

## Glossary Research Terms

### *abstract:*

A brief, comprehensive summary of a research report that includes the research problem, a description of the participants, and an overview of the method, results and conclusions.

### *action research:*

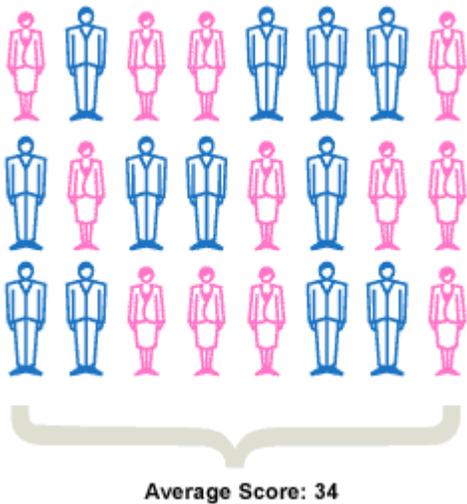
A type of research in which educators examine their own practice and evaluate strategies to improve practice and education outcomes. Most action research studies use descriptive research designs.

### *aggregated data:*

Data for which individual scores on a measure have been combined into a single group summary score.

### *Example:*

*In education research, it is common to aggregate individual student scores on an achievement test into a mean score for each school. Researchers then use the aggregate school achievement score for data analyses. Aggregating data reduces the sample size and obscures differences among individual scores.*



### *analysis of variance (ANOVA):*

A statistical technique used to test for statistically significant differences between two or more different groups of observations. An ANOVA produces  $F$ , an inferential test statistic.

### *attitude scale:*

A questionnaire that gathers information about participants' attitudes or beliefs concerning a particular topic based on the degree of intensity that they indicate in their responses.

### *bivariate correlation:*

A statistical correlation between two variables.

### *case study:*

A data collection method in which a single person, entity or phenomenon is studied in depth over a sustained period of time and through a variety of data.

### *Example:*

*A researcher conducts a yearlong case study of a school district that was awarded a grant to improve teacher quality. The researcher documents the processes used to implement the grant, interviews teachers and administrators, observes staff development, and measures student achievement before and after the grant was awarded.*

### *central tendency:*

A score in a set of scores or a frequency distribution that is typical or representative of all the scores. Measures of central tendency are the mean, median and mode.

*coding:*

In qualitative research, the process used to reduce information into categories or themes for data analysis and interpretation.

*coefficient of determination:*

For bivariate correlations, the coefficient of determination is defined as  $r^2$ , which is interpreted as the proportion of variation in the scores that is explained by the relationship between the variables. Note: Correlations indicate statistical, not causal, relationships.

*Example:*

*A researcher finds a correlation of  $r = .60$  between years of teaching experience and student achievement. The coefficient of determination of  $r^2 = .36$  means that 36% of the variation in achievement scores can be explained by the relationship between the two variables. (Conversely, 64% of the variation in achievement scores cannot be explained by the relationship.)*

*comparative descriptive research design:*

A research design in which data are collected to describe and compare two or more groups of participants or entities.

*Example:*

*A researcher identifies high-poverty schools in the state that have either high or low student achievement. The researcher describes the alignment or match between each school's curriculum and state standards and compares the high- versus the low-achieving schools to determine whether the degree of alignment is different.*

*comparison groups:*

The groups of participants who are being compared in a study, either based on different group characteristics or on having different treatments.

*confidence interval:*

A range of values that indicates the confidence or probability of observing a particular score or value in a population, usually expressed as standard deviation units above and below the mean. The wider the interval, the greater the confidence or probability that a particular value will be observed.

*Example:*

*Based on a random sample of 4th-grade reading scores, a researcher calculates the following 90% confidence interval for the mean of the population of 4th-grade reading scores:  $67 \pm 3.2$ . This indicates there is a 90% probability that the mean reading score of the population is between 63.8 and 70.2.*

*construct validity:*

The degree to which variables in a research study are considered by the education and research communities as acceptable representations of the constructs that the study concerns.

*Example:*

*One-on-one instruction is a valid representation of the construct of tutoring, while whole-class instruction would not be considered valid. Student scores on a standardized mathematics test are a valid representation of the construct of student achievement, while student scores on a survey about attitudes toward school would not be considered valid.*

*content validity:*

The degree to which the items on a measuring instrument (e.g., test or questionnaire) adequately cover the content that the instrument is designed to measure.

*control:*

The strategy used in scientific research to regulate the effects of variables in a study that are not intended to influence the results or conclusions.

*Example:*

*A researcher conducts a study of two different teacher preparation courses on how to teach mathematics. The researcher controls for differences among preservice students by randomly assigning the students to one of the two courses. The researcher controls for differences among course instructors by having a single instructor teach both courses.*

*control group:*

The group of participants in an experiment who do not receive the treatment that is being studied.

*convenience sample:*

A sample of participants selected for a research study based on their availability.

*Example:*

*A teacher educator conducts a research study of the preservice students enrolled in the traditional and alternative teacher preparation programs at the institution where the teacher educator is a faculty member. The sample is one of convenience because the preservice students are selected for the study based on their availability to participate.*

*correlation coefficient:*

A number that indicates the strength and direction of the statistical association between two or more variables.

Correlation coefficients vary between  $-1.00$  and  $+1.00$ . The higher the numerical value, the stronger the association. A correlation of  $0.00$  indicates the absence of an association. A positive sign means that as one variable increases, so does the other. A negative sign means that as one variable increases, the other variable decreases.

*Example:*

*A correlation coefficient of  $+0.63$  between the number of education courses and teacher test scores means that the more education courses that a teacher candidate completed, the higher the test score. A correlation of  $-0.63$  means that the more education courses that a teacher candidate completed, the lower the test score. Neither correlation coefficient, however, can support the existence of a causal relationship between courses and test scores because correlation is not causation.*

*correlational research:*

A type of research that seeks to establish an association or correlation between two or more variables. The fact that two or more variables are associated does not necessarily mean that one is a cause of the other(s).

*correlational research design:*

A research design in which data are collected to describe the statistical association between two or more variables.

*Example:*

*Bivariate correlation:*

*In School District X, a researcher collects data on beginning teachers' scores on the state licensing test (variable 1) and data on the achievement gains of each teacher's students (variable 2). The researcher then uses correlational statistics to measure the association between the two variables.*

*Multivariate correlation (also referred to as multiple regression):*

*In School District X, a researcher collects data on beginning teachers' scores on the state licensing test (variable 1), the number of college courses that each teacher completed in mathematics (variable 2), the amount of time that each teacher spent in school-based field experiences prior to certification (variable 3), and the achievement gains in mathematics by each teacher's students (dependent variable). The researcher uses multiple regression statistics to measure the association between the three teacher variables and student achievement gains and to estimate student achievement gains based on the contribution of each of the teacher variables to that association.*

*covariate:*

A variable that is correlated with another variable, such that when there is a change in one variable, there is a corresponding change in the other variable. Analysis of covariance is a statistical method that controls for the influence of covariates on the dependent variable in a research study.

*Example:*

*A researcher conducts a study on the influence of teacher professional development on principals' ratings of teacher performance. The researcher designates teaching experience as a covariate to statistically control its influences on principal ratings.*

*criterion variable:*

The dependent variable that is being predicted in a regression analysis.

*criterion-referenced test:*

A test for which a score is interpreted by comparing it to levels of performance established for the test by professionals in the field that the test addresses.

*Example:*

*Scores on the Colorado Student Assessment Program are assigned to the following categories based on the proficiency that students demonstrate in relation to state content standards: unsatisfactory, partially proficient, proficient and advanced.*

*cross-sectional research:*

A data-collection strategy in which data are collected at one point in time from participants who are at different developmental or grade levels. The purpose is to draw conclusions about differences between developmental groups.

*Example:*

*A researcher conducts a study of a new standards-based mathematics curriculum to determine whether the curriculum benefits students differently depending on their grade levels. The researcher compares gains in mathematics achievement by 2nd, 4th and 6th graders after their school adopts the new curriculum.*

*data:*

Factual information gathered as evidence for a research study.

*data-analysis plan:*

The plan for analyzing data in a research study. In a quantitative research study, the data-analysis plan provides details on statistical procedures. In a qualitative research study, the data-analysis plan provides details on coding procedures.

*data-collection instrument:*

A tool used to collect data in a research study such as a test, observation protocol or questionnaire.

*degrees of freedom (df):*

In statistics, the number of scores in a sample that are free to vary, calculated as sample size minus one ( $n - 1$ ). The degrees of freedom are used in the calculation of inferential statistics.

*dependent variable:*

The variable that is measured in a study. In an experimental research study, the dependent variable is affected by the independent variable. In a correlational research study, the dependent variable is associated with one or more other variables.

*Example:*

*Experimental research study:*

*A researcher randomly assigns teachers in a large elementary school to receive one of three types of professional development: (1) a class on instructional strategies, (2) a training program on how to increase student motivation or (3) a teacher discussion group. The researcher measures the differences in achievement gains among the students of the three teachers. The dependent variable is student achievement gains.*

*Correlational research study:*

*A researcher collects data on beginning teachers' scores on the state licensing test (variable 1) and data on the achievement gains of each teacher's students (variable 2). The researcher then uses the association between the two variables to estimate student achievement gains. The dependent variable is student achievement gains.*

*descriptive research:*

A type of research that has the goal of describing what, how or why something is happening.

*descriptive statistics:*

Statistics used to describe, organize and summarize data.

*Example:*

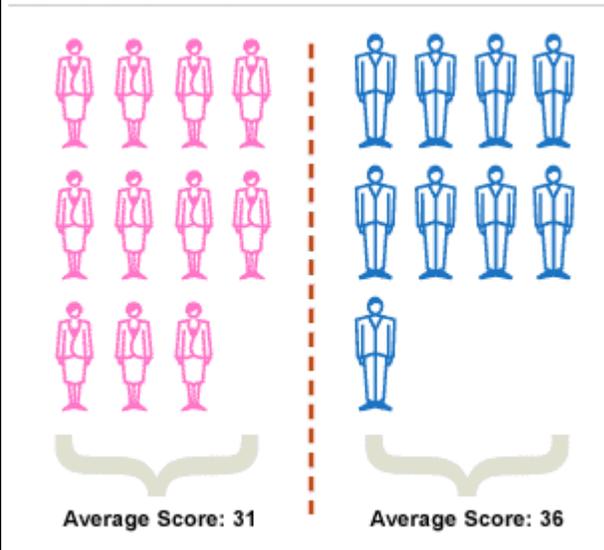
*Commonly used descriptive statistics include the mean, median, and standard deviation.*

*disaggregated data:*

Aggregated or grouped data that have been separated into individual component scores.

*Example:*

The No Child Left Behind Act requires schools to disaggregate student achievement data into the scores obtained by subgroups of students based on race/ethnicity, disability, socioeconomic level, gender, migrant status and English language proficiency.



*disconfirming evidence:*

A method used to verify the accuracy of data analyses in qualitative research by searching for evidence that negates the themes and categories that the researcher used to code and analyze the data.

*education research:*

The systematic gathering of empirical information to answer questions related to education.

*effect size:*

The degree to which a practice, program or policy has an effect based on research results, measured in standard deviation units. (Effect size is also referred to as practical significance.) A statistic commonly used to measure effect size is Cohen's  $d$ , which social scientists interpret as the following:  $d = .2$ , small;  $d = .5$  to  $.8$ , medium; and  $d = .8$  and higher, large.

*Example:*

A researcher finds an effect size of  $d = .5$  for the effect of an after-school tutoring program on reading achievement. This means (provided that the research study is valid) that the average student who participates in the tutoring program will achieve one-half standard deviation above the average student who does not participate. If the standard deviation is eight points, then the effect size translates into four additional points, which might increase a student's ranking on the test

*empirical information:*

Information based on something that can be observed. Students' test scores, observations of teachers' classroom instruction, principals' interview responses and school dropout rates are examples of empirical information in education research.

*empirical research:*

Research that seeks systematic information about something that can be observed in the real world or in the laboratory.

*ERIC:*

The Educational Resources Information Center, a federally funded source for literature on education research, including a searchable online database. (See <http://www.eric.ed.gov>)

*error:*

Inaccuracies in implementing a research study, including during sampling, treatment delivery, data recording or data analysis. Errors increase the variability of the data and threaten the validity of research conclusions.

*ethnography:*

A data-collection method in which information is collected about a group of individuals in their natural setting, primarily through observations.

*Example:*

*A researcher uses ethnography to study the challenges that face three beginning teachers at one elementary school. The researcher observes and documents the teachers in their classrooms, on the playground, in the teachers' lounge, at staff meetings, at parent conferences and in staff development sessions.*

*evaluation design:*

The plan for how data will be collected in an evaluation study. The evaluation design should be appropriate for the evaluation questions that the study addresses.

*evaluation question:*

The question(s) that an evaluation seeks to answer about a program. Evaluation questions can address program processes, program outcomes, links between the processes and outcomes, and explanations for the outcomes.

*evaluation study:*

A study designed to judge the effectiveness of an education program. Evaluation studies use some of the same research designs that research studies employ.

*Example:*

*A school district hires an evaluator to conduct a study on the effectiveness of an after-school tutoring program. The evaluator collects data about the student participants, their achievement before and after tutoring, the type and amount of tutoring that occurred, and the characteristics of the tutors. The evaluator also collects achievement data from a comparison group of students who applied too late to receive tutoring. The evaluation results include data about changes in student achievement as well as data about whether the program was implemented as planned.*

*experimental research:*

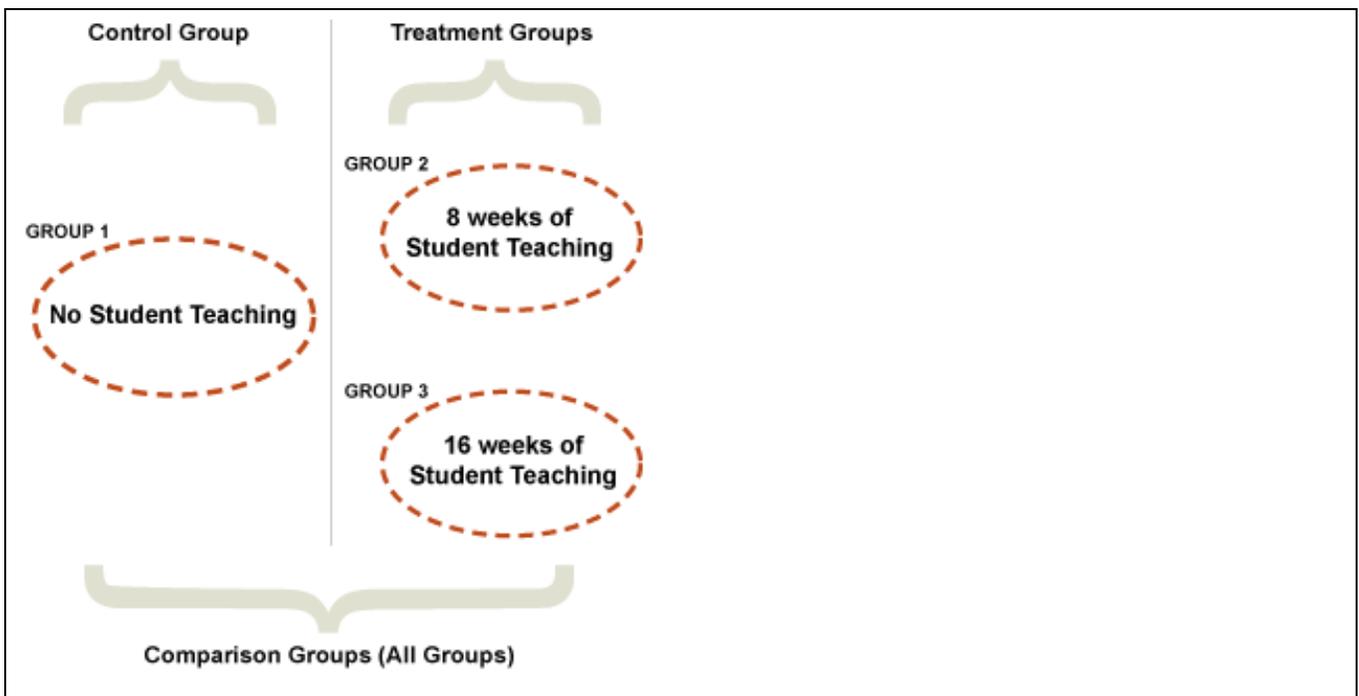
A type of research that has the goal of determining whether something causes an effect.

*experimental (true) research design:*

A research design in which (1) an independent variable is directly manipulated to measure its effect on a dependent variable, and (2) participants are randomly assigned to different groups that receive different amounts of the independent variable. (Also referred to as randomized field trials or randomized controlled trials.)

*Example:*

*A researcher randomly assigns 30 teacher preparation candidates to participate in one of three student teaching programs: (1) no student teaching, (2) eight weeks of student teaching or (3) 16 weeks of student teaching. After the candidates graduate, the researcher compares their scores on a performance-based teacher licensing test. The type of student teaching is the independent variable, and performance on the teacher-licensing test is the dependent variable. Groups 1 and 2 are the treatment groups because they participate in student teaching. Group 3 is the control group because the participants do not participate in student teaching. Together the three groups make up the comparison groups.*



**Independent Variable:** Type of Student Teaching  
**Dependent Variable:** Performance on Licensing Test

*ex post facto research:*

Descriptive research that examines the influence of a preexisting independent variable or treatment.

*Example:*

A researcher conducts a study to compare two reading programs. The participants are students in School A, which has been using Reading Program A for three years, and students in neighboring School B, which has been using Reading Program B for three years. This study is *ex post facto* because the research concerns effects from a preexisting treatment.

*external validity:*

The degree to which results from a study can be generalized to other participants, settings, treatments, and measures.

*extraneous variables:*

Variables in a research study that are not intended to influence the results or conclusions. Researchers use various methods to control the influence of extraneous variables.

*Example:*

A researcher conducts a study of the effects of two different reading curricula on 1st-grade reading achievement. *Extraneous variables in this study include students' verbal abilities and teachers' characteristics. The researcher needs to control the influence of these extraneous variables on achievement, possibly by having one teacher instruct both curricula and by randomly assigning students to the curricula.*

*factor analysis:*

A statistical procedure that reduces a set of items on a measuring instrument into a smaller number of dimensions called factors.

*Example:*

A researcher creates a 24-item questionnaire on teachers' classroom practices in language arts. A factor analysis reduces the 24 items into three factors. Factor one has eight items related to using drills and worksheets, factor two has six items related to independent reading, and factor three has 10 items related to whole-class instruction.

*focus group:*

A group of participants who are interviewed together and encouraged to share their opinions on a particular topic.

*frequency distribution:*

The frequency of occurrence of scores in a set. Frequency distributions can be represented in graphs or tables.

*Example:*

*Scores on a Mathematics Test: 51,52,51,55,55,53,58,50,55,58*

*Frequency Table:*

Mathematics Score	Frequency	Percent	Percentile
50	1	10.0	10.0
51	2	20.0	30.0
52	1	10.0	40.0
53	1	10.0	50.0
55	3	30.0	80.0
58	2	20.0	100.0
Total	10	100.0	

*generalization:*

The replication of research results in different contexts and with different populations.

*goodness-of-fit statistics:*

Statistics used to evaluate how well a set of scores or results conforms to a predicted frequency distribution or to a hypothesized model.

*grounded theory:*

A qualitative research method in which the researcher creates a theory from the categories that emerge from an extensive collection of qualitative data.

*hierarchical linear modeling (HLM):*

A statistical technique used to analyze data from participants who exist within different levels of a hierarchical structure.

*Example:*

*Student achievement data reflect influences from the family, classroom, grade, school, district, and state. Through HLM, the influences of these different levels on student achievement can be estimated.*

*history effect:*

A threat to the validity of research conclusions due to events that occur in the time between a pretest and a posttest. The longer the time span between a pretest and posttest, the more likely the occurrence of history effects.

*Example:*

*A researcher randomly assigns eight elementary schools to participate in Reform Model A and eight elementary schools to participate in Reform Model B. The researcher measures student achievement prior to implementation of the reform models (the pretest). After one school year, the researcher measures student achievement again (the posttest). Events that occur between the pretest and posttest can influence the results. For example, perhaps a large number of teachers in B schools enroll in graduate school, which improves their teaching.*

*hypothesis, null:*

A statement that an independent variable or treatment will have no effect. Researchers attempt to demonstrate through data that the null hypothesis is false.

*hypothesis, research:*

A statement about the researcher's expectations concerning the results of a study.

*Example:*

*Directional research hypothesis: A new standards-based mathematics curriculum will benefit elementary students at all grade levels.*

*Non-directional research hypothesis: A new standards-based mathematics curriculum will have different effects on elementary students depending on grade level.*

*independent variable:*

In experimental research, the variable that the researcher varies or manipulates to determine whether it has an effect on the dependent variable.

*Example:*

*As part of an experiment, a researcher randomly assigns teachers in a large elementary school to receive one of three types of professional development: (1) a class on instructional strategies, (2) a training program on how to increase student motivation, or (3) a teacher discussion group. The researcher measures the differences in achievement gains among the students of the three teachers. The independent variable is professional development.*

*inferential statistics:*

Statistics used to make inferences about a population based on the scores obtained from a sample. Inferential statistics are based on the mathematics of probability theory. Commonly used inferential statistics include *t*, *F* and Chi Square.

*internal validity:*

The degree to which the conclusions of a research study are supported by evidence and can be trusted.

*inter-rater reliability:*

The degree of agreement in the ratings that two or more observers assign to the same behavior or observation.

*intervening variable:*

An unmeasured variable that is assumed to intervene between a treatment or independent variable and a behavior or dependent variable. Most intervening variables are internal and cannot be observed. Their existence is inferred based on external measures.

*Example:*

*Learning is an intervening variable because it cannot be observed but is assumed to occur between instruction and performance based on measures such as tests.*

*intervention:*

A procedure, technique or strategy that is designed to modify an ongoing process. In research studies, the intervention also is referred to as a treatment. Most interventions in education are designed to modify directly or indirectly the student-learning process.

*interview:*

A data-collection method in which the researcher asks questions of individuals or groups and records the participants' answers. The interviewer usually asks the questions orally in a face-to-face interaction or over the telephone, but electronic interviews administered through e-mail also are possible.

*interview protocol:*

The planned questions and accompanying probes asked during an interview. Structured interview protocols ask specific objective questions in a predetermined order. Unstructured interview protocols ask open-ended questions and the order depends on interviewees' answers.

*latent variable:*

An unobserved and unmeasured variable that is hypothesized to have an influence on a dependent variable. Latent variables can be analyzed through the statistical technique of structural equation modeling (SEM).

*Likert Scale:*

A response scale in which participants respond to questionnaire items about their beliefs and attitudes by indicating varying degrees of intensity between two extremes such as like/dislike and agree/disagree.

*literature review:*

A comprehensive and systematic summary of past empirical research and/or evaluation studies on a specific topic. (Another term for a literature review is research synthesis.)

*longitudinal research:*

A data-collection strategy in which data are collected from the same participants at different points in time. The purpose is to draw conclusions about individual change over time.

*Example:*

*A researcher studies the mathematics achievement of students who were taught a new standards-based mathematics curriculum when they were in 6th grade. The researcher compares students' performances in mathematics achievement in grades 7, 8, and 9 to the performances of another group of students at each of those grade levels who were not taught the new curriculum in 6th grade. The purpose of the research is to determine whether change in mathematics performance over time is related the type of 6th-grade mathematics curriculum.*

*matching:*

A procedure used to select participants for comparison groups based on participant characteristics that are related to the dependent variable. Matching is frequently used in quasi-experimental studies when random assignment to groups is not feasible.

*Example:*

*A researcher assigns 15 teacher preparation candidates who have a seminar on Wednesdays to participate in eight weeks of student teaching. The researcher finds a group of 15 teacher preparation candidates who have a seminar on a different day and who are similar to the Wednesday group in the number and type of courses completed. The researcher assigns this second group of candidates to participate in 16 weeks of student teaching.*

*mean:*

In general, the average score in a set of scores or frequency distribution, calculated as the sum of the scores divided by the number of scores.

*Example:*

*The mean of the following set of five scores is 11:*

*9, 10, 10, 12, 14*

*median:*

The middle score in a set of scores or frequency distribution such that 50% of the scores are at or below the median score.

*Example:*

*The median of the following set of five scores is 10:*

*9, 10, 10, 12, 14.*

*member checking:*

A method used to verify the accuracy of data analyses in qualitative research by asking participants to review the findings and comment on the accuracy of the themes and categories that the researcher identified.

*meta-analysis:*

A comprehensive, systematic, quantitative review of past empirical research studies on a specific topic. Most meta-analyses examine only quantitative studies. Effect-size statistics are calculated to produce an overall conclusion about the various studies on the topic.

*Example:*

*A researcher conducts a meta-analysis of computer-assisted instruction in reading. The researcher examines 40 studies and calculates an overall effect size of  $d = .25$ , indicating a small positive effect of computer-assisted instruction on reading achievement.*

*mixed methods:*

The use of both quantitative and qualitative data-collection strategies in the same study. By providing more and different types of information related to the same research question, this approach can increase the reliability and applicability of research conclusions.

*mode:*

The most frequent score in a set of scores or a frequency distribution.

*Example:*

*The mode for the following set of five scores is 10:*

*9, 10, 10, 12, 14.*

*mortality:*

A threat to the validity of research conclusions due to the loss of participants from a study sample (also referred to as sample attrition).

*multiple methods:*

The use of more than one research method in a single research study, such as an experimental research study that includes descriptive research to verify that a treatment was implemented correctly.

*Example:*

*A researcher conducts an eight-week study of the effects of cooperative learning on student achievement. The researcher randomly assigns half of a teacher's students to participate in cooperative learning groups and the other half to participate in small-group instruction. To verify treatment implementation, the researcher conducts systematic observations of both the cooperative learning and the small-group instruction groups. This study uses both experimental and descriptive research methods.*

*multiple regression analysis:*

A statistical technique that determines the linear association between a set of predictor variables and a dependent variable and identifies the combination of predictor variables that best estimates the dependent variable (also referred to as the criterion variable).

*Example:*

*In School District X, a researcher collects data on beginning teachers' scores on the state licensing test (predictor 1), the number of college courses in mathematics that each teacher completed (predictor 2), the amount of time spent in school-based field experiences prior to certification (predictor 3), and the achievement gains in mathematics by each teacher's students (criterion variable). The researcher uses multiple regression statistics to measure the association between the three teacher variables and student achievement gains and to estimate student achievement gains based on the contribution of each of the teacher variables to that association.*

*N (n):*

The number of scores in a population (N) or a sample (n) of scores.

*narrative descriptions:*

Verbal descriptions of the information obtained from qualitative research such as descriptions of interview results.

*narrative review:*

A type of literature review in which research studies and their results are interpreted through narrative descriptions and qualitative comparisons.

*normal curve:*

The bell-shaped curve that results from the graph of a normal frequency distribution.

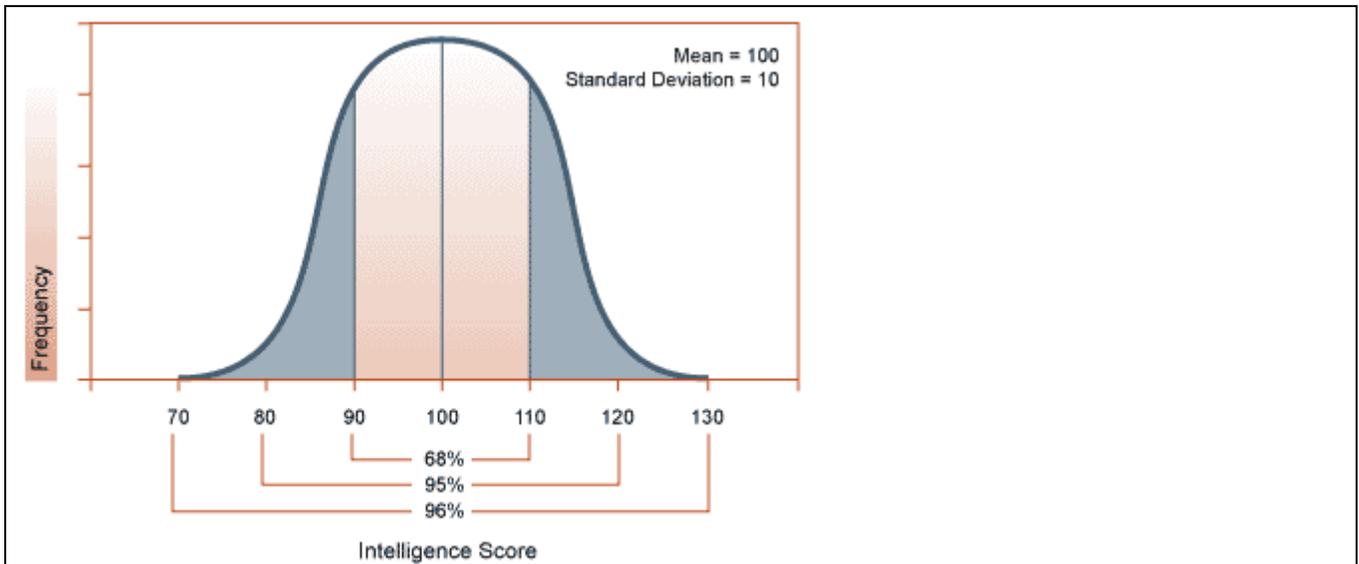
*normal curve equivalent (NCE) scores:*

Percentile scores from a normal frequency distribution that have been converted so there is an equal interval between each NCE score.

*normal distribution:*

A symmetrical frequency distribution in which the scores form a bell-shaped curve, and the mean, median and mode have the same value.

*Example:*



*norm-referenced test:*

A test for which a score is interpreted by comparing it to the scores of a comparison or norming group of persons who took the test. The similarity of an individual to the persons in the comparison group influences the accuracy of interpretation.

*Example:*

The SAT, which students take to gain admission to institutions of higher education, is a norm-referenced test. A score on the SAT is interpreted with reference to the scores of other students who took the test. A score of 500 on the SAT is considered average because that is the average score of the comparison or norming group of students.

*observation:*

The collection of data by documenting the occurrence of events in a setting. Observation is a common method of data collection in qualitative research.

*observation protocol:*

The plan for conducting observations of an event or behavior, including the frequency and duration of observations, and the definition of what will be observed.

*operational definition:*

A definition of a variable based on the methods used to measure or produce it.

*Example:*

An operational definition of student proficiency might be a score on an achievement test that is at or above 60% correct. An operational definition of an after-school tutoring program might be one-to-one tutoring of children by adults in reading and mathematics for two hours immediately after school, twice a week.

*percent:*

The proportion of participants who obtain a particular score in a frequency distribution.

*Example:*

In the following frequency distribution, 30% of the participants obtained a mathematics score of 55.

Mathematics Score	Frequency	Percent	Percentile
50	1	10.0	10.0
51	2	20.0	30.0
52	1	10.0	40.0
53	1	10.0	50.0
55	3	30.0	80.0
58	2	20.0	100.0
Total	10	100.0	

*percentile:*

The percent of participants who score at or below a particular score in a frequency distribution (also referred to as percentile rank).

*Example:*

*In the following frequency distribution, 80% of the participants obtained a mathematics score of 55 or lower, which means that a score of 55 is at the 80th percentile.*

Mathematics Score	Frequency	Percent	Percentile
50	1	10.0	10.0
51	2	20.0	30.0
52	1	10.0	40.0
53	1	10.0	50.0
55	3	30.0	80.0
58	2	20.0	100.0
Total	10	100.0	

*peer reviewed:*

A research study that has been critiqued by other researchers prior to publication or presentation at a research conference. (The quality of peer review varies among different publications and professional organizations.)

*phenomenological study:*

A qualitative research method in which the researcher conducts an in-depth and extensive study of participants' experiences of an event or situation from the participants' perspectives.

*pilot test:*

A trial run of all or some parts of a research study. Researchers often pilot test their data-collection procedures and instruments.

*population:*

All individuals or entities belonging to the group that is being studied.

*Example:*

*Examples of populations are all elementary school teachers in the United States, all schools in the Midwest, all 4th-grade students in Colorado, and all high school teachers in School District X.*

*practical significance:*

The degree to which a practice, program or policy has enough of an effect to justify its adoption. Practical significance usually is measured with statistics that calculate effect sizes.

*predictor variable:*

The variable in a regression analysis used to predict the value of a dependent variable.

*pretest-posttest research:*

Research in which participants take a pretest that measures the dependent variable prior to the administration of a treatment and a posttest that measures the dependent variable after the treatment is completed. The most valid approach to implementing pretest-posttest research is to randomly assign participants to two or more groups, one of which receives the treatment. The pretest-posttest difference scores are then compared for the groups.

*Example:*

*A researcher randomly assigns middle school students to participate in either an inquiry-based science unit or a traditional science unit. The students complete a test on problem solving before and after the unit. Because the problem-solving skills of the students in the inquiry-based group improved more than those of the students in the traditional group, the researcher concludes that inquiry-based units facilitate problem-solving skills.*

*primary source:*

A report on an original research study, usually written by the researcher(s), which includes details about the method and results.

*procedure:*

The specific steps that are taken to implement a research study.

*professional wisdom:*

The judgment that individuals acquire through experience, including the ability to incorporate local circumstances into practices and policies.

*proxy:*

A measure used to approximate the data sought when it is difficult to obtain a more precise measure due to constraints involving data collection or time.

*Example:*

*Average passing rate on state licensing tests by teacher candidates is a proxy measure for the quality of teacher preparation delivered by teacher education institutions.*

*purposive sample:*

A sample of participants selected for a research or evaluation study based on the information that they can provide related to the study.

*Example:*

*A researcher conducts case studies of four teacher preparation programs that received recognition for their effectiveness in preparing teacher candidates. The sample is purposive because the programs were chosen based on their recognition.*

*qualitative data:*

Narrative descriptions or observations.

*qualitative research:*

Research in which the data are narrative descriptions or observations. In most qualitative research, there is an emphasis on the influence of context.

*Example:*

*A researcher observes how teachers deliver instruction related to different reading curricula in two different schools. The researcher also interviews the teachers to understand their approaches to the different curricula and how their approaches might be influenced by school characteristics.*

*quantitative data:*

Numbers and measurements.

*quantitative research:*

Research in which the data are numbers and measurements. In quantitative research, there is an emphasis on control of the variables in the study.

*Example:*

*A researcher randomly assigns students to different reading curricula. At the end of the school year, the researcher examines the students' scores on a reading achievement test to determine whether the different curricula had different effects on reading.*

*quasi-experimental research design:*

A research design in which (1) an independent variable is manipulated to measure its effects on a dependent variable, and (2) participants are not randomly assigned to comparison groups.

*Example:*

*A researcher assigns 15 teacher preparation candidates who have a seminar on Wednesdays to participate in eight weeks of student teaching. The researcher assigns 15 teacher preparation candidates who have a seminar on Tuesdays to participate in 16 weeks of student teaching. After the candidates graduate, the researcher compares their scores on a performance-based teacher-licensing test. The amount of student teaching is the independent variable, and candidate performance on the teacher-licensing test is the dependent variable. The researcher does not randomly assign*

*candidates to the comparison groups. As a result, differences between the groups' performance on the test could be due to the amount of student teaching or due to other characteristics of the teacher candidates. The researcher should demonstrate that the candidates in the two groups do not differ in characteristics that are related to teaching performance.*

*random assignment:*

The assignment of participants to comparison groups using chance procedures so that every participant has the same probability of being selected to a group.

*random sample:*

A sample that is randomly drawn from a population so that each member of the population has an equal probability of being chosen for the sample.

*randomized trials:*

A "true experimental" research design in which (1) an independent variable is directly manipulated to measure its effect on a dependent variable (i.e., the treatment trial), and (2) participants are randomly assigned to different groups that receive different amounts of the independent variable (i.e., the treatment). (Also referred to as randomized field trials and randomized controlled trials.)

*range:*

The difference between the highest and lowest score in a set of scores or frequency distribution.

*Example:*

*The range for the following set of five scores is 5: 9, 10, 10, 12, 14.*

*raw score:*

An original score on a test or other measuring instrument prior to any score transformations.

*reactive measure:*

A measure toward which a participant is likely to react due to interactions with the researcher or the participant's assumption that certain responses are desirable.

*Example:*

*Interview questions are reactive measures because participants respond to actions by the interviewer that indicate approval or disapproval of their answers.*

*regression analysis:*

A statistical technique that uses the relationship between two variables, X and Y, to predict the value of X based on observations of Y.

*regression toward the mean:*

The tendency for extreme scores to move toward the average or mean score when a test or other measure is repeated. Regression effects threaten the validity of research conclusions in studies in which participants are chosen because of their extreme scores on a measure.

*Example:*

*Researchers often study schools in which students have extremely low achievement scores. If these students improve their achievement following a treatment or intervention, the improvement could be due to regression effects instead of treatment effects. In such studies, it is important to have comparison schools of students who also have extremely low achievement scores but who do not receive the treatment.*

*reliability (of a measuring instrument):*

The extent to which a measuring instrument produces consistent results when it is administered again under similar conditions.

*Example:*

*A reading test is reliable if students obtain similar scores when they take alternate but equivalent forms of the test within a short time span.*

*reliability coefficient:*

A correlation coefficient that indicates the degree of relationship between two sets of scores that result from persons

taking a test again under similar conditions. Reliability coefficients also indicate the degree of relationship among a set of items on a questionnaire or test.

*Example:*

*A test-retest reliability coefficient of .91 for a mathematics achievement test indicates that the test produces consistent results. A reliability coefficient of .51 for the internal consistency of an attitude questionnaire indicates that the questionnaire items have only a moderate relationship to one another.*

*repeated measures:*

A research study in which participants are measured two more times on the same dependent variable.

*Example:*

*A researcher conducts a study of the effects of an inquiry-based science unit on students' problem-solving skills. The researcher tests the students three times in the month following the unit to examine the duration of the effects.*

*replicate:*

To repeat a research study using the same method and similar participants. A successful replication obtains the same results as the original study.

*representative sample:*

A subset of a population used in a research study whose characteristics are generally reflective of the characteristics of the larger population that the sample is taken to represent. If a sample is not representative of the larger population, then any conclusions based on the sample might not hold for the larger population.

*Example:*

*To find out whether senior boys in a high school have different academic interests than senior girls, a researcher interviews 10% of the senior boys and girls. If this 10% does not have roughly the same proportion of white and minority students as the entire class, however, any conclusions the researcher draws from the sample might not reflect the interests of all of the senior boys and girls.*

*research design:*

The plan for how data will be collected in a research study. The research design should be appropriate for the research question that the study addresses. Research designs include simple descriptive, comparative descriptive, correlational, experimental and quasi-experimental.

*research ethics:*

The system of moral values established for the conduct of research and codified by professional associations and the United States Federal Government.

*research method:*

In a research report, the details on how a research study was conducted, including the research design, the data-collection instruments, and the procedure.

*research problem:*

The purpose of the research study, usually described in more general terms than research questions.

*Example:*

*A researcher conducts a study of a new standards-based mathematics curriculum to determine whether the curriculum benefits students at different grade levels differently. The research problem is whether the new mathematics curriculum has different effects at different grade levels.*

*research question:*

The question that a research study is designed to answer. Research questions include: What is happening? How is it happening? Why is it happening? Is something causing an effect?

*research synthesis:*

A comprehensive and systematic summary and review of past empirical research and/or evaluation studies on a specific topic. (Another term for a research synthesis is literature review.) Research syntheses can be quantitative or qualitative. Meta-analysis is the term used for a quantitative synthesis, and narrative review is the term used for a qualitative synthesis.

*researcher bias:*

Errors in the results of a research or evaluation study due to influences from the researcher's or evaluator's expectancies concerning study outcomes.

*Example:*

*A curriculum developer designs a new mathematics program for middle school students. If the developer conducts research on the effectiveness of the curriculum, the developer's expectancies could produce a positive bias in the results. To avoid researcher bias, persons and agencies that are external and independent from program developers should conduct the research.*

*response rate:*

The proportion of participants in a study who respond to a data-collection instrument; typically refers to the number of persons who complete and return a mailed questionnaire.

*rival explanation:*

An alternate explanation for research results that rivals the researcher's conclusions.

*Example:*

*A researcher randomly assigns eight elementary schools to participate in Reform Model A and eight elementary schools to participate in Reform Model B. The researcher measures student achievement prior to implementation of the reform models (the pretest). After one school year, the researcher measures student achievement again (the posttest). Because the students in the schools that used Reform Model B experienced achievement gains that were significantly higher than the students in schools that used Reform Model A, the researcher concludes that Model B caused greater achievement gains. The main rival explanation is that events that occurred between the pretest and posttest could have influenced the results. For example, perhaps a large number of teachers in Model B schools enrolled in graduate school, which improved their teaching. The researcher should demonstrate that historical events did not influence the results for either of the comparison groups.*

*sample:*

A subset of individuals or entities from a population.

*Example:*

*For the population of all 4th-grade students in Kansas, the 4th-grade students in the eastern half of the state would constitute a sample of the population (but not a random sample).*

*sample attrition:*

A threat to the validity of research conclusions due to the loss of participants from a study sample (also referred to as mortality).

*Example:*

*A researcher conducts a study of an after-school reading program on achievement gains. Twenty percent of the children drop out of the program. Conclusions about the effectiveness of the program are threatened by sample attrition because the students who remained could have special characteristics, for example, more motivation than those who left. Program effectiveness could be due to these individual characteristics and not the program characteristics.*

*sample size:*

The number of participants (e.g., students) or entities (e.g., schools) in a study sample. Large samples are preferred because, if randomly selected, they are more representative of the population than small samples.

*scaled questionnaire:*

A data-collection instrument that gathers information about participants' attitudes or beliefs concerning a particular topic based on the degree of intensity that they indicate in their responses. (Also called an attitude scale.)

*Example:*

*A scaled questionnaire on high school students' attitudes toward school might include a response scale and items such as the following:*

*Response Scale – Strongly Disagree, Disagree, Agree, Strongly Agree;*

*Item 1 – Teachers at my school are happy that I am in their classes;*

*Item 2 – I look forward to attending school each day.*

*scientifically-based research:*

According to the No Child Left Behind Act, research that is rigorous, systematic, objective, empirical, peer reviewed,

and relies on multiple measurements and observations, preferably through experimental or quasi-experimental methods. According to the National Research Council (2000), six principles underlie all scientific research:

- Pose significant questions that can be investigated empirically
- Link research to relevant theory
- Use methods that permit direct investigation of the question
- Provide a coherent and explicit chain of reasoning
- Replicate and generalize across studies
- Disclose research to encourage professional scrutiny and critique.

*secondary source:*

A description and/or summary of one or more prior research studies.

*selection bias:*

Systematic effects on the dependent variable that occur due to characteristics of the study participants.

*Example:*

*A researcher conducts a study on the influence of student teaching on teaching performance. The researcher assigns 20 teacher preparation candidates who attend college during the day to participate in 16 weeks of student teaching. The researcher assigns 20 candidates who are night students to eight weeks of student teaching. Selection bias in this study is likely because the characteristics of day and night students, such as age and motivation, might be different. The results could be due to these differences instead of the amount of student teaching.*

*simple descriptive research design:*

A research design in which data are collected to describe persons, organizations, settings or phenomena.

*Example:*

*A researcher surveys administrators of 10 alternative teacher preparation programs in order to describe the characteristics of the different programs.*

*standard deviation:*

A measure of the variability of the scores in a set of scores or a frequency distribution, equivalent to the average distance of the scores from the mean.

*Example:*

*The mean for the following set of five score is 11 and the standard deviation is 2: 9, 10, 10, 12, 14. The scores vary on average about two points from the mean.*

*For the following set of five scores, the mean is 10 and the standard deviation is 0: 10, 10, 10, 10, 10. There is no variation among the scores.*

*standard error of estimate:*

In a graph of the relationship between two variables, a measure equivalent to the average distance between the actual data points and the regression line.

*standard score:*

A score that transforms an original or raw score into standard deviation units in order to locate the score's position within a frequency distribution. Standard scores also are known as z-scores and are calculated as:  $z = \frac{\text{Raw Score} - \text{Mean}}{\text{Standard Deviation}}$ . The sign of a standard score (plus or minus) indicates whether it is above or below the mean.

*Example:*

*For the following set of five scores, the mean is 11 and the standard deviation is 2: 9, 10, 10, 12, 14. The score of 12 has a standard score of  $+1.00$ . The score of 9 has a standard score of  $-1.00$ .*

*standardized test:*

A test that has standard items and standard procedures for administration and scoring. Standardized tests are prepared by commercial test developers who establish the validity and reliability of the tests.

*Example:*

*The tests that are administered as part of the National Assessment of Educational Progress (NAEP) are standardized tests (see <http://nces.ed.gov/nationsreportcard/>).*

*statistical control:*

The use of statistics to isolate the effects of an extraneous variable on the dependent variable in a research study.

*Example:*

*A researcher conducts a correlational study of the relationship of student achievement in mathematics to the amount of time spent on whole-class instruction. To statistically control for the influence of students' prior achievement, the researcher uses a multiple regression analysis in which the predictor variables are prior achievement and instructional time, making it possible to estimate the separate effects of each variable on student mathematics achievement, the dependent variable.*

*statistical power:*

The likelihood that an inferential statistical test (e.g., *t*-test, Analysis of Variance) will detect a statistically significant result when an actual treatment effect exists. The power of a statistical test increases as the sample size increases.

*statistically significant:*

A result that has a low probability (e.g., 5 %) of occurring by chance. Because it is unlikely that a statistically significant result has occurred by chance, the result is said to reflect non-chance factors in the study, such as the effects of a treatment.

*statistics:*

Methods and rules for organizing and interpreting quantitative observations.

*stratified random sample:*

A sample of research participants that is randomly selected from different groups or strata in the population. The groups are defined based on one or more characteristics that might influence research results.

*Example:*

*In a study of the influence of state standards on mathematics achievement, a researcher divides the state's population of middle school students into males and females. The researcher randomly selects participants for the study from within each group. The proportion of male and female participants selected for the sample reflects the proportion of males and females in the middle school student population.*

*structural equation modeling (SEM):*

A statistical technique that tests a hypothesized network of linear relationships between observed and unobserved variables (also called latent variables).

*Example:*

*A researcher hypothesizes that teachers' years of experience and their perceptions of school culture influence how much they learn from staff development, which in turn influences student achievement. Teacher experience, perceptions of school culture, and student achievement are observed variables, and teacher learning is an unobserved or latent variable. The researcher uses SEM to test whether the hypothesized model is supported by the data that the researcher collects on the observed variables.*

*subjects:*

The participants whose behavior is examined in a research study.

*survey:*

A data-collection method in which participants provide information through self-report on questionnaires or in interviews.

*test:*

A data-collection instrument that gathers information about participants' knowledge and skills related to a particular topic based on their responses to a standard set of questions.

*theory:*

A set of interrelated principles proposed as an explanation for phenomena or observations (also referred to as a conceptual framework).

*Example:*

*Freud's theory of personality and Piaget's theory of child development are examples of social science theories. An example of a conceptual framework is an explanation of teacher professional development - in which teacher learning*

*influences instruction, which in turn influences student achievement.*

*threats to validity:*

Specific factors in a research study that threaten the validity or accuracy of research conclusions. (Also referred to as rival explanations.)

*Example:*

*The loss of participants from the treatment or control group is a threat to validity because those who remain in the study could be different from those who left. Also, if more participants leave one group than the other, then the two groups are no longer equivalent in non-treatment characteristics.*

*treatment:*

The program, policy or practice that is being studied through research or evaluation. Treatments are often interventions of some type such as a special reading program for low-achieving students. In an experimental research study, the treatment is the independent variable.

*treatment diffusion:*

The adoption of elements of the treatment in a research study by the participants who are in a control or a comparison group. Treatment diffusion (also called treatment spillover) threatens the validity of a conclusion that a treatment has no effect because both groups of participants experience the treatment.

*Example:*

*A researcher randomly assigns teachers in an elementary school either to participate in weekly professional development on integrating technology (the treatment group) with instruction or to have an extra weekly planning time (the control group). Treatment diffusion is likely because treatment teachers can discuss the new techniques they are learning with control teachers, who then might adopt these techniques.*

*treatment fidelity:*

The degree to which the treatment (e.g., a program or intervention) in a research or evaluation study is implemented as planned or intended.

*treatment group:*

The group of participants in an experiment who receive some amount of the independent variable (i.e., the program, policy or practice being studied).

*triangulation:*

Comparison of results obtained from the use of multiple research methods and/or data-collection strategies in a single study.

*Example:*

*A researcher randomly assigns half of the students in an after-school program to receive tutoring in reading and the other half to participate in a physical education class. The researcher examines students' gains in reading achievement and also interviews the students in each group about the effects of the after-school activity. The interview data are used to confirm the information about the effects of the after-school program obtained from the achievement data.*

*t-test:*

A statistical technique used to make inferences about a population of study participants based on a sample of these participants or to test for statistically significant differences between two different groups of observations.

*validity (of a measuring instrument):*

The degree to which an instrument measures what it is designed to measure and the degree to which it is used appropriately.

*Example:*

*A valid test of mathematics should measure mathematics knowledge or skills and should be correlated with other measures of mathematics ability. A valid use of this test is to make inferences about knowledge of mathematics, but using the test to make inferences about reading skills would be invalid.*

*validity (of a research study):*

The degree to which the conclusions of a research study are supported by evidence and can be trusted (also referred to as internal validity).

*variability:*

The amount of differences among scores in a distribution (i.e., a set of scores); the degree to which the scores are spread out or are clustered together. When all of the scores in a distribution are the same, there is no variability among the scores.

*variable:*

A characteristic or quantity that can change and have different values.

*Example:*

*Variables studied in education include characteristics of students (e.g., achievement), teachers (e.g., certification), schools (e.g., curriculum), districts (e.g., leadership), teacher preparation programs (e.g., accreditation), and states (e.g., education funding).*

*verification methods:*

Methods used in qualitative research to confirm the validity and reliability of the data coding and analyses.

## References and Resources

- Cooper, H. (1998). *Synthesizing research: A guide for literature reviews* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Creswell, J.W. (2002). *Research design: Qualitative, quantitative and mixed method approaches*. Thousand Oaks, CA: Sage Publications.
- Isaac S. & Michael, W. B. (1995). *Handbook in research and evaluation* (3rd ed.). San Diego: EdITS.
- Gravetter, F. J. and Wallnau, L. B. (1988). *Statistics for the behavioral sciences* (2nd ed.). St. Paul: West Publishing.
- McMillan, J. H. (2000). *Educational research: Fundamentals for the consumer* (3rd ed.). New York: Addison Wesley Longman.
- Shadish, W. R., Cook, T.D. and Campbell, D. T. (2002). *Experimental and quasi-experimental designs for causal inference*. Boston: Houghton Mifflin.
- Shanahan, T. (2000). "Research synthesis: Making sense of the accumulation of knowledge in reading." In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, and R. Barr (Eds.), *Handbook of reading research, volume III* (pp. 209–226). Mahwah, NJ: Lawrence Erlbaum and Associates
- Weiss, C. H. (1998). *Evaluation: Methods for studying programs and policies* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.