



Visit Series website

Technical Terms for a Non-Technical Audience

Attrition

Occurs when members of the original study sample are not included in the final analysis sample. High attrition rates can cause the intervention and comparison groups to become dissimilar which, in turn, creates bias in study results.

Baseline equivalence

Evidence that the intervention and comparison groups are similar on important known characteristics, such as a pre-test, prior to beginning the study. Baseline equivalence ensures an apples-to-apples comparison, meaning that both groups are similar on the outcome of interest at the start of the study.

Clustering

A characteristic of the way participants are organized such that some participants (e.g., students) are nested within other participants (e.g., teachers or classrooms). Clustering can happen at multiple levels (e.g., students within teachers, teachers within schools, schools within districts).

Comparison or control group

A group of study participants (e.g., students, teachers, or schools) who are assigned to refrain from implementing a program, practice, or policy under study. Often the comparison group is asked to conduct “business as usual” so that the study can estimate what would have happened had the program not been implemented. The comparison group in a randomized control trial is typically referred to as the *control* group.

Confounding factor

A component of the study design that (1) is not part of the program or intervention being tested and (2) is present for all participants in one study group and absent for all the participants in the other study group. One common example is the N=1 confounding factor, which occurs when a single unit (e.g., a single participant or school) is assigned to one of two study groups. In this case, it is impossible to know whether differences in study outcomes are due to the intervention or unique features of the unit of study.

Effect size

A standardized quantitative measure of the effect of a program or intervention. An effect size converts the impact of an intervention—typically the difference between the intervention and comparison group on an outcome measure—into a standard metric. This is helpful when comparing across different outcomes and across different studies.

External validity

The extent to which it is possible to generalize from the context of the research study to other populations, times and settings.



Group design

A research design in which at least two groups of participants (e.g., students, teachers, or schools) are involved in a study and compared on outcomes at the end of the study.

Internal validity

The degree of certainty that A caused B, where A is the program and B is the outcome or result.¹

Intervention group

The group of study participants (e.g., students, teachers, or schools) who are assigned to implement a program, practice or policy under study.

Intervention

A program, practice, or policy that is introduced with the goal of achieving a desired result such as higher student performance. The intervention is commonly referred to as the *treatment*.

Multiple comparisons

Multiple statistical tests carried out to test for the effect of a program or intervention in a single study. Multiple comparisons can be problematic when a statistical model is not appropriately adjusted. In unadjusted models, some results may erroneously show statistical significance due to random chance.

Over-alignment

Occurs when an outcome measure is closely tailored to the intervention, giving the intervention group in a study an unfair advantage. For example, a list of vocabulary word definitions taught to students in the intervention group but not the comparison group, which is later administered as a test to use as an outcome measure, will likely be over-aligned with the intervention group.

Quasi-experimental design (QED)

A study design in which participants (e.g., students, teachers, or schools) are assigned to the intervention group or the control group using a non-random process. A QED typically uses existing demographic and pre-test data to select an equivalent group of participants to compare to the treatment group. Data are used to match participants based on known or observed characteristics.

Randomized control trial (RCT)

A study in which participants (e.g., students, teachers, or schools) are assigned to the intervention group or the control group using a random process.

Regression discontinuity design (RDD)

A study design that uses a continuous scoring rule to assign the program to study participants. Participants with scores below a pre-set cutoff value are assigned to the

¹ Definitions for internal and external validity adapted from Bickman, L., and Rog, D. (2009). *The Sage Handbook of Applied Social Research Methods*, Second Edition. Thousand Oaks, CA: Sage.



treatment group and participants with scores above the cutoff value are assigned to the comparison group, or vice versa.

Reliability

The degree to which a data collection instrument (e.g., a test, survey) results in similar findings for the same individual or group across repeated measurements.

Selection bias

Differences in study outcomes that can be attributed to dissimilar characteristics between the two study groups. When one group has characteristics (e.g., higher than average pre-test scores, ethnicity, age, poverty status) that do not exist in the other group, it is impossible to know whether differences in study outcomes are due to the intervention or the unique characteristics embedded in one of the two study groups.

Single case design (SCD)

A study design in which the study sample involves one individual case. A case may be an individual or a cluster of participants such as a classroom, school, or district. The case is measured several times before and after the program or intervention is introduced, and then results are compared during different phases of the study (i.e., before, during, and after the program).²

Validity

The degree to which a data collection instrument (e.g., a test, survey) or study measures what it intends to measure.

² Definition adapted from Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M. & Shadish, W. R. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.