknown value of a population there is the risk of being wrong. They decide ahead of time how much error in their prediction is acceptable. For example, when predicting the percentage of children enrolled in home-based child care, a researcher could select a sample size with an error of +/- 5 percent points.

Estimation

The process by which data from a sample are used to indicate the value of an unknown quantity in a population.

Ethnographic Decision Models

A qualitative method for examining behavior under specific circumstances. An EDM is often referred to as a decision tree or flow chart and comprises a series of nested "if-then" statements that link criteria (and combinations of criteria) to the behavior of interest.

Ethnographic Interviewing

A research method in which face-to-face interviews with respondents are conducted using open-ended questions to explore topics in great depth. Questions are often customized for each interview, and topics are generally probed extensively with follow-up questions.

Ethnography

Literally meaning "folk-" or "people-" "writing," ethnography is a field method focused on recording the details of social life occurring in a society. A primary objective is to gain a rich, *thick* understanding of a setting and of the members within a society. Ethnographers seek to learn the language, thoughts, and practices of a society by participating in the rituals and

observing the everyday routines of the community. Ethnography is primarily based upon participant observation, direct observation, and in-depth interviewing.

Evaluation Research

The use of scientific research methods to plan intervention programs, to monitor the implementation of new programs and the operation of existing programs, and to determine how effectively programs or clinical practices achieve their goals.

Event History Analysis

Event history analysis deals with data obtained by observing individuals over time, focusing on events occurring for the individuals under observation. It is similar to linear or logistic regression analysis, except that the dependent variable is a measure of the likelihood or timing of an event (e.g., birth of first child, divorce, job changes).

Exogeneity

The condition of being external to the process under study. For example, a researcher may study the effect of parental characteristics on their children's behaviors. A parent's religious upbringing is exogenous to their children's behaviors because it is impossible for children's current behavior to impact parent's upbringing, which occurred prior to the birth of the child. The opposite of exogeneity is endogeneity.

Experimental Control

Processes used to hold the conditions uniform or constant under which an investigation is carried out.

Experimental Design

A research design used to establish cause-and-effect relationships between the independent and dependent variables by means of manipulation of variables, control and randomization. A true experiment involves the random allocation of participants to experimental (treatment or intervention) and control (non-treatment or non-intervention) groups, manipulation of the independent variable (e.g., new reading or math curriculum), and the introduction of a control group for comparison purposes (e.g., classes in program that will not use new reading or math curriculum). Participants are assessed after (or before and after) the manipulation of the independent variable in order to assess its effect on the dependent variable (e.g., the reading or math skills of children in both the experimental and control groups are assessed); differences between the groups are assumed to be due to the treatment effect.

Experimental Group

In experimental research, the group of subjects who receive the experimental treatment or intervention under investigation.

Explanatory Analysis

A method of inquiry that focuses on the formulating and testing of hypotheses. For example, instead of, or in addition to, describing Black and White differences in the reading and math skills of preschool children, the analysis focuses on testing whether factors that may contribute to these differences (e.g., resources available to children at home and in their child care programs) are in fact associated with those differences.

Exploratory Factor Analysis

A multivariate statistical method used to uncover the underlying structure (factors) of a set of variables. It is a form of factor analysis whose goal is to identify the underlying relationships between measured variables. It is especially useful when there is no hypothesis about the underlying structure of the data and the pattern of relationships between the variables. Researchers often use exploratory factor analysis to develop scales from multiple survey or assessment items. It is a way of reducing the number of items to be used in an analysis to a smaller and more meaningful set.

Exploratory Study

An exploratory study is conducted to begin to understand a research problem or question when there are few or no earlier studies and/or limited theory to refer to or rely upon to answer the question. The focus is on gaining insights and familiarity for later investigation. Many variables are often taken into account and compared, using a variety of techniques in the search for patterns.

External Validity

The degree to which the results of a study can be generalized beyond the study sample to a larger population.

Extraneous Variable

A variable that interferes with the relationship between the independent and dependent variables and which therefore needs to be controlled for in some way.

Extrapolation

Predicting the value of unknown data points by projecting beyond the range of known data points.

Face Validity

The extent to which a survey or a test appears to actually measure what the researcher claims it measures. For example, a researcher may create survey questions that s/he claims measure gender role attitudes. To have face validity, other researchers who read the survey questions must also agree that the questions do appear to measure gender role attitudes.

Factor Analysis

A form of multivariate analysis that includes a large number of variables or objects and aims to identify a smaller number of factors that are more understandable. It is a way of identifying patterns in the data and overlap in the patterns. There are two basic types: exploratory factor analysis and confirmatory factor analysis. Definitions of the two types can be found elsewhere in the glossary.

Field Experiments

An experimental study that is not conducted in a laboratory, but instead in real-life settings such as early childhood classrooms and schools. Field experiments, like lab experiments, generally randomize subjects (or other units such as classrooms or schools) into treatment and control groups and compare outcomes between these groups. For example, to evaluate the effectiveness of a new math curriculum, a sample of 4-year-old classrooms may be randomly

assigned to one of two groups: (1) classrooms that will use the new curriculum (treatment) and (2) classrooms that will continue to use the old curriculum.

Field Notes

A text document that detail behaviors, conversations, or setting characteristics as recorded by a qualitative researcher. Field notes are the principle form of data gathered from direct observation and participant observation.

Field Research

Research conducted where research subjects live or where the activities of interest take place.

Field Work

Observing human behavior or interviewing individuals within their own communities. Field work is generally used to collect qualitative data. It often involves long-term relocation of researchers to the community under study. Data collection generally takes place over an extended period of time. The term is also used more broadly to describe the tasks performed by members of a research team in schools, early childhood programs, and communities. This might include: working with school and program staff to select samples of classes and children; conducting in-person interviews with teachers and other program staff and children's parents; and administering standardized assessments to the children.

Fixed Effects Regression

Regression techniques that can be used to eliminate biases associated with the omission of unmeasured characteristics. Biases are eliminated by including an individual-specific intercept term for all cases.

Floor

The lowest limit of performance that can be assessed or measured by an instrument or process. Individuals who perform near to or below this lower limit are said to have reached the floor, and the assessment may not be providing a valid estimate of their performance levels.

Focus Group

An interview conducted with a small group of people, all at one time, to explore ideas on a particular topic. The goal of a focus group is to uncover additional information through participants' exchange of ideas.

Forecasting

The prediction of the size of a future quantity (e.g., unemployment rate next year).

Frequency Distribution

The frequency with which values of a variable occur in a sample or a population. To graph a distribution, first the values of the variables are listed across the bottom of the graph. The number of times the value occurs are listed up the side of the graph. A bar is drawn that corresponds to how many times each value occurred in the data. For example, a graph of the distribution of women's heights from a random sample of the population would be shaped like a bell. Most women's height are around 5'4" This value would occur most frequently, so it would have the highest bar. Heights that are close to 5'4", such as 5'3" and 5'5" would have slightly shorter bars. More extreme heights, such as 4'7" and 6'1" would have very short bars.

GIS (Geographical Information Systems)

A computer system that enables one to assemble, store, manipulate, and display geographically referenced information.

Game Theory

Game theory is the study of how and why people make decisions using mathematical models of conflict and cooperation. It analyzes the strategies that two or more individuals use in dealing with situations where the outcome of an individual's choice of action depends on the actions of other individuals. The prisoner's dilemma is an example of how game theory can be used to show why two completely rational individuals might not cooperate, even if it appears that it is in their best interests to do so.

Generalizability

The extent to which conclusions from analysis of data from a sample can be applied to the population as a whole.

Generalized Linear Mixed Model

A generalized linear mixed model (GLMM) is an extension to the generalized linear models (GLMs) in which the predictor contains random effects in addition to fixed effects.

Generalized Linear Models

The generalized linear models (GLMs) are a broad class of statistical models that include linear regression, analysis of variance, logistic regression, log-linear models, etc. These models can be used when analyzing the effects of independent variables with different distributions. They do not require that the dependent variable be normally distributed.

Gini Coefficient

A measure of inequality or dispersion in a group of values (e.g.; racial inequality in a population). The larger the coefficient the greater the dispersion.

Grounded Theory

Grounded theory (GT) is an inductive research methodology used in the social sciences. It involves the construction of theory from the data collected in research and analyses of those data. Thus, it is quite different from the traditional deductive approach, where the researcher collects and analyzes data to test an existing theory and a set of research hypotheses derived from that theory. Grounded theory is used widely in qualitative research.

Hawthorne Effect

Refers to the process where research subjects change their behavior in response to being studied. That is, people behave differently because they are being observed. For example, teachers in a classroom might change their discipline practices if they are part of a study that uses classroom observations as a data collection approach.

Heterogeneity

The degree of dissimilarity among cases with respect to a particular characteristic.

Heterogeneous Treatment Effects

Randomized experimental designs test the average effect of a treatment or intervention.

However, the treatment might affect different research subjects or groups of subjects in different ways. The effects might be larger for some subjects and smaller for others, and it may have no effect on some subjects in the treatment group. The study of treatment effect heterogeneity is the study of these differences across subjects and groups of subjects. The findings from these studies provide important information that can be used to develop theories about the conditions under which the treatment is effective or ineffective.

Heteroskedastic

A distribution characterized by a changing (non-constant) variance or standard deviation. Heteroskedasticity is problematic in statistical models because estimated standard errors will be inefficient and biased. Consequently, traditional significance test will not be valid.

Hierarchical Linear Modeling (HLM)

A multi-level modeling procedure that is used to analyze the variance in the outcome variables when the predictor variables are at varying hierarchical levels (child, classroom, programs) and nested such that children are nested in classrooms and classrooms are nested in programs. For example, HLM can be used to estimate the effects of child characteristics (e.g., race/ethnicity, gender, social economic status) on children's academic skills for designs that include children nested in classrooms. HLM enables a researcher to estimate effects within individual units (children), formulate hypotheses about cross level effects (children and classrooms) and partition the variance and covariance components among levels (share of variance explained by child and by classroom characteristics).

Hierarchical Model

A hierarchical model is a type of linear regression model that is used when the data one is analyzing are organized into a tree-like structure (or hierarchy). That is, data at one level is nested under another level. For example, children are nested in classrooms. Hierarchical models are used when predicting outcomes (e.g., children's test scores) using variables from the different levels (child level variables such as gender and prior test scores and classroom level variables such as class size and quality of teacher-child interactions).

Histogram

A visual presentation of data that shows the frequencies with which each value of a variable occurs. Each value of a variable typically is displayed along the bottom of a histogram, and a bar is drawn for each value. The height of the bar corresponds to the frequency with which that value occurs.

Hypothesis

A statement that predicts the relationship between the independent (causal) and dependent (outcome) variables.

Hypothesis Testing

Statistical tests to determine whether a hypothesis is accepted or rejected. In hypothesis testing, two hypotheses are used: the null hypothesis and the alternative hypothesis. The alternative hypothesis is the hypothesis of interest; it generally states that there is a relationship between two variables. The null hypothesis states the opposite, that there is no relationship between two variables.

Imputed Response

A missing survey response that is filled in by the data analyst. The method to fill in the missing response is based on careful analysis of patterns of missing data. Imputation is done to allow for statistical analysis of surveys that were only partially completed.

In-depth Interviewing

A research method in which face-to-face interviews with respondents are conducted using open-ended questions to explore topics in great depth. Questions are often customized for each interview, and topics are generally probed extensively with follow-up questions.

Independence

The lack of a relationship between two or more variables. For example, annual snow fall and the Yankee's season record are independent, but annual snow fall and coat sales are not independent.

Independent Variable

The variable that the researcher expects to be associated with an outcome of interest. For example, if a researcher wants to examine the relationship between parental education and children's language development, parent education (years of schooling or highest level of education completed) is the independent variable. Sometimes this variable is referred to as the treatment variable or the causal variable.

Independent and Identically Distributed (IID)

A collection of two or more random variables $\{X_1, X_2, \dots\}$ is independent and identically distributed if the variables are independent and also have the same probability distribution.

Index

A type of composite measure (i.e., a measure that is created from more than one data item, such answers to a series of survey questions) that summarizes responses to several specific observations (e.g., items on a parent questionnaire that ask whether their child participates in a variety of extracurricular activities) and represents a more general dimension (e.g., extracurricular activities might include music and dance lessons, different sports, and youth clubs). An index is often created by simply summing the responses to a series of yes/no questions.

Index Variable

A variable that is a summed composite of other variables that are assumed to reflect the same underlying construct. For example, a count of the number of caregiving activities (e.g., bathing and feeding) a father engages in with his infant child.

Indicator

An observation or measure that is assumed to be evidence of the attributes or properties of some phenomenon. Indicators are monitored over time and are used to assess progress toward the achievement of intended outcomes, goals, and objectives. Child well-being indicators include children's letter knowledge, frequency of pro- and anti-social behaviors, being read to on a regular basis by family members and attending high quality early childhood program.

Indicator Variable

In statistics, an indicator variables has only two possible values, which are typically coded 0 and 1 to identify the presence (1) or absence (0) of a characteristic. Some examples would include high school graduate (yes/no), preschool enrollment (enrolled, not enrolled), and program type (publicly funded/private). Also referred to as a dummy variable and are often used in regression analysis.

Indirect Effect

A condition where one variable affects another indirectly through an intervening variable. For example, parental income may have an indirect effect on children's school readiness skills if income affects the quality of child care a child receives which in turn affects the child's early reading, math and social skills. Also referred to as mediation.

Inductive Approach to Data Analysis

The general inductive approach to data analysis allows research findings to emerge from frequent and dominant themes observed in the data. This approach is generally associated with qualitative research, where researchers develop models and theory from a careful review, analysis and summary of the raw data.

Inductive Method

A method of study that begins with specific observations and measures, from which patterns and regularities are detected. These patterns lead to the formulation of tentative hypotheses, and ultimately to the construction of general conclusions or theories

Informal Interview

An informal or conversational interview is a type of qualitative interview where the researcher begins by engaging an individual in a conversation. As the conversation unfolds, the researcher formulates specific questions, often spontaneously, and begins asking them informally. It is used when the researcher wants maximum flexibility to pursue topics and ideas as they emerge during the interview.

Informed Consent

The agreement between concerned parties about the data-gathering process and/or the disclosure, reporting, and/or use of data, information, and/or results from a research experiment.

Instrument Error

A type of non-sampling error caused by the survey instrument (or questionnaire) itself, such as unclear wording, asking respondents for information they are unable to supply or the instrument being changed in some way during the course of the research.

Instrumental Variables

Instrumental variables are used to estimate causal relationships or effects when controlled experiments are not feasible or when a treatment is not successfully delivered to every unit in a randomized experiment's treatment group. An instrumental variable is related to the explanatory variable of interest (e.g., participation in a child care program or enrollment in a before- and after-school program) but has no independent effect on the dependent variable

(e.g., children's reading and math scores or social skills). For example, region of the country may be associated with participation in child care or before- and after-school programs, but would not be expected to relate to measures of children's development.

Intent-to-Treat

Intent-to-treat (ITT) is an approach for estimating the effects of a treatment or intervention that includes all those who were randomly assigned to the treatment or control groups, regardless of the treatment they actually received. It includes those who were assigned to the treatment but who did not participate in the intervention and those who withdrew before the study was complete. Estimate of the treatment effect is generally conservative, because the outcomes of those who did not receive the treatment or the full treatment are generally smaller. It is best used when the researcher has access to outcome data on all those who were randomly assigned at the beginning of the study.

Inter-Rater Reliability

A measure of the consistency between the ratings or values assigned to a behavior that is being rated or observed; usually expressed as a percentage of agreement between two raters/observers, or as a coefficient of agreement which can be stated as a probability.

Interaction Effect

An interaction effect is observed when the effect of one explanatory (independent) variable differs depending on the level of another explanatory variable. This is distinct from a "main effect," which is the effect of a single explanatory variable on the dependent variable. In the research literature, interaction effects are sometimes called simultaneous effects or moderators.

Intercept

In regression analysis, the predicted value of Y (dependent variable) when all X variables (independent variables) are zero (0). It is also called the Y-intercept or the constant term. It is the place where the regression line crosses the Y-axis.

Internal Validity

The extent to which researchers provide compelling evidence that the causal (independent) variable causes changes in the outcome (dependent) variable. To do this, researchers must rule other potential explanations for the changes in the outcome variable.

Interval Scale

A scale of measurement where the distance between any two adjacent units of measurement is the same but the zero point is arbitrary. Scores on an interval scale can be added and subtracted but cannot be meaningfully multiplied or divided.

Interval Variable

A variable wherein the distance between units is the same but the zero point is arbitrary.

Intervening Variable

An intervening variable or mediating variable is something that effects or helps to explain the relationship between an independent and a dependent variable. An intervening variable is the link between an independent and a dependent variable. It is predicted by the independent

variable, and it predicts the dependent variable. In early childhood research, intervening variables are often used to identify the processes that underlie the relationships between independent variables and dependent variables. For example, the time parents spent reading to children may explain the relationship between parental education and children's reading scores on a standardized test.

Intervention

The situation or variable introduced to affect the dependent variable (outcome); manipulations of the subject or the subject's environment performed for research purposes. In early education, interventions often involve the introduction and use of one or more instructional approaches designed to improve children's learning. Other interventions might include professional development training designed to improve the quality of teaching in the classroom and ultimately children's learning.

Interviewer Error

A type of non-sampling error caused by mistakes made by the interviewer. These may include influencing the respondent in some way, asking questions in the wrong order, or using slightly different phrasing (or tone of voice) than other interviewers. It can include intentional errors such as cheating and fraudulent data entry.

Jackknife Technique

A (usually) computer-intensive resampling method used to estimate population parameters (for example, means and percentage), and/or to gauge uncertainty in these estimates (e.g., standard error). The name is derived from the approach that involves removing each observation (i.e., cut with a knife) one at a time (or two at a time for the second-order Jackknife, and so on), calculating the mean for each new sample (original sample minus the omitted case) and then averaging the means of the new samples.

Kurtosis

A statistic that measures how outlier-prone a distribution is. The kurtosis of a normal distribution is 0. If the kurtosis is different from 0, then the distribution produces outliers that are either more extreme (positive kurtosis) or less extreme (negative kurtosis) than are produced by the normal distribution.

Latent Growth Model

Latent growth modeling (LGM) is a class of statistical methods that are used to study change (growth) in behavior or attitudes over time. Traditional approaches to the study of change such as regression analysis and ANOVA focus on mean change (average change of a group or subgroups of study participants) and treat differences in change between participants as error. LGM examines individual (within-person) change over time as well as differences in the individual change (between-person). It is used to model change over time and to investigate factors that affect the level and rates of change. It can be used to examine differences in which groups of people change.

Latent Variables

In statistics, latent variables are variables not directly observed and measured but inferred from other observed and measured variables. Mathematical models (e.g., factor analysis, structural

equation modeling, item response theory models) are used to examine the relationships between a set of observed variables (indicators) in order to identify the latent variable. For example, the latent variable 'teacher attitudes toward math' may be modeled from a series of survey items asking about their feelings toward math and how they feel when doing math.

Least Squares

A commonly used method for calculating a regression equation. This method minimizes the difference between the observed data points and the data points that are estimated by the regression equation.

Level of Significance

See significance level.

Likert Scale

A Likert Scale is a type of rating scale used to measure attitudes, values, or opinions about a subject. Survey respondents are asked to indicate their level of agreement or disagreement with a series of statements. The responses are often scaled and summed to give a composite measure of attitudes or opinions about a topic.

Limited Dependent Variable

A limited dependent variable is a variable with the range of possible values "restricted in some important way." Examples include binary variables that have only two values (e.g., child attends child care or not; child is promoted to next grade or not). Also, variables that can only take on certain values (e.g., discrete variables that have a limited set of categories or continuous variables that can only have positive values such as hours worked or wages earned).

Linear Regression

A statistical technique used to find a linear relationship between one or more (multiple) continuous or categorical predictor (or independent) variables and a continuous outcome (or dependent) variable.

Literature Review

A comprehensive survey of the research literature on a topic. Generally the literature review is presented at the beginning of a research paper and explains how the researcher arrived at his or her research questions.

Local Average Treatment Effect

The local average treatment effect (LATE) is the average effect of a treatment for the group of individuals who complied with the random assignment. That is, it is the estimated effect when those assigned to the treatment group received the treatment and those assigned to the control group did not.

Log-Linear Analysis

A technique used in statistics to examine the relationship between two or more categorical variables. The technique is used to test whether the variables are independent of one another or associated in some way. For example, a log-linear model might be used to test whether there is an association between children being 'red shirted' (children whose parents hold them back from entering kindergarten even though they meet the age eligibility requirement for

kindergarten entry) and their sex (boys versus girls) and race/ethnicity (African American, Hispanic, White, Asian). The model would test for both main effects and interaction effects.

Logistic Regression

Logistic regression (or logit regression) is a special form of regression used to analyze the relationship between predictor variables and a dichotomous outcome variable. A dichotomous variable is a variable with only two possible values, e.g. child attends center-based child care (Yes/No). For example, a researcher might use logistic regression to study the relationships between parental education, household income, and parental employment and whether children receives child care from someone other than their parents (receives nonparent care/does not receive nonparent care).

Logit Model

A special form of regression used to analyze the relationship between predictor variables and a categorical outcome variable.

MANOVA (Multivariate Analysis of Variance)

MANOVA is an extension of analysis of variance (ANOVA). ANOVA is used to examine whether there are statistical differences in the group means for a single continuous dependent variable. MANOVA is used when there are two or more dependent variables (outcome variables). It is a statistical test that measures group differences on several dependent variables.

Main Effect

The effect of a predictor (or independent) variable on an outcome (or dependent) variable.

Markov Chain Monte Carlo Methods

Markov chain Monte Carlo (MCMC) methods are used to estimate the properties of a distribution by examining random samples from the distribution. Researchers using a Monte Carlo approach draw a large number of random samples from a normal distribution, and calculate the sample mean of those. The random samples are generated by a special sequential process. Each random sample is used as a stepping stone to generate the next random sample (hence the chain). A special property of the chain is that, while each new sample depends on the one before it, new samples do not depend on any samples before the previous one (this is the "Markov" property).

Matched Samples

Two samples in which the members are paired or matched explicitly by the researcher on specific attributes, such as IQ or income. Also refers to samples in which the same attribute or variable is measured twice on each subject under different circumstances; also referred to as repeated measures.

Maxima

The maxima are points where the value of a function is greater than other surrounding points.

Maximum Likelihood Estimation

Maximum-likelihood estimation (MLE) is one of the most widely used methods for estimating the parameters of a statistical model (for example, means and variances) from sample data.

Using the sample data, MLE obtains estimates of the population parameters such that the probability (likelihood) of obtaining the observed data is maximized.

Mean

A descriptive statistic used as a measure of central tendency. To calculate the mean, all the values of a variable are added and then the sum is divided by the number of values. For example, if the age of the respondents in a sample were 21, 35, 40, 46, and 76, the mean age of the sample would be $(21+35+40+46+76)/5 = 43.6$

Measurement Error

The difference between the value measured in a survey or on a test and the "true" value, if the difference is due to factors beyond the control of the respondent. Some factors that contribute to measurement error include the environment in which a survey or test is administered (e.g., administering a math test in a noisy classroom could lead students to do poorly even though they understand the material), poor measurement tools (e.g., using a ruler that is only marked in feet to measure height would lead to inaccurate measurement), rater effects (e.g., if a police man in uniform conducted interviews with individuals about drug use, they might not feel comfortable revealing their drug use). There are many more such factors that can contribute to measurement error.

Measures of Association

Statistics that measure the strength and nature of the relationship between variables. For example, correlation is a measure of association

Median

A descriptive statistic used to measure central tendency. The median is the value that is the middle value of a set of values. 50% of the values lie above the median, and 50% lie below the median. For example, if a sample of individuals are ages 21, 34, 46, 55, and 76 the median age is 46.

Member Checking

During open-ended interviews, the practice of a researcher restating, summarizing, or paraphrasing the information received from a respondent to ensure that what was heard or written down is in fact correct.

Meta-Analysis

A statistical technique that combines and analyzes data across multiple studies on a topic. In early childhood and education research, a meta-analysis combines a number of studies (usually conducted by a number of different researchers in a variety of contexts) to quantify the effect a given independent or treatment variable (e.g., full-day versus part-day kindergarten and class size) has on a given outcome (e.g., children's academic skills and prevalence of positive and negative classroom behavior).

Methodology

The principles, procedures, and strategies of research used in a study for gathering information, analyzing data, and drawing conclusions. There are broad categories of methodology such as

qualitative methods or quantitative methods; and there are particular types of methodologies such as survey research, case study, and participant observation, among many others.

Metropolitan Statistical Area (MSA)

A term used by the U.S. Census Bureau to designate an area of adjacent counties (except in New England where they are defined by adjacent cities). Metropolitan Statistical Areas (MSAs) are often used to geographically understand labor markets because individuals often look for work outside of the city or county in which they live.

Micro-Ethnography

Micro-ethnography, also known as ethnographic microanalysis of interaction, describes how interaction is socially and culturally organized in particular situational settings (e.g., classrooms or neighborhoods). It is a process that includes data collection, content analysis, and comparative analysis of everyday situations for the purpose of formulating insights.

Minima

The minima are points where the value of a function is less than other surrounding points.

Missing Completely at Random (MCAR)

The term implies that all respondents are equally likely/unlikely to respond to the item and that the estimate is approximately unbiased. To ignore the missing data and restrict analyses to those records with reported values for the variables in the analysis, implicitly invokes the assumption that the missing cases are a random subsample of the full sample, that is, they are missing completely at random (MCAR). This is a strong assumption.

Missing Data

Values in a data set values that were not recorded. Missing values can have many causes including a respondent's refusal to answer survey questions, an interviewer incorrectly coding a response, or questions that do not apply to a respondent. The more missing data there are in a data set, the greater the likelihood of bias. There are several coding strategies that can "fill in" missing data for statistical analyses. These strategies are called imputation (see Data Imputation).

Missing Data Imputation

A method used to fill in missing values (due to nonresponse) in surveys. The method is based on careful analysis of patterns of missing data. Types of data imputation include mean imputation, multiple imputation, hot deck and cold deck imputation. Data imputation is done to allow for statistical analysis of surveys that were only partially completed.

Misspecification

Misspecification occurs when the predictor (independent) variables in a statistical model are incorrect. The most common cause of model misspecification is that important predictor (independent) variables are left out of the model. Misspecification often leads to incorrect estimates of the effects of the predictor (independent) variables that are included in the model on the outcome (dependent) variable.

Mixed-Method Research

A approach to conducting research that involves collecting, analyzing and integrating data from

quantitative (e.g., experiments, surveys) and qualitative (e.g., focus groups, observations, qualitative interviews) methods. The goal is to provide greater breath and depth of understanding of a topic.

Mode

A descriptive statistic that is a measure of central tendency. It is the value that occurs most frequently in the data. For example, if survey respondents are ages 21, 33, 33, 45, and 76, the modal age is 33.

Moderating Variable

A moderating variable, also called a moderator variable, is a variable that affects the strength or direction of the relationship between a dependent and independent variable. It is represented by an interaction term in regression models. For example, a researcher might examine whether maternal depression moderates the relationship between the quality of mother-child interactions and children's development. The researcher might hypothesize that this relationship will be weaker for children whose mothers are depressed than for other children.

Moving Average

A form of average which has been adjusted (or "smoothed") to allow for seasonal or cyclical components of a time series.

Multi-Dimensional Scaling

Multidimensional scaling (MDS) is a tool used by researchers to identify and quantify the relationships among the responses of study participants to a group of items in a survey or other data collection instrument. The outcome of MDS is a visual representation of the patterns of similarities, dissimilarities and distances among these participants.

Multi-Site Trials

Experiments that are conducted at multiple locations, often by several research organizations. Multi-site studies provide larger samples, which increases the statistical power to detect significant treatment effects. Given differences in the populations in the different communities, samples are often more diverse, which can increase the likelihood that the effects are not solely limited to a single population.

Multi-Stage Sampling

Multistage sampling is a probability sampling technique where sampling is carried out in several stages. It is often used to select samples when a single frame is not available to select members for a study sample. For example, there is no single list of all children enrolled in public school kindergartens across the U.S. Therefore, researchers who need a sample of kindergarten children will first select a sample of schools with kindergarten programs from a school frame (e.g., National Center for Education Statistics' Common Core of Data) (Stage 1). Lists of all kindergarten classrooms in selected schools are developed and a sample of classrooms selected in each of the sampled schools (Stage 2). Finally, lists of children in the sampled classrooms are compiled and a sample of children is selected from each of the classroom lists (Stage 3).

Multicollinearity

A situation in which two or more predictor (independent) variables in a sample are highly

related to each other. When using regression analysis, this can lead to incorrect estimates of their individual effects on the outcome (dependent) variable. Multicollinearity violates an underlying assumption of regression that each predictor (independent) variable has an independent impact on the outcome (dependent) variable.

Multilevel Data

Multilevel data are organized at more than one level. That is, the data are nested. In research that involves children enrolled in early childhood programs such as Head Start, the children in the study are often nested in classrooms, which are nested in Head Start centers.

Multilevel Modeling

A model involving variables measured at more than one level of a hierarchy. An obvious hierarchy consists of children nested in classes, and classes nested in schools. Measurements can be obtained for child characteristics, class and teacher characteristics, or school characteristics. Multilevel models are also known as hierarchical linear models or random coefficient models. Multilevel are used to solve the statistical problems caused by dealing with hierarchically nested data.

Multinomial Distribution

A distribution that arises when a response variable is categorical in nature. For example, if a researcher recorded the type of child care a child used, then the distribution of the counts in these categories would be multinomial. The multinomial distribution is a generalization of the binomial distribution to more than two categories. If the categories for the response variable can be ordered, then the distribution of that variable is referred to as ordinal multinomial.

Multinomial Logit Model

A special form of regression used to analyze the relationship between predictor variables and a categorical outcome variable. The multinomial logit is used when the categorical outcome variable has more than two values, e.g., marital status could be never married, married, or divorced.

Multiple (Linear) Regression

A statistical technique used to find the linear relationship between an outcome (dependent) variable and several predictor (independent) variables.

Multivariate Analysis

Any of several statistical methods for examining more than one predictor (independent) variable or more than one outcome (dependent) variable or both. Allows researchers to examine the relation between two variables while simultaneously controlling for the influence of other variables. For example, a researcher might examine the relationships between children's race/ethnicity and their early reading skills while controlling for the influence of family social economic status (SES).

Multivariate Normal Distribution

The multivariate normal distribution is a generalization of the one-dimensional (univariate) normal distribution. The multivariate normal distribution has two or more random variables

and every linear combination of the values of these variables has a univariate normal distribution.

Multivariate Probit Model

A model that is used to estimate several correlated bivariate outcomes jointly. For example, if parental decisions regarding sending their children to public schools and voting in favor of a school budget are correlated (both decisions are binary), then the multivariate probit model would be appropriate for jointly predicting these two choices on an individual-specific basis.

Mutually Exclusive

Said of variables, events or conditions that can be placed into one category and no other. If there is no overlapping part between two events, we say they are mutually exclusive. However, mutually exclusive doesn't mean the two events are independent.

Narrative Research

Narrative research is a group of qualitative approaches that rely on the written or spoken words or visual representation of individuals to understand how people create meaning in their lives. The focus is most often on the lives of individuals as told through their own stories. The emphasis is on the story, what it says and how it is told.

Natural Experiment

A natural experiment is an empirical study in which individuals or groups of individuals are sorted by "nature" (i.e., by conditions or events that were not controlled by the researcher) into something like treatment and control groups. They differ from Randomized Controlled Trials (RCTs). The key assumption of a natural experiment is that there are comparable groups, and one group is randomly affected by forces outside their control ("nature"). For example, researchers often study the effects of nature disasters on a range of outcomes.

Negative Binomial Regression

Negative binomial regression is a type of generalized linear model that researchers use when the dependent variable is a count of the number of times an event occurs. For example, a researcher might use negative binomial regression to study the relationships between several variables (e.g., children's general health status and type of early care and education program) and the number of school days children have missed during the school year.

Nominal Data

See categorical data.

Nominal Scale

A scale that allows for the classification of elements into mutually exclusive categories based on defined features but without numeric value.

Nonlinear Models

A nonlinear model describes nonlinear relationships between the dependent and independent variable(s). A linear model assumes that the dependent variable changes by a fixed amount for each unit change in the independent variable. A nonlinear model, on the other hand, does not make this assumption. Instead of the relationship between the dependent and independent

variable being represented by a straight line, a nonlinear relationship is characterized by one or more curves.

Nonlinear Trends

When analyzing time series data, a linear trend is one where the data increase by a constant amount at each successive time period. A linear trend is represented by a straight line. However, data do not always increase by the same amount. For example, data that increase by varying amounts at each successive time period show a nonlinear, curvilinear trend.

Nonparametric Statistics

Nonparametric statistics refer to the group of statistical methods that require fewer assumptions about the distribution of the data. For example, nonparametric tests of significance such as the Chi-square test does not require the data to fit a normal distribution. Nonparametric statistical methods are used when analyzing nominal, ordinal or ranked data.

Nonresponse Error

A type of error that is caused when a portion of the sample with particular characteristics do not respond to a survey. For example, individuals who are trying to dodge bill collectors might be less likely to answer their telephone and therefore may be less likely to respond to a telephone survey. This could lead to biased statistical results because individuals who do not pay their bills would be less likely to answer the survey. Researchers try to correct for this problem by determining the characteristics of those who were less likely to answer the survey (for example, they might be younger and have lower incomes) and controlling for those characteristics in the analysis, by imputing missing data, or by using sampling weights (analysis weights) that adjust for nonresponse.

Nonresponse Rate Bias

A source of bias that occurs when non-respondents differ in important ways from respondents.

Nonsampling Error

Errors that can affect the accuracy of research findings other than errors associated with selecting the sample (sampling error). They can occur in any phase of a research study (planning and design, data collection, or data processing). They include errors that occur when units in the target population are missing from the sampling frame, coverage error, nonresponse to surveys, measurement errors due to interviewer or respondent behavior, errors introduced by how survey questions were worded or by how data were collected (e.g., in-person interview, web survey), and processing error (e.g., errors made during data entry or when coding open-ended survey responses).

Nonsignificant Result

The result of a statistical test that indicates that there is not sufficient evidence to conclude that the predictor (independent) variable had an impact on the outcome (dependent) variable.

Normal Curve

The bell-shaped curve that is formed when data with a normal distribution are plotted.

Normal Distribution

A distribution of the frequency of data points that resembles a bell shape. In a normal

distribution, one half of the data points (e.g., scores on a standardized test) are above and one half are below the mean or median of the full set of data. The further a data point is from the mean/median, the less likely it is to occur. And, the closer it is, the more likely it is to occur. The normal distribution exhibits important mathematical properties that are necessary for performing most statistical tests (e.g., data have a normal distribution or are at least symmetrical with half of the values above and half below the mean or median).

Null Hypothesis

This hypothesis states that there is no difference between groups. The alternative hypothesis states that there is some real difference between two or more groups.

Observation Unit

The actual unit observed during a study in order to measure something about it. In child care and early education research typical observation units include programs and schools, classrooms and teachers, children and their parents.

Odds Ratio

A way to express a probability; the ratio of the odds of having a response or experience to the odds of not having it.

Omitted Variable Bias

A form of bias in research resulting from the absence of key variables into the research design that would influence the results. When there is omitted variable bias, the results of the study could be due to alternative explanations that are not addressed in the study.

Omitted Variable Sensitivity Analysis

Omitted variable sensitivity analysis is used to assess the influence of not including one or more variables on the relationship between an independent variable X and a dependent or outcome variable Y. It quantifies how large the effect of the omitted variable or variables would have to be in order to invalidate or explain away the association between X and Y. For example, the correlations between the omitted variable and the independent and dependent variables would have to be of a specified value in order to invalidate the finding that X and Y are significantly associated.

One-Way ANOVA

A test of whether the mean for more than two groups are different. For example, to test whether the mean income is different for individuals who live in France, England, or Sweden, one would use a one-way ANOVA.

Open-Ended Data

Data derived from open-ended inquiries, such as interview questions, to which responses are not predetermined, such as would be the case with multiple choice or true/false questions.

Optimal Matching

Matching is a technique that is used to evaluate the effect of a treatment (intervention) when those who receive the treatment and those who do not have not been randomly assigned. The goal of matching is, for every treated subject (participant), to find one (or more) subjects who have not received the treatment but who have similar characteristics (e.g., age, gender,

achievement test scores) to those who have. Matching enables a comparison of outcomes among those treated and those not treated in order to estimate the effect of the treatment. Optimal matching is a global matching approach that looks to minimize the distance (differences) between matched subjects.

Ordinal Data

Data that are categorical, but that can also be ranked (ordered). However, the distance between the categories is not known and may not be equal. For example, parents might rate their satisfaction with their child's child care provider as "very dissatisfied," "dissatisfied," "satisfied," and "very satisfied." using numerical values of 1, 2, 3 and 4, respectively. A parent with a satisfaction score of 1 is more dissatisfied than a parent with a score of 2, but not necessarily twice as dissatisfied. And the difference between scores of 1 and 2 and scores of 3 and 4 are not necessary the same.

Ordinal Scale

A scale that allows for classification and labeling into mutually exclusive categories based on features that are ranked or ordered with respect to one another, although equal differences between numbers do not reflect an equal magnitude of difference.

Ordinary Least Squares Estimation

A commonly used method for calculating a regression equation. This method minimizes the difference between the observed data points and the data points that are estimated by the regression equation.

Outcomes

Outcomes are the measured behaviors, attitudes, or other characteristics of a sample or population that research seeks to explain. There may be one or more than one outcome of interest in a single research study. Outcomes may be measured at different levels (e.g., communities, schools/early childhood programs, classrooms, families and children).

Outlier

An observation in a data set that is much different than the other observations in the data set. The data point is unusually larger or unusually smaller compared to the other data points.

Oversampling

A sampling procedure in which a large proportion of subjects with a particular characteristic are sampled. Oversampling is used to ensure that researchers have enough data from groups with particular characteristics to yield good estimates for that group. For example, researchers often over sample African-Americans because just 12% of the population is African-American. This ensures that enough African-Americans are in the sample to yield good models and estimates for African-Americans.

P-Value

The probability that the results of a statistical test were due to chance. A p-value greater than .05 is usually interpreted to mean that the results were not statistically significant. Sometimes researchers use a p-value of .01 or a p-value of .10 to indicate whether a result is statistically significant. The lower the p-value the more rigorous the criteria for concluding significance.

Paired Comparison Method

The paired comparison method is a research design that yields scores based on respondents' ratings of pairs of items. For example, a respondent is presented with a set of binary items and asked to indicate which of the choices he/she prefers or is more applicable.

Paired T-Test

This test, which is sometimes called the dependent sample t-test, is usually used to determine whether the mean difference between two sets of observations for the same subjects is zero. In a paired sample t-test, each participant or subject is measured twice. It is often used to determine whether an intervention brought about a change in some characteristic of respondents (e.g., respondents' math knowledge). To perform a paired t-test, respondents' math knowledge would be measured prior to the intervention, then the intervention would be performed (e.g., teaching a class on math), then respondent's math knowledge would be measured after the intervention. The change from before to after the intervention is used to assess whether the intervention was successful.

Panel Study

A type of longitudinal study in which data are collected from the same group of individuals (a panel) at two or more points in time. Although the sample selected for a panel study often include individuals (e.g., children, young adults), they may sample from other populations such as households, schools, and classrooms and collect data on these over a period of time.

Parameter

In statistics, a parameter is a characteristic of a population. It is a numerical quantity that tells us something about a population and is distinct from a statistic, which is a characteristic of a sample.

Participant Observation

A field research method whereby the researcher develops knowledge of the composition of a particular setting or society by taking part in the everyday routines and rituals alongside its members. A principle goal of participant observation is develop an understanding of a setting from a member's perspective, which may be accomplished through both informal observations and conversations as well as in-depth interviews.

Participant-As-Observer

The investigator takes part in the group activity that the researcher plans to study. The researcher also reveals to the group that s/he is studying the group's activities.

Participatory Action Research

Participatory action research (PAR) is a type of action research that involves stakeholders as equal partners. For example, researchers may work together with representatives from American Indian tribes to design a study that can be used to improve how health and education services are delivered to their people.

Path Analysis

A special use of multiple regression where the goal is to discern and assess the effects of a set of variables on an outcome. Path analysis is a form of analysis that looks explicitly at cause. The

pattern of relationships among variables is described by a path diagram with arrows used to indicate the directions of the causal relationships between them. Multiple regression is used to estimate the strength and direction of the relationships.

Pearson's Correlational Coefficient

Usually denoted by r , this is a measure of the degree to which two variables are associated. Pearson's correlation coefficient is used when the two variables are continuous. The coefficient can have values ranging from -1 to +1. If the coefficient is between 0 and +1, the variables are positively correlated, which means they both tend to increase (or decrease) in tandem. For example, children's height and weight are positively correlated because as height increases weight also tends to increase. If the coefficient is between 0 and -1, the variables are negatively correlated, which means as one increases the other decreases. For example, the number of days children are absent from school is negatively correlated with children's reading and math test scores because as the number of days children are absent increases their scores on reading and math tests decrease (when lower test scores indicate poorer performance in reading and math). The closer the coefficient is to either -1 or +1, the stronger the association between the two variables. This is also called a Product Moment Correlation.

Percentage

A proportion times 100.

Percentile

The percent of observations in a sample that have a value below a given score.

Pile Sorting

A task used to elicit judgments of similarity among items in a specific domain. The technique uses a set of index cards on which the name or short description of a domain item is written; the respondent is asked to sort them into piles according to their similarity.

Pilot Study

A small scale research study that is conducted prior to the larger, final study. The pilot study gives researchers a chance to identify any problems with their proposed sampling scheme, methodology, or data collection process. These studies are very useful in assessing strengths and weakness of a potential study. For example a pilot study might be conducted in a small sample of early childhood classrooms to identify potential difficulties when videotaping teacher-child interaction. Findings from the pilot are used to find solutions to these problems and to modify the approach that will be used in the larger study.

Point Estimate

A statistic calculated from a sample that is an estimate of some single characteristic of the population. For example, the sample mean is the point estimate of the population mean.

Poisson Distribution

A distribution that describes the number of events that occur in a certain time interval or spatial area. For example, the number of child care arrangements during a given period of time.

Poisson Regression

Poisson regression is a form of regression analysis used to model or predict outcomes (Y

variables) that are numerical counts (e.g., number of days a student is absent from school). It is best used for counts of rare events (e.g., number of unexcused absences over the school year).

Pooled Cross-Sectional Data

Pooled cross-sectional data (repeated cross-sectional data) are collected from different (independent) samples from the same population at different points in time. These data are contrasted with panel data that are collected from the sample at different time points. The early care and education program participation surveys administered as part of the National Household Surveys Program are examples of pooled cross-sectional data.

Population

In statistics, the population includes all members of a clearly defined group. The population can be comprised of a group of individuals (e.g., all children ages zero to 5) or of organizations (e.g., all programs providing early childhood education to 3- and 4-year old children). Samples are drawn from the population and the statistical results that are derived from random samples can be used to estimate characteristics of the whole population.

Potential Outcomes Framework of Causality

The potential outcomes framework of causality is based on the idea that every subject has different potential outcomes depending on their "assignment" to a condition (for example, assignment to treatment and control group or to different types or levels of treatment). It recognizes that the effect of assignment on an individual subject can only be directly observed for one condition. Statisticians use this framework to assess causality in observation (nonexperimental) studies.

Power

The degree to which a statistical test will detect significant differences between groups in a sample, when the differences do in fact exist. Sometimes statistical tests are not "powerful" enough to detect significant differences between groups in a sample that actually do exist in the population. The primary reason that a statistical test is not powerful is a small sample.

Practical Action Research

Practical action research is intended to address a specific problem. For example, a team of researchers, teachers, and parents may work together to design and execute a study whose purpose is to improve the communication between children's classroom teachers and their parents with an emphasis on ongoing bi-directional communication.

Precision

Precision of measurement is the degree to which repeated measurements under the same set of conditions show the same results. The standard error of measurement (SEM) is one way of calculating the precision of a measure or instrument (e.g., assessment of children's early literacy skills). The smaller the SEM, the more precise the measurement of the instrument. Precision of estimates refers to how close estimates derived from different samples are to each other. The standard error is a measure of precision of sample estimates. When the standard error is small, sample estimates are more precise; when the standard error is large, sample estimates are less precise.

Predictive Validity

To find out if a test or other measure assesses what it is intended to assess the correlation between the test score or value of the measure and a future external criterion. For example, the predictive validity of an assessment administered to children during the fall of the kindergarten year might be evaluated by correlating children's scores on this assessment with their scores on a standardized test administered in third grade. In general, the higher the correlation the stronger the predictive validity of the test.

Predictor Variable

The variable whose effect on an outcome variable is being modeled. A predictor variable is also called an "independent" variable.

Pretest

In research, pretest has several meanings. In experimental or evaluation research, it refers to a measure (e.g., children's performance on a test of their early literacy skills) taken before the experimental manipulation or intervention is applied or is introduced. It is also used to refer to the process of testing materials (e.g., questionnaires) and procedures (e.g., approach for collecting information about children from program records for use in selecting a sample of children) before they are used in a main study. For example, a parent questionnaire that will be used in a large study of preschool-age children may be administered first to a small (often non-random) sample of parents in order to identify any problems with how questions are asked and whether the response options that are offered to parents are adequate.

Primary Sampling Units

Primary sampling units refer to the sampling units that are selected at the first (primary) stage of a multi-stage sample. For example, a study of the instructional practices in early childhood classrooms might use a multi-stage sample design to select the classrooms to be studied. Using this approach, early childhood programs are selected first and then classrooms are selected from all the classrooms in the selected programs. Programs are the primary sample units (PSU) and classrooms are the secondary sampling units.

Principal Component Analysis

Principal component analysis is a procedure that is used to reduce a set of variables to a smaller set of variables, which are called principal factors. The principal factors retain most of the important information that is found in the larger set and are much easier to analyze and interpret.

Probability

A description of the likely occurrence of a particular event. Probability is conventionally expressed on a scale from 0 to 1; a rare event has a probability close to 0, a very common event has a probability close to 1.

Probability Sampling

A random sample of a population, which ensures that each member of the population has a chance of being selected for the sample.

Probability of Selection

In probability samples, the probability of selection is the probability that a member of the population will be selected to participate in the study sample.

Probit Models

A probit model is a type of regression where the dependent variable can only have two values. For example, a child from a low income family is either enrolled in a Head Start program or not.

Product Moment Correlation Coefficient

See Pearson's Correlation Coefficient.

Program Evaluation

Research that is conducted in order to determine the effectiveness of an intervention program.

Projection

Estimates of the future size and other demographic characteristics of a population, based on an assessment of past trends and assumptions about the future course of demographic behavior.

Propensity Score Matching

Propensity score matching is a statistical matching technique that is used to estimate the effect of a treatment or intervention when data come from a nonrandomized (observational) design. It uses a set of observable characteristics to predict the probability that participants will be assigned the treatment. Its purpose is to eliminate or reduce systematic differences between those who received the treatment and those who did not; thus, mimicking a randomized controlled trial design.

Proxy Variable

A variable used to "stand in" for another variable. Proxy variables are used when the variable of interest is not available in the data, either because it was not collected in the data or because it was too difficult to measure in a survey or interview.

Psychometric Properties

There are two important psychometric properties of any test or measure: reliability and validity. Reliability refers to the consistency of a measure and validity refers to the accuracy of the measure. (See definitions of reliability and validity elsewhere in the glossary.)

Purposive Sampling

A sampling strategy in which the researcher selects participants who are considered to be typical of the wider population. Since the sample is not randomly selected, the degree to which they actually represent the population being studied is unknown.

Q-Methodology/Q-Sort

Q Methodology is a research method used in psychology and social sciences to study people's point of viewpoint or judgments on a variety of topics. Participants are asked to sort decks of cards that contain statements, words, pictures, etc. into piles according to some criterion. For example, a child's mother may be asked to read and sort a series of statements about her child's behavior towards her in different situations.

Qualitative Research

A field of social research that is carried out in naturalistic settings and generates data largely through observations and interviews. Compared to quantitative research, which is principally concerned with making inferences from randomly selected samples to a larger population, qualitative research is primarily focused on describing small samples in non-statistical ways.

Quartiles

A set of three values that divides the total frequency into four equal parts.

Quasi-Experimental Research

Research in which individuals cannot be assigned randomly to two groups, but some environmental factor influences who belongs to each group. For example, if researchers want to look at the effects of smoking on health, they cannot ethically assign individuals to a group that smokes and a group that does not smoke. Researchers might rely on some environmental factor, for example an ad campaign that discourages smoking, to examine changes in health following the campaign. The theory behind quasi-experimental designs is that following an environmental intervention, individuals' characteristics play a smaller role in determining whether they smoke or do not smoke, and thus membership in these groups is closer to random assignment.

Questionnaire

A survey document with questions that are used to gather information from individuals to be used in research.

Quota Sampling

A non-probability sampling method in which a given number of subjects are selected from a specific group or groups. For example, a researcher might design a sample of 200 parents of newborns that sets quotas of 100 mothers and 100 fathers. Widely used in opinion polling and market research.

R-Squared

A measure of how well the independent, or predictor, variables predict the dependent, or outcome, variable. A higher R-square indicates a better model. The R-square denotes the percentage of variation in the dependent variable that can be explained by the independent variables. An Adjusted R-squared is a better comparison between models that have with different numbers of variables and different sample sizes than is the R-Squared. Please see Adjusted R-squared for more information.

Random Coefficient Model

Linear regression models assume that all individuals come from a population with a single slope (β). Random coefficient models relax this assumption and allow the slope to vary across individuals and for the slopes to be predicted by other variables in the model.

Random Effects Model

A random effects model is a statistical model (e.g., regression or analysis of variance) that assumes the independent variable(s) are random. Used when the levels of the variable(s) in the data are a subset of all possible values of the variable(s) (e.g., household income and children's

height and weight in inches and ounces). In contrast, a fixed effects model assumes that all the levels of the variable(s) of interest are in the data (e.g., children's race/ethnicity or sex).

Random Error

Error in measurement that is due to factors that cannot be controlled. Random errors are always present and are unpredictable. Sources of random error in measurement include fluctuations in the values obtained by an instrument (e.g., small differences in children's weight when measured twice using the same scale) and changes in what is being measured.

Random Sampling

A sampling technique in which individuals are selected from a population at random. Each individual has a chance of being chosen, and each individual is selected entirely by chance.

Random Selection

Random selection refers to the process of selecting individuals (schools, programs, classrooms) from the population to participate in a study. In random selection, each individual is chosen by chance and has a fixed and known probability of selection into the study sample.

Random Variable

A variable that numerically measures some characteristic of a sample, or population (e.g., height). The value of the variable will differ depending on which individual is measured (i.e., people are of different heights). The variable is said to be random because the variation in the value of the variable is due, at least in part, to chance (i.e., some people are just taller than other people).

Randomization

Assigning individuals in a sample to either an experimental group or a control group at random.

Range

A measure of how widely the data (values) for a specific variable are dispersed or spread. The larger the range the more dispersed the data. The range is calculated by subtracting the value of the lowest data point from the value of the highest data point. For example, in a sample of children between the ages of 2 and 6 years the range would be 4 years. When reporting the range, researchers typically report the lowest and highest value (Range = 2 - 6 years of age).

Rank Order Scale

A set of behaviors, objects or statements presented to research subjects which they are asked to rank (put them in some order) according to a specific criterion (e.g., size, importance, frequency). For example, parents may be presented with a number of factors that may affect their choices of child care and asked to order them in their importance.

Rating Scale

A rating scale is a measuring instrument for which judgments are made in order to rate a subject or case at a specified scale level with respect to an identified characteristic or characteristics.

Ratio

A ratio is a relationship between the number in two groups of objects. It tells us how many

times the number in the first group contains the number in the second group. For example, if we have a 10 elementary schools and 3 middle schools in a community, the ratio of elementary to middle schools is 10 to 3 (10:3), or roughly 3 elementary schools to every 1 middle school.

Ratio Scale

A scale in which the difference between the values on the scale are equivalent and the scale has a fixed zero point; values on the scale can be meaningfully measured against each other.

Raw Score

A score obtained from a test, assessment, observation, or survey that has not been converted to another type of score such as a standard score, percentile, ranking, or grade. Although a raw score often provides little useful information, there are exceptions. For example, if all children in a classroom are administered the exact same 20 items on a math test, the raw score is a measure of each child's absolute performance on the test and can be used to compare the math performance of the children in the class.

Refusal Rate

The proportion of all potentially eligible cases (subjects) in a sample that refuses to do an interview, survey, or to provide other requested data (for example, child record data).

Regression Analysis

A statistical technique that measures the relationship between a dependent (outcome) variable and one or more independent (predictor) variables (see linear, logistic and multiple regression).

Regression Coefficient

A coefficient that is calculated for each independent (predictor) variable. The regression coefficient indicates how much the dependent (outcome) variable will change, on average, with each unit change in the independent variables.

Regression Discontinuity Design

A regression discontinuity design (RDD) is a quasi-experimental design that provides unbiased estimates of the treatment effect (effect of an intervention). Participants are assigned to the treatment and control groups based on their score on an assignment variable (e.g., scores of a standardized assessment) that identifies the need for the intervention. Participants below a predetermined cut-point on the assignment variable are assigned to the treatment group (receive intervention) and those at or above the cut-point are assigned to the control group (do not receive intervention). The effect of the treatment is determined by comparing scores on the outcome variable(s) for participants in the two groups who are at or near the cut-point.

Regression Equation

A mathematical equation that indicates the relationship between a dependent (outcome) variable and one or more independent (predictor) variables. The equation indicates the extent to which the dependent variables can be predicted by knowing the value of the independent variables.

Reliability

Reliability is the degree to which an assessment or other measurement tool produces stable and consistent results. Reliability indicates the degree to which a measure will provide the same

result for the same person, across similar groups, and irrespective of who administers the assessment or collects the data. A reliable measure will always give the same result on different occasions, assuming that what is being measured has not changed during the intervening period.

Repeated Measures Analysis of Variance

Repeated measures analysis of variance (ANOVA), also called a within-subjects ANOVA or ANOVA with correlated samples, is a statistical method used to detect mean differences between groups that are not independent. It is used when examining differences in the means for a single group of subjects over three or more time points or under three or more different conditions. For example, a researcher might use a repeated measures ANOVA to evaluate the effects of a year-long intervention on children's reading achievement. A sample of children might be exposed to the intervention and their reading achievement measured at three time points (fall of program year, mid-year, and spring of the program year).

Replicability

The degree to which a scientific investigation can be easily repeated to see if its findings and outcomes can be tested again or by others. Replicability is an ideal in social science research, and is related to the reliability of study findings.

Representativeness

The idea that research subjects in a sample, as a group, represent the population from which the sample was selected.

Research Method

The approaches, tools, and techniques that researchers use to study a problem. These methods include laboratory experiments, field experiments, surveys, case studies, focus groups, ethnographic research, action research, and so forth.

Research Question

A clear statement in the form of a question of the specific issue that a researcher wishes to answer using data from one or more sources. Examples include: Do children who attend center-based early care and education programs have stronger academic and social skills than children who are cared for in a home-based child care setting? Does the Black-White achievement gap narrow or widen as children move through the elementary school grades?

Respondent

The person who responds to a survey questionnaire and provides information for analysis.

Response Categories

Pre-determined categories that limit the responses that can be given to a questionnaire item or an item on a standardized test. Response categories are required for closed-ended questions where a respondent must choose from the options provided. Response categories are often used when asking parents about their current employment. They are asked to indicate whether they are working, not working, looking for work, or attending school or a training program.

Response Rate

The number of complete interviews or surveys divided by the number of individuals who were

originally asked or selected to be interviewed or complete a survey (all eligible sample selections). The response rate is calculated using the same denominator as the refusal rate.

Rigorous Research

Rigorous research is research where the design, methods of data collection and analysis are appropriate to meet the stated objectives of an investigation.

Robustness

The state whereby a statistic remains useful even when one or more of its assumptions are violated.

Sample

A group that is selected from a larger group (the population). By studying the sample the researcher tries to draw valid conclusions about the population.

Sample Size

The number of subjects in a study. Larger samples are preferable to smaller samples, all else being equal.

Sampling

The process of selecting a subgroup of a population (i.e. sample) that will be used to represent the entire population.

Sampling Bias

Distortions that occur when some members of a population are systematically excluded from the sample selection process. For example, if interviews are conducted over the phone, only individuals with telephones will be in the sample. This could produce bias if the researcher intends to draw conclusions about the entire population, including those with a phone and those without a phone.

Sampling Design (Sample Design)

The part of the research plan that specifies the method of selection and the number of individuals or organizations (schools, programs) who will be selected and asked to participate in the study. The sampling design (sample design) specifies the target population, the frame or list from which cases from that population will be selected, the approach that will be used to select the sample members (simple random sampling, stratified sampling, cluster sampling, or combinations of these), the number of sample units to be selected to achieve the study objectives.

Sampling Distribution

The frequency with which data values appear in the sample. The sampling distribution can be characterized by the mean and the variance of the sample.

Sampling Error

This is the error that occurs because all members of the population are not sampled and measured. The value of a statistic (e.g., mean or percentage) that is calculated from different samples that are drawn from the same population will not always be the same. For example, if several different samples of 5 people are drawn at random from the U.S. population, the

average income of the 5 people in those samples will differ. (In one sample, Bill Gates may have been selected at random from the population, which would lead to a very high mean income for that sample.) It is not incorrect to have sampling error, and in fact statistical techniques take into account that sampling error will occur.

Sampling Frame

A list of the entire population eligible to be included within the specific parameters of a research study.

Scale

A group of survey questions that measures the same concept. For example, a researcher may be interested in individuals' gender role attitudes, and use several questions to determine their attitudes. This group of questions make up a gender role attitude scale.

Scaled Score

A mathematical transformation of a raw score so that scores can be compared across individuals and over time. The purpose of scaled scores is to report scores for all study participants on a consistent scale.

Scatter Plot

A display of the relationship between two quantitative or numeric variables. A scatter plot shows the value of one variable plotted against the value of another variable.

Selection Bias

Error due to systematic differences in the characteristics of those who are selected for a study and those who are not. For example, if a survey about parental attitudes toward child care in general is administered by randomly selecting parents whose children are enrolled in a university child care center, only parents whose children attend that program will be included in the sample. This will exclude parents whose children go to other child care centers and parents whose children are cared for in other types of settings (e.g., home-based child care). It may also result in a sample of parents who have children of a specific age (e.g., 2-4 year olds). The result of these exclusions is to introduce selection bias. Selection bias is a very serious problem in research, and it can negate research findings if the researcher does not carefully address the issue within the research study.

Selective Observation

The act of only attending to observations that correspond to current belief.

Self-Selection Sampling

Self-selection sampling is a type of non-probability sampling where the decision to take part in the study is left up to the potential participants themselves. For example, individuals may be invited to participate in a public opinion survey as part of a local news broadcast, or an invitation to participate in a survey may be sent to all members of an organization asking them to express their opinions about a particular issue.

Semantic Differential Scale

A type of rating scale that is designed to measure the meaning of things or objects. In research a semantic differential scale or a series of such scales are often used to measure attitudes.

Participants are presented an object, event or concept that is followed by a series of opposing adjectives separated by a sequence of unlabeled categories. Participants are asked to indicate their position relative to the two adjectives. For example, teachers may be asked how they would rate the professional development opportunities at their program using bipolar adjectives such as good-bad, interesting-boring, relevant-irrelevant.

Semi-Structured Interview

A method of data collection in which the interviewer uses a pre-determined list of topics or questions to gather information from a respondent. The interviewer, however, may stray from the list to follow-up on things the respondent says during the interview.

Sensitivity Analysis

Sensitivity analysis is a tool that is used to assess the robustness of the findings or conclusions based on primary analyses of data. It is used to assess the impact, effect or influence of key assumptions or variations—such as different methods of analysis, definitions of outcomes, protocol deviations, missing data, and outliers—on the overall conclusions of a study.

Sequential Hypothesis Testing

In statistics, sequential hypothesis testing or sequential analysis is a type of statistical analysis where the sample size is not set in advance. Instead, data are evaluated as they are collected, and sampling is stopped in accordance with a pre-defined stopping rule as soon as significant results are observed. An advantage of this approach is that fewer sampled cases may be required to meet the objectives of the research, thus reducing costs.

Significance Level

The probability of rejecting the null hypothesis in a statistical test when it is true (Type I Error). The significance level is set before the statistical analysis is undertaken. A commonly used significance level is .05, which indicates a 5% risk of concluding that a difference exists (group means are different or a correlation is different from zero) when there is no actual difference. If a statistical test (e.g., t-test or F-test) indicate that the changes of finding the observed results by chance are unlikely ($p < .05$) the findings are classified as statistically significant.

Simple Linear Regression

A statistical technique that measure the relationship between a dependent (outcome) variable and one independent (predictor) variable.

Simple Random Sampling

The basic sampling technique where a group of subjects (a sample) for study is selected from a larger group (a population). Each individual is chosen entirely by chance and each member of the population has an equal chance of being included in the sample.

Simulation

Simulation is a tool used by researchers to study complex problems and processes. Data are created according to a known model or theory and data analysis is used to explore how well the data fit the model under different sets of assumptions and conditions. Unlike in the real world, the researcher controls all of the factors affecting the data and can manipulate these systematically to see how each alone and in combination affect their findings.

Skewness

The tendency of the distribution of a statistic to depart from symmetry. Distributions can be skewed with more values to the right (positive) or to the left (negative). When the distribution is skewed, the median is a better measure of the midpoint of the distribution than the mean.

Slope

The coefficient of the independent variable indicating the change in dependent variable per unit change in the independent variable.

Snowball Sampling

A strategy used to gather a sample for a research study in which study participants give the researcher referrals to other individuals who fit the study criteria. Snowball samples cannot be generalized to the population because they are not selected randomly. Snowball samples are usually used to investigate groups that have some unique, rare, or unusual quality and groups where members know each other through an organization or common experience. For example, snowball samples might be used to identify parents who homeschool their children and attend local support groups.

Social Desirability

The tendency for respondents to give answers that are socially desirable or acceptable, but may not reflect their actual attitudes or behavior.

Social Network Analysis

Social network analysis (SNA) is the process of investigating social structures by mapping and measuring the relationships between people individually or in groups. Some examples of SNA include the study of friendship and acquaintance networks, family networks, and service delivery networks. SNA provides both a visual and a mathematical analysis of human relationships.

Sociogram

A display of networks of relationships among variables, designed to enable researchers to identify the nature of relationships that would otherwise be too complex to understand well enough to be able to describe.

Spurious Relationship

A statistical association between two variables that is produced by a third variable rather than by a causal link between the two original variables. For example, American children start school at the same time of year that the leaves begin to fall from the trees. This does not mean that leaves falling from trees affects when children start school or vice versa, instead both leaves falling from trees and children starting school occur during autumn.

Standard Deviation

A measure of variability or dispersion of a set of data. The standard deviation (SD) is the square root of the variance. It is calculated based on the difference between each individual observation and the mean observation.

Standard Error

A measure of the extent to which the sample mean fluctuates. The standard error is the

standard deviation (SD) of the sample means. Conceptually, the standard error of the mean would be calculated by selecting multiple samples at random from a population, calculating the mean for each of the samples, then calculating the standard deviation of these sample means. Because only one sample is generally drawn from a population for a research study, the standard error is calculated by dividing the sample deviation by the number of the observations in the sample. Generally speaking, the larger the sample, the smaller the standard error.

Standard Score

An indicator of the relative standing of a score within a normal distribution of scores, defined by its mean and standard deviation. It is calculated by subtracting the mean score from each score and dividing by the standard deviation of the scores in the distribution. The standard score, also known as the Z-score, tells us where a score lies in relation to the mean. Scores above the mean have a positive Z-score value and those below the mean have a negative value. Standard scores are used in norm-referenced assessment to compare one student's performance on a test to the performance of other students in the same age group or grade. Standard scores estimate whether a student's scores are above average, average, or below average compared to peers.

Standardization

A scale transformation procedure that involves manipulating data from different types of scales so that they can then be compared. It consists of subtracting the sample mean for the scale from each score on the scale and dividing by the scale's standard deviation.

Standardized Test

A standardized test is a test that is administered, scored, and interpreted in the same way for all test-takers. Scores on standardized test are often, but not always, created so that an individual's test score or the mean of a group's test scores can be compared to the mean of the population of individuals of the same age or grade. The Peabody Picture Vocabulary Test (PPVT) and the Woodcock-Johnson III Tests of Achievement are examples of commonly used standardized tests in early childhood research.

Standardized Variables

The values of variables from different types of scales are transformed so that they can be compared with one another.

Standardized, Open-Ended Interview

A standardized, open-ended interview is a type of qualitative interview that is similar in many ways to a survey. Questions are carefully scripted and written prior to the interview. The researcher asks a uniform series of questions in the same order to each interviewee. The questions are open-ended to capture more details and individual differences in responses.

Statistic

A measure of the characteristics of a sample (e.g., the mean is a statistic that measures the average of a sample). It gives an estimate of the same value for the population from which the sample was selected.

Statistical Analysis

Statistical analysis is the process of collecting, examining, manipulating, summarizing and interpreting quantitative or numerical data for the purpose of identifying patterns, trends, and relationships in the data. It can include the use of descriptive statistics such as percentages, means, variances and correlations and/or the use of inferential statistics such as t-tests, chi-square tests, regression, and analysis of variance (ANOVA).

Statistical Control

Statistical control refers to the technique of removing the effect of one independent variable or set of variables from the effects of the remaining variables on the dependent variable in a multivariate analysis. In statistical analyses with multiple independent variables and a single dependent or outcome variable, the effects of all other variables are held constant or their impact is removed to better analyze the relationship between the outcome variable and each of the independent variables. For example, if one was using multiple regression to examine the relationships between parental education, family income, and employment on parental choices for child care, the effects of income and education would be controlled when estimating the relationship between parent employment and child care choice.

Statistical Inference

Statistical inference is the process of drawing conclusions about or inferring the properties of a population based on analysis of data from a sample of that population.

Statistical Significance

Statistical significance refers to the probability or likelihood that the difference between groups or the relationship between variables observed in statistical analyses is not due to random chance. If there is a very small probability that an observed difference or relationship is due to chance (e.g., $p < .05$), the results are said to reach statistical significance. This means that the researcher concludes that there is a real difference between two groups or a real relationship between the observed variables. See Significance Level for additional information.

Stratification

Grouping the study population into subgroups by their homogenous characteristics before sampling so as to improve the representativeness of a sample.

Stratified Sampling

A statistical method for testing different theorized models, including the "structures" of relationships among the observed indicators and their underlying concepts.

Stratified sampling

A type of probability sample where the units in a population of interest are divided into mutually exclusive and collectively exhaustive groups or strata. A probability sample (e.g., simple random sample) is then drawn from each stratum. Stratified sampling provides greater precision than a simple random sample of the same size. It is also used to ensure representation of different groups (e.g., programs in different regions of the country) in the sample.

Structural Model

A theoretical model assumed to underlie the data that expresses the relationship between the dependent variable and the independent variables.

Structured Interview

The interviewer asks respondents the same questions using a predetermined series of interview questions. Deviations from the predetermined series of questions are not allowed in the interview process.

Structured Observation

Structured observation (or systematic observation) is a data collection method in which researchers gather data by observing the actions of individuals alone or during interactions with others. Research specifies in advance which behaviors are to be observed and how they are to be classified. Observers typically note the occurrence and frequency of the behaviors. Structured observations often are conducted in laboratory settings, but can be done in natural settings (e.g., children's homes) as well.

Subjectivity

A reflection of the person's mind, or thoughts, which is the result of his/her experiences, moods or attitudes.

Subjects

Those who participate in research and from whom data are collected.

Subsample

A sample selected from a larger sample. Subsamples are often selected to control data collection costs. For example, in a longitudinal study, the original (larger) sample may be divided into groups based on the per unit cost of data collection and a subsample (smaller sample) of high cost units (e.g., cases that have been difficult to complete in prior rounds or cases that entail high travel costs) selected.

Survey Research

A research approach designed to systematically collect data about a group of individuals. Data are obtained through direct questioning, using a written questionnaire (e.g., mail and online surveys) or interview (e.g., phone and in-person interviews). Data can be collected at a single point in time (cross-sectional survey) or from the same participants at multiple time points (longitudinal survey). While survey research most often involves collecting data from a sample of the population, population surveys or census surveys use survey methods (e.g., questionnaires, interviews).

Systematic Bias

When the recorded data from a sample is systematically higher or lower than the true data values within the population. Systematic bias can occur as a result of sampling bias or measurement bias. Sampling bias is an error in sampling when some subgroup of the target population is unintentionally left out of the sampling process. Measurement bias is an error in data collection when some occurrence distorts the responses in the same way (e.g., a test is

administered in a noisy classroom). Bias is a serious error in data collection and should be handled through a researcher's careful attention to sources of bias.

T Distribution

The T distribution (or Student's t-distribution) is a theoretical probability distribution that is similar to a normal distribution (i.e., it is a symmetrical bell-shaped distribution). It is used when testing differences in group means for small sample sizes (sample $n < 30$).

T-Test

A statistical test that is used to compare the means of two samples (independent t-test), the means of one sample at different times (paired sample t-test) or the mean of one sample against a known mean (one sample t-test). The test is appropriate for small sample sizes (less than 30) although it is often used when testing group differences for larger samples. It is also used to test whether correlation and regression coefficients are significantly different from zero.

Target Population

The population to which the researcher would like to generalize her or his results based on analysis of a sample. The sample is selected from a target population.

Test-Retest Reliability

The degree to which a measure produces consistent results over several administrations.

Theoretical Sampling

The selection of individuals within a naturalistic research study based on emerging findings as the study progresses to ensure that key issues are adequately represented.

Theory

General statement that describes a hypothesized relationship between different phenomena or characteristics. Theories should be specific enough to be testable with a well-designed research study.

Time Series

A sequence of observations or data which are ordered in time. Time series data are often collected at a regular interval, but the interval between time points does not have to be equal. The Head Start Program Information Report (PIR) is an example of a time series database. Data on Head Start enrollments and the characteristics of the programs serving Head Start children are collected and reported annually.

Time Series Analysis

Time series analysis or trend analysis is a set of statistical techniques used to identify patterns or associations in data across time. These methods can be used with continuous or discrete numeric data to examine univariate and multivariate time series models.

Time-Invariant Effects

Time-invariant effects refer to effects of a variable on another variable that are the same across time. For example, the effect of parental education on children's math achievement at time 1 is the same as its effect on math achievement at time n.

Time-Varying Covariate

A time-varying covariate (TVC) is a variable with values that can change across time, but its effect on the dependent (outcome) variable is assumed to be constant. For example, the amount of mathematics instructional time a child receives could be a TVC in a study of children's growth in math achievement from kindergarten through grade 5. Although the number of hours of math instruction children receive can change at each grade, the estimated relationship between math instruction and math achievement might remain constant across time.

Trajectory Modeling (Group-Based)

Group-based trajectory modeling is used to identify subgroups within a population that follow distinctive developmental trajectories. For example, within a cohort of young children there may be three such groups with different math achievement trajectories from kindergarten through grade 5: (1) accelerated growth in achievement; (2) steady growth in achievement; (3) declining growth in achievement.

Treatment Effect

Change in the outcome variable that is due to some intervention. Sometimes used to describe the change in an outcome variable that is due to changes in the independent (predictor) variable, even if the independent variable is not an intervention.

Treatment Offered

In experimental research, individuals or groups of individuals are randomly assigned to a treatment (intervention) or control group. Those assigned to the treatment group are offered the opportunity to participate in the treatment and those assigned to the control group are not. Intent-to-treat analyses assess the effect of being offered the treatment, while treatment-on-the-treated analyses assess the effect of participating in the treatment.

Treatment-on-the-Treated

Treatment-on-the-treated (TOT) is an estimate of the effect of a treatment or intervention on those who received the treatment regardless of their original assignment to treatment or control groups.

Triangulation of Data

Using of a variety of research methods or sources to collect data pertaining to a specific research problem or question. This process helps to enhance the validity of data since it does not overly rely on any particular method. It can also result in a more comprehensive understanding of the phenomenon under study.

Two-Tailed Test

A type of test that is used when a researcher is unsure of whether the independent (predictor) variable has a positive or negative effect on the dependent (outcome) variable, or whether the mean for group A is greater than or less than the mean for group B. A two-tailed test is also used when comparing the mean for a sample to a given value of x , when the mean can be either greater than or less than x . By contrast, a one-tailed test will test either that the mean is significantly greater than x or that the mean is significantly less than x , but not both.

Two-Way ANOVA

A statistical test to study the effect of two categorical independent variables on a continuous outcome variable. Two-way ANOVAs analyze the direct effect of the independent variables on the outcome, as well as the interaction of the independent variables on the outcome.

Type I Error

An error that occurs when a researcher concludes that a statistically significant difference between two groups or the relationship between two variables exists (based on the analysis of the sample), when in fact the difference or relationship does not exist in the population from which the sample was selected (null hypothesis is true). The amount of risk of a Type I Error a researcher is willing to accept is decided upon prior to analyzing the data. This probability of a Type I Error is also called a significance level.

Type II Error

An error that occurs when a researcher concludes that no significant relationship between two variables (based on analysis of sample data) when in fact the relationship does exist in the population from which the sample was drawn. The probability of not making a type II error is also called the power of a statistical test.

Typology

A descriptive, categorical scheme that may serve as the foundation for data coding or data analysis in a qualitative study. The types or categories comprising a typology should be mutually exclusive--separate and distinct with no overlap--and exhaustive--all the data must fall into some category.

Unbalanced Scale

A scale where the number of favorable and unfavorable categories is not the same.

Unbiased

A statistic that is free of systematic bias. Systematic bias occurs when the recorded data from a sample is systematically higher or lower than the true data values within the population. Systematic bias can occur as a result of sampling bias or measurement bias. Sampling bias is an error in sampling when some subgroup of the target population is unintentionally left out of the sampling process. Measurement bias is an error in data collection when some occurrence distorts the responses in the same way (e.g., a test is administered in a noisy classroom). Bias is a serious error in data collection and should be handled through a researcher's careful attention to sources of bias.

Unconditional Longitudinal Models

When modeling change or growth over time and its relationship to one or more predictor variables, researchers first test several models that include either no predictors or that include only time as a fixed effect. These unconditional models are used to initially identify the amount of variation in the outcome variable and the variation in outcome growth rates that is due to interindividual differences.

Unequal Variances (Heteroscedasticity)

Heteroscedasticity refers to values on a variable (dependent) that are unequally spread

(unequal variances) across the values on a second, predictor (independent) variable. Heteroscedasticity is the absence of homoscedasticity (equal variance in the dependent variable across values/levels of the independent variable), which is a key assumption of linear regression analysis.

Unfolding Questions

Unfolding questions include a sequence of questions designed to yield more complete and accurate data than would a single question on the same topic. They are used by survey researchers in an attempt to reduce missing data and measurement error. For example, asking respondents who are unwilling or unable to provide their exact household a set of follow-up questions designed to identify within which of a narrower set of income categories their income falls has been shown to significantly reduce the amount of missing income data.

Unit of Analysis

The individuals, groups of people, or objects that are being analyzed in a study. For example, if an analysis examines children's well-being, children are the unit of analysis. If an analysis examines family income, families are the unit of analysis. For some analyses, classrooms (the mean quality of early childhood classrooms) and programs (percentage of programs offering before- and after-care) are the units of analysis.

Univariate Analysis

Examination of the properties of one variable only and not the relationship between variables. Generally univariate analysis is performed by examining frequencies of response or values (counts and frequency distribution), one or more measures of central tendency (mode, mean and median) and the spread of responses or values (range, variance, standard deviation).

Unstructured Interview

An interview in which the researcher asks open-ended questions. The researcher aims to give respondents the latitude to talk freely on a topic and to influence the direction of the interview. There is no predetermined plan about the specific information to be gathered from these types of interviews.

Validity

The degree to which data and results are accurate reflections of reality. Validity refers to the concepts that are investigated, the people or objects that are studied; the methods by which data are collected; and the findings that are produced.

Variable

A measurable attribute or characteristics of a person, group or object that varies within the sample under investigation (e.g. age, weight, IQ, child care type). In research, variables are typically classified as dependent, independent, intervening, moderating, or as control variables (See definitions elsewhere in glossary).

Variance

A commonly used measure of dispersion for variables. The variance is calculated by squaring the standard deviation. The variance is based on the square of the difference between the values for each observation and the mean value.

Weighted Score

A score adjusted by such factors as the importance of the attribute assessed or the reliability and validity of the assessment from which the score was derived, or a combination of such factors.

Weighting

A process used to ensure that statistics produced from a sample are representative of the population from which the sample was drawn. In survey research, weights are used to adjust for differences in the probabilities of selection (i.e., members of the sample did not have an equal chance of being selected) and differences in the rates of nonresponse (i.e., some groups of respondents had higher rates of non-response than others). Most large surveys, and especially those that are national in scope, include one or more weights and statistics produced from such surveys should be based on weighted data. Using weights is one way to help reduce the risk of nonresponse bias in the survey findings.

Z Score

A score that is produced by subtracting the mean value from an individual data value and dividing by the standard deviation. This standardizes data values and allows for individual data values from different distributions (distributions with different means and standard deviations) to be compared.

Z Test

A statistical test that is used to compare the means of two independent samples or the mean of one sample with some fixed value. The test assumes that the populations from which the samples are drawn are normally distributed. It is used when testing differences for large samples (over 30 observations) and for smaller samples in which the variance of the population is known.

Prepared by: Child Care & Early Education Research Connections and Jerry West

Last updated: March 2019

Research Connections is a partnership between the National Center for Children in Poverty at the Mailman School of Public Health, Columbia University, and the Interuniversity Consortium for Political and Social Research at the Institute for Social Research, the University of Michigan, supported by a grant from the Office of Planning, Research and Evaluation in the Administration for Children and Families, U.S. Department of Health and Human Services. Contents are solely the responsibility of the authors and do not necessarily represent the official views of the Office of Planning, Research and Evaluation, the Administration for Children and Families, or the U.S. Department of Health and Human Services.