RESEARCH-BASED PRACTICE SYMPOSIUM - GLOSSARY OF RESEARCH TERMS

ANECDOTAL DATA/ ANECDOTAL EVIDENCE

Information that comes from individual, often biographical, accounts of experiences. Anecdotes are often appealing because of their richness and association with credible individuals, but they are not always factual and are not collected through RESEARCH. Thus, they are subject to bias and inaccuracy. They should not be the sole or primary basis for decision-making.

CASE STUDY

A Research design in which an array of DATA (including both quantitative and qualitative data) are collected about an individual, group, or organization to provide a detailed analysis of events or conditions within their real-life contexts. While similar in "feel" to Anecdotes, case studies involve collection of observable data coupled with systematic analysis. Findings from case studies should not be used to make claims about CAUSATION or GENERALIZABILITY.

CAUSATION

A relationship between two variables in which one causes the other to happen. A CORRELATION between two variables does not necessarily mean that one causes the other.

CONTROL/ CONTROL GROUP

In a RANDOM ASSIGNMENT study, the group not assigned to participate in the activity, program, or condition being studied. The control group can receive a condition that is similar to the TREATMENT (e.g., a different curriculum), a typical condition (e.g., the curriculum already being used), or nothing (e.g., no curriculum).

CORRELATION

Statistical correlation indicates the degree of linear relationship between two variables. A correlation coefficient always lies between -1 and +1. -1 indicates a perfect *negative* linear relationship between two variables; +1 indicates a perfect *positive* linear relationship; 0 indicates a lack of any linear relationship. Correlation is also frequently used to describe any relationship or association

CROSS-SECTIONAL

Cross-sectional DATA refer to observations of many individuals (e.g., students) at a given point in time. Annual student achievement scores are cross-sectional since they are "snapshots" of student performance in a given year. Cross sectional *studies* are those that record data from a sample of participants at a given point in time. Such studies can be problematic because they do not identify trends or show any type of growth. Care should be taken in making conclusions from such data.

DATA

Factual information organized and used to reason or make decisions. Data can include quantitative, statistical information as well as qualitative, descriptive information. Information that comes strictly from opinion or conjecture should not be used as data.

DESCRIPTIVE DATA

Descriptive data refers to statistical summaries and descriptions of a data set. Measures of central tendency (e.g. MEAN, MEDIAN) and variation (e.g. RANGE, STANDARD DEVIATION) are the main descriptive statistics.

EFFECT SIZE

Effect size is a measure of the magnitude of difference between TREATMENT and CONTROL groups. There are many ways to determine effect size, but it is commonly measures as the proportion of a Standard Deviation by which the outcomes of a treatment group exceeds the outcomes of a control group. Effect sizes are useful in determining whether findings that are statistically Significant are also substantively important. Effect sizes are used in META-ANALYSIS as a common measure that can be calculated for different studies and then combined into an overall summary.

EFFECTS

An effect reflects the contribution of a TREATMENT (e.g., curriculum) or FACTOR (e.g., student ability) to an OUTCOME (e.g., student achievement)

EVIDENCE

Refers to analyses and RESEARCH about educational programs that are helpful in forming a conclusion about their effectiveness. Evidence can be positive or negative and includes many forms of information, from descriptive DATA to case studies to RANDOM ASSIGNMENT studies.

EXPERIMENT

See RANDOM ASSIGNMENT

FACTOR

A variable that may affect an OUTCOME. FACTORS may be manipulated by researchers, as in a treatment condition (e.g., a curriculum). They may also not be manipulated, but still affect the outcome (e.g., student ability).

RESEARCH-BASED PRACTICE SYMPOSIUM - GLOSSARY OF RESEARCH TERMS

GENERALIZABILITY

Refers to the degree to which findings from a study based on a SAMPLE can be assumed to apply to a POPULATION. In general, RESEARCH designs that employ random sampling (in which members of a population are selected at random and all members have an equal chance of being selected) have the highest degree of generalizability. Research designs that include very small numbers of participants who are not selected at random have the least generalizability.

LONGITUDINAL

Longitudinal DATA refer to observations of a given unit made over time. Annual achievement scores of a group of students over a five-year period, for example, are longitudinal data. Longitudinal *studies* are those that record data for participants or variables over time. These studies are more appropriate for making decisions about programs or strategies since they can identify trends and growth. Studies that collect data over multiple years at the school level do not do this; these are CROSS-SECTIONAL studies.

MATCHED COMPARISON

A type of RESEARCH design in which the outcomes of a group of participants in a TREATMENT condition are compared to those of a group with similar characteristics not participating in a Treatment. For example, students being instructed using a particular curriculum program could be identified and compared to a group of students in the same grade and with the same demographics who are not being taught with that curriculum. Using statistical modeling, differences in achievement between the two groups can be attributed to the curriculum. However, there may still be differences between the two groups in unobservable characteristics (e.g., teacher quality). Similar terms: QUASI-EXPERIMENT

MEAN

For a POPULATION or a SAMPLE, the mean is the arithmetic average of all values (the sum of all values divided by the number of values).

MEDIAN

In a POPULATION or a SAMPLE, the median is the value that has just as many values above it as below it. If there is an even number of values, the median is the average of the two middle values. The median can also be defined as the 50th percentile.

META-ANALYSIS

Meta-analysis is a statistical technique that takes the results of two or more studies of the same research question and combines them into a single statistical analysis, typically incorporating EFFECT SIZES. The purpose of meta-analysis is to gain greater accuracy and statistical power by taking advantage of the large SAMPLE size resulting from the cumulation of results over multiple studies. Not to be confused with a REVIEW OF RESEARCH, which is a non-statistical synthesis of research studies.

OUTCOME

Refers to the results of a group in a RESEARCH study on the variable of interest. For example, in a study of a curriculum program, the outcome would be the achievement scores of students participating in the study, separated into TREATMENT and CONTROL outcomes.

P-VALUE

Statistical SIGNIFICANCE is expressed as a p-value, which represents the probability that a difference found between groups in a SAMPLE would occur even if there is, in fact, no difference between the groups in the POPULATION. A p-value of 0.05, for example, means that there is a 5% chance that a difference does not exist in a population even if it is found in a sample drawn from that population.

POPULATION

In statistical analysis, a population refers to the universe of individuals or groups from which a SAMPLE is drawn. Findings from samples can be used to make inferences about the population. For example, findings from a sample of study of 4th grade students could be used (if the study was properly designed and conducted) to make claims about all 4th grade students.

QUASI-EXPERIMENT (QE)

See MATCHED COMPARISON

RANDOM ASSIGNMENT

Random assignment is a method of assigning participants in a study to the TREATMENT or CONTROL group in which each individual is chosen for a group entirely by chance. Carried out correctly, random assignment results in groups that are similar on average in both observable and unobservable characteristics and any differences in outcomes between the two groups can be attributed to the treatment alone. The use of this method is sometimes referred to as "randomization". Similar terms: RANDOMIZED CONTROLLED TRIAL (RCT), EXPERIMENT

RESEARCH-BASED PRACTICE SYMPOSIUM - GLOSSARY OF RESEARCH TERMS

RANDOMIZED CONTROLLED TRIAL (RCT) — See RANDOM ASSIGNMENT

RANGE — Range is a measure of dispersion. It is defined as the difference between the highest and the lowest values.

REGRESSION ANALYSIS — Regression analysis provides a "best-fit" mathematical equation for the relationship between the dependent variable (OUTCOME) and independent variable(s) (FACTORS). This typically involves assuming a linear relationship between the dependent/outcome variable and the independent/explanatory variables.

RESEARCH — Systematic, empirical inquiry or investigation (based on observation and not theory or opinion) into a subject in order to discover or revise facts, theories, or application. Research on educational programs can be conducted in many ways. The term "research" refers to the act of conducting inquiry, not the outcomes or quality of inquiry.

REVIEW OF RESEARCH — Non-statistical analysis of a body of knowledge through summary, classification, and comparison of prior RESEARCH studies. Not to be confused with a META-ANALYSIS, which is a statistical procedure for aggregating the results of multiple research studies.

SAMPLE — A sample is a portion of a POPULATION. A sample is chosen to make inferences about the population by examining or measuring the elements in the sample. For example, a group of students could be selected to be representative of all students in a school or district rather than collecting DATA on every student in the school or district.

SCALED SCORE — A scaled score is a conversion of a student's raw test score to a common scale that allows for numerical comparison among students. Because most student testing programs use multiple versions of a test, the scale is used to control slight variations from one version of a test to the next. Scaled scores are quite useful for comparing test scores over time. On the TAKS, for example, a vertical scale score can be used to compare scores over time and across grade levels. The regular scale score can only be used to compare scores over time at the same grade level.

SIGNIFICANCE (STATISTICAL) — Refers to the probability that a difference in OUTCOMES between groups (for example, a TREATMENT group and a CONTROL group) is due to chance. The accepted standard to make a claim that a difference is <u>not</u> due to chance is a 95% or greater probability.

STANDARD DEVIATION — The standard deviation is a measure of the dispersion of values in a SAMPLE. It is the positive square root of the variance. The standard deviation expresses dispersion in the same units as the original values in the sample or POPULATION. For example, the standard deviation of a series of achievement scores is measured using the same range of scores used to score the test.

TREATMENT/ TREATMENT GROUP — In a RANDOM ASSIGNMENT study, the group assigned to participate in the activity, program, or condition being studied.