Appendix F:
Glossary from Methods in Behavioral Research
Appendix F: Glossary from Methods in Behavioral Research

alternative explanation  Part of causal inference; a potential alternative cause of an observed relationship between variables.

analysis of variance  See F test.

archival research  The use of existing sources of information for research. Sources include statistical records, survey archives, and written records.

autonomy (Belmont Report)  Principle that individuals in research investigations are capable of making a decision of whether to participate.

bar graph  Using bars to depict frequencies of responses, percentages, or means in two or more groups.

baseline  In a single case design, the subject's behavior during a control period before introduction of the experimental manipulation.

beneficence (Belmont Report)  Principle that research should have beneficial effects while minimizing any harmful effects.

carry-over effect  A problem that may occur in repeated measures designs if the effects of one treatment are still present when the next treatment is given.

case study  A descriptive account of the behavior, past history, and other relevant factors concerning a specific individual.

ceiling effect  Failure of a measure to detect a difference because it was too easy (also see floor effect).

central tendency  A single number or value that describes the typical or central score among a set of scores.

cluster sampling  A method of sampling in which clusters of individuals are identified. Clusters are sampled, and then all individuals in each cluster are included in the sample.

coding system  A set of rules used to categorize observations.

cohort  A group of people born at about the same time and exposed to the same societal events; cohort effects are confounded with age in a cross-sectional study.

conceptual replication  Replication of research using different procedures for manipulating or measuring the variables.

conclusion validity  Extent to which the conclusions about the relationships among variables reached on the basis of the data are correct.

concurrent validity  The construct validity of a measure is assessed by examining whether groups of people differ on the measure in expected ways.

confederate  A person posing as a participant in an experiment who is actually part of the experiment.

confidence interval  An interval of values within which there is a given level of confidence (e.g., 95%) where the population value lies.

confounding  An uncontrolled variable varies systematically with an independent variable; it is impossible to separate the effect of the independent variable from the confounding variable.

confounding variable  A variable that is not controlled in a research investigation. In an
experiment, the experimental groups differ on both the independent variable and the confounding variable.

**construct validity** The degree to which a measurement device accurately measures the theoretical construct it is designed to measure.

**content analysis** Systematic analysis of the content of written records.

**content validity** An indicator of construct validity of a measure in which the content of the measure is compared to the universe of content that defines the construct.

**control series design** An extension of the interrupted time series quasi-experimental design in which there is a comparison or control group.

**convergent validity** The construct validity of a measure is assessed by examining the extent to which scores on the measure are related to scores on other measures of the same construct or similar constructs.

**correlation coefficient** An index of how strongly two variables are related to each other.

**correlational method** See nonexperimental method.

**counterbalancing** A method of controlling for order effects in a repeated measures design by either including all orders of treatment presentation or randomly determining the order for each subject.

**covariation of cause and effect** Part of causal inference; observing that a change in one variable is accompanied by a change in a second variable.

**Cronbach’s alpha** An indicator of internal consistency reliability assessed by examining the average correlation of each item (question) in a measure with every other question.

**cross-sectional method** A developmental research method in which persons of different ages are studied at only one point in time; conceptually similar to an independent groups design.

**curvilinear relationship** A relationship in which increases in the values of the first variable are accompanied by both increases and decreases in the values of the second variable.

**debriefing** Explanation of the purposes of the research that is given to participants following their participation in the research.

**degrees of freedom (df)** A concept used in tests of statistical significance; the number of observations that are free to vary to produce a known outcome.

**demand characteristics** Cues that inform the subject how he or she is expected to behave.

**dependent variable** The variable that is the subject's response to, and dependent on, the level of the manipulated independent variable.

**descriptive statistics** Statistical measures that describe the results of a study; descriptive statistics include measures of central tendency (e.g., mean), variability (e.g., standard deviation), and correlation (e.g., Pearson r).

**discriminant validity** The construct validity of a measure is assessed by examining the extent to which scores on the measure are not related to scores on conceptually unrelated measures.

**effect size** The extent to which two variables are associated. In experimental research, the magnitude of the impact of the independent variable on the dependent variable.

**electroencephalogram (EEG)** A measure of the electrical activity of the brain.

**electromyogram (EMG)** A measure of the electrical activity of muscles, including muscle tension.

**empiricism** Use of objective observations to answer a question about the nature of behavior.

**error variance** Random variability in a set of scores that is not the result of the independent variable. Statistically, the variability of each score from its group mean.

**exact replication** Replication of research using the same procedures for manipulating
and measuring the variables that were used in the original research.

**Experimental method** A method of determining whether variables are related, in which the researcher manipulates the independent variable and controls all other variables either by randomization or by direct experimental control.

**Experimenter bias (expectancy effects)** Any intentional or unintentional influence that the experimenter exerts on subjects to confirm the hypothesis under investigation.

**External validity** The degree to which the results of an experiment may be generalized.

**Extraneous variable** See third variable.

**F test (analysis of variance)** A statistical significance test for determining whether two or more means are significantly different. F is the ratio of systematic variance to error variance.

**Face validity** The degree to which a measurement device appears to accurately measure a variable.

**Factorial design** A design in which all levels of each independent variable are combined with all levels of the other independent variables. A factorial design allows investigation of the separate main effects and interactions of two or more independent variables.

**Falsifiability** The principle that a good scientific idea or theory should be capable of being shown to be false when tested using scientific methods.

**Field experiment** An experiment that is conducted in a natural setting rather than in a laboratory setting.

**Filler items** Items included in a questionnaire measure to help disguise the true purpose of the measure.

**Floor effect** Failure of a measure to detect a difference because it was too difficult (also see ceiling effect).

**Frequency distribution** An arrangement of a set of scores from lowest to highest that indicates the number of times each score was obtained.

**Frequency polygon** A graphic display of a frequency distribution in which the frequency of each score is plotted on the vertical axis, with the plotted points connected by straight lines.

**Functional MRI** Magnetic resonance imaging uses a magnet to obtain scans of structures of the brain. Functional magnetic resonance imaging (fMRI) provides information on the amount of activity in different brain structures.

**Galvanic skin response (GSR)** The electrical conductance of the skin, which changes when sweating occurs.

**Haphazard (convenience) sampling** Selecting subjects in a haphazard manner, usually on the basis of availability, and not with regard to having a representative sample of the population; a type of nonprobability sampling.

**History effect** As a threat to the internal validity of an experiment, refers to any outside event that is not part of the manipulation that could be responsible for the results.

**Hypothesis** A statement that makes an assertion about what is true in a particular situation; often, a statement asserting that two or more variables are related to one another.

**Independent groups design** An experiment in which different subjects are assigned to each group. Also called between-subjects design.

**Independent variable** The variable that is manipulated to observe its effect on the dependent variable.

**Inferential statistics** Statistics designed to determine whether results based on sample data are generalizable to a population.

**Informed consent** In research ethics, the principle that participants in an experiment be informed in advance of all aspects of the research that might influence their decision to participate.
Institutional Review Board (IRB)  An ethics review committee established to review research proposals. The IRB is composed of scientists, nonscientists, and legal experts.

Instrument decay  As a threat to internal validity, the possibility that a change in the characteristics of the measurement instrument is responsible for the results.

Interaction  The differing effect of one independent variable on the dependent variable, depending on the particular level of another independent variable.

Internal consistency reliability  Reliability assessed with data collected at one point in time with multiple measures of a psychological construct. A measure is reliable when the multiple measures provide similar results.

Internal validity  The certainty with which results of an experiment can be attributed to the manipulation of the independent variable rather than to some other, confounding variable.

Interrater reliability  An indicator of reliability that examines the agreement of observations made by two or more raters (judges).

Interrupted time series design  A design in which the effectiveness of a treatment is determined by examining a series of measurements made over an extended time period both before and after the treatment is introduced. The treatment is not introduced at a random point in time.

Interval scale  A scale of measurement in which the intervals between numbers on the scale are all equal in size.

Interviewer bias  Intentional or unintentional influence exerted by an interviewer in such a way that the actual or interpreted behavior of respondents is consistent with the interviewer’s expectations.

IV × PV design  A factorial design that includes both an experimental independent variable (IV) and a nonexperimental participant variable (PV).

Justice (Belmont Report)  Principle that all individuals and groups should have fair and equal access to the benefits of research participation as well as potential risks of research participation.

Latin square  A technique to control for order effects without having all possible orders.

Longitudinal method  A developmental research method in which the same persons are observed repeatedly as they grow older, conceptually similar to a repeated measures design.

Main effect  The direct effect of an independent variable on a dependent variable.

Manipulation check  A measure used to determine whether the manipulation of the independent variable has had its intended effect on a subject.

Matched pairs design  A method of assigning subjects to groups in which pairs of subjects are first matched on some characteristic and then individually assigned randomly to groups.

Maturation effect  As a threat to internal validity, the possibility that any naturally occurring change within the individual is responsible for the results.

Mean  A measure of central tendency, obtained by summing scores and then dividing the sum by the number of scores.

Measurement error  The degree to which a measurement deviates from the true score value.

Median  A measure of central tendency; the middle score in a distribution of scores that divides the distribution in half.

Meta-analysis  A set of statistical procedures for combining the results of a number of studies in order to provide a general assessment of the relationship between variables.

Mixed factorial design  A design that includes both independent groups (between-subjects) and repeated measures (within-subjects) variables.
mode A measure of central tendency, the most frequent score in a distribution of scores.

moderator variable A variable that influences the nature of the relationship between two other variables (an independent variable and a dependent variable). In a factorial design, the effect of the moderator variable is revealed as an interaction.

mortality The loss of subjects who decide to leave an experiment. Mortality is a threat to internal validity when the mortality rate is related to the nature of the experimental manipulation.

multiple baseline design Observing behavior before and after a manipulation under multiple circumstances (across different individuals, different behaviors, or different settings).

multiple correlation A correlation between one variable and a combined set of predictor variables.

naturalistic observation Descriptive method in which observations are made in a natural social setting. Also called field observation.

negative case analysis In field observation, an examination of observations that do not fit with the explanatory structure devised by the researcher.

negative linear relationship A relationship in which increases in the values of the first variable are accompanied by decreases in the values of the second variable.

nominal scale A scale of measurement with two or more categories that have no numerical (less than, greater than) properties.

nonequivalent control group design A quasi-experimental design in which nonequivalent groups of subjects participate in the different experimental groups, and there is no pretest.

nonequivalent control group pretest-posttest design A quasi-experimental design in which nonequivalent groups are used, but a pretest allows assessment of equivalency and pretest-posttest changes.

nonexperimental method Use of measurement of variables to determine whether variables are related to one another. Also called correlational method.

nonprobability sampling Type of sampling procedure in which one cannot specify the probability that any member of the population will be included in the sample.

no relationship Outcome of research in which two variables are not related; changes in the first variable are not associated with changes in the second variable.

null hypothesis The hypothesis, used for statistical purposes, that the variables under investigation are not related in the population, that any observed effect based on sample results is due to random error.

one-group posttest-only design A quasi-experimental design that has no control group and no pretest comparison; a very poor design in terms of internal validity.

one-group pretest-posttest design A quasi-experimental design in which the effect of an independent variable is inferred from the pretest-posttest difference in a single group.

operational definition Definition of a concept that specifies the operation used to measure or manipulate the concept.

order effect In a repeated measures design, the effect that the order of introducing treatment has on the dependent variable.

ordinal scale A scale of measurement in which the measurement categories form a rank order along a continuum.

panel study In survey research, questioning the same people at two or more points in time.

partial correlation The correlation between two variables with the influence of a third variable statistically controlled for.

path analysis A method used to develop models of possible relationships among a set of variables that were studied with the nonexperimental method.
Pearson product-moment correlation coefficient A type of correlation coefficient used with interval and ratio scale data. In addition to providing information on the strength of relationship between two variables, the Pearson product-moment correlation coefficient indicates the direction (positive or negative) of the relationship.

peer review The process of judging the scientific merit of research through review by peers of the researcher—other scientists with the expertise to evaluate the research.

pie chart Graphic display of data in which frequencies or percentages are represented as “slices” of a pie.

pilot study A small-scale study conducted prior to conducting an actual experiment; designed to test and refine procedures.

placebo group In drug research, a group given an inert substance to assess the psychological effect of receiving a treatment.

population The defined group of individuals from which a sample is drawn.

positive linear relationship A relationship in which increases in the values of the first variable are accompanied by increases in the values of the second variable.

posttest-only design A true experimental design in which the dependent variable (posttest) is measured only once, after manipulation of the independent variable.

power The probability of correctly rejecting the null hypothesis.

prediction A statement that makes an assertion concerning what will occur in a particular research investigation.

predictive validity The construct validity of a measure is assessed by examining the ability of the measure to predict a future behavior.

pretest-posttest design A true experimental design in which the dependent variable is measured both before (pretest) and after (posttest) manipulation of the independent variable.

probability The likelihood that a given event (among a specific set of events) will occur.

probability sampling Type of sampling procedure in which one is able to specify the probability that any member of the population will be included in the sample.

program evaluation Research designed to evaluate programs (e.g., social reforms, innovations) that are designed to produce certain changes or outcomes in a target population.

pseudoscience Claims that are made on the basis of evidence that is designed to appear scientific; such evidence is not based on the principles of the scientific method, however.

psychobiography A type of case study in which the life of an individual is analyzed using psychological theory.

purposive sample A type of haphazard sample conducted to obtain predetermined types of individuals for the sample.

quasi-experimental design A type of design that approximates the control features of true experiments to infer that a given treatment did have its intended effect.

quota sampling A sampling procedure in which the sample is chosen to reflect the numerical composition of various subgroups in the population. A haphazard sampling technique is used to obtain the sample.

randomization Controlling for the effects of extraneous variables by ensuring that the variables operate in a manner determined entirely by chance.

ratio scale A scale of measurement in which there is an absolute zero point, indicating an absence of the variable being measured. An implication is that ratios of numbers on the scale can be formed (generally, these are physical measures such as weight or timed measures such as duration or reaction time).

reactivity A problem of measurement in which the measure changes the behavior being observed.
regression equation  A mathematical equation that allows prediction of one behavior when the score on another variable is known.

regression toward the mean  Also called statistical regression; principle that extreme scores on a variable tend to be closer to the mean when a second measurement is made.

reliability  The degree to which a measure is consistent.

repeated measures design  An experiment in which the same subjects are assigned to each group. Also called within-subjects design.

replication  Repeating a research study to determine whether the results can be duplicated.

research hypothesis  The hypothesis that the variables under investigation are related in the population—that the observed effect based on sample data is true in the population.

response rate  The percentage of people selected for a sample who actually completed a survey.

response set  A pattern of individual response to questions on a self-report measure that is not related to the content of the questions.

reversal design  A single-case design in which the treatment is introduced after a baseline period and then withdrawn during a second baseline period. It may be extended by adding a second introduction of the treatment. Sometimes called a “withdrawal” design.

role-playing  A procedure for studying behavior in which individuals are asked to indicate how they would respond to a given situation rather than being observed in action in the situation.

sampling  The process of choosing members of a population to be included in a sample.

sampling frame  The individuals or clusters of individuals in a population who might actually be selected for inclusion in the sample.

selection differences  Differences in the type of subjects who make up each group in an experimental design; this situation occurs when participants elect which group they are to be assigned to.

sensitivity  The ability of a measure to detect differences between groups.

sequential method  A combination of the cross-sectional and longitudinal design to study developmental research questions.

simple main effect  In a factorial design, the effect of one independent variable at a particular level of another independent variable.

simple random sampling  A sampling procedure in which each member of the population has an equal probability of being included in the sample.

single case experiment  An experiment in which the effect of the independent variable is assessed using data from a single participant.

split-half reliability  A reliability coefficient determined by the correlation between scores on half of the items on a measure with scores on the other half of a measure.

standard deviation  The average deviation of scores from the mean (the square root of the variance).

statistical significance  Rejection of the null hypothesis when an outcome has a low probability of occurrence (usually .05 or less) if, in fact, the null hypothesis is correct.

stratified random sampling  A sampling procedure in which the population is divided into strata followed by random sampling from each stratum.

systematic observation  Observations of one or more specific variables, usually made in a precisely defined setting.

systematic variance  Variability in a set of scores that is the result of the independent variable; statistically, the variability of each group mean from the grand mean of all subjects.

temporal precedence  Part of causal inference; the cause precedes the effect in a time sequence.

testing effect  A threat to internal validity in which taking a pretest changes behavior without any effect on the independent variable.
**test-retest reliability**  A reliability coefficient determined by the correlation between scores on a measure given at one time with scores on the same measure given at a later time.

**third variable**  When describing the relationship between two variables, a third variable is any other variable that is extraneous to the two variables of interest. True experiments control for the possible influence of third variables.

**true score**  An individual's actual score on a variable being measured, as opposed to the score the individual obtained on the measure itself.

**t-test**  A statistical significance test used to compare differences between means.

**Type I error**  An incorrect decision to reject the null hypothesis when it is true.

**Type II error**  An incorrect decision to accept the null hypothesis when it is false.

**validity**  See construct validity, external validity, internal validity.

**variability**  The amount of dispersion of scores about some central value.

**variable**  Any event, situation, behavior, or individual characteristic that varies—that is, has at least two values.

**variance**  A measure of the variability of scores about a mean; the mean of the sum of squared deviations of scores from the group mean.