



Definitions: Research Designs and Terminology

Types of Research Designs

Experimental research design	<p>Experimental research studies use randomized control trials (RCTs) to investigate the effects of a particular treatment or intervention on participants. In human research studies, RCTs are studies in which eligible and interested subjects volunteer for the intervention and are randomly assigned to a treatment group or a control group.</p> <ul style="list-style-type: none"> • In the treatment group, participants are offered the intervention or treatment being tested. Within the treatment group, some study subjects enroll, or “take up” the offered services, whereas other subjects may decide not to participate. When analyzing impact results, researchers must decide how to handle subjects who were assigned to the treatment group but did not participate. <ul style="list-style-type: none"> • Intent-To-Treat (ITT) analyses include all randomized study subjects and keep participants in the intervention group to which they were randomized, regardless of the treatment that they actually received (or did not receive).¹ The ITT approach conserves the integrity of the study’s random assignment. • Treatment-on-Treated (TOT) analyses are another approach, which analyze the impact of <i>actual</i> program participation, adjusting for no shows (members of the treatment group who did not participate) and crossovers (members of the control group who manage to receive the treatment).² • The TOT approach may be able to detect impacts that were hidden in an ITT analysis, but it is also more vulnerable to biases as the groups are no longer truly randomized. In response, statistical procedures have been developed to estimate TOT effects while simulating the experimental design as closely as possible.³ • In the control group, participants receive alternative treatment, which may consist of conventional treatment (e.g. usual services offered by the study site or related services available through outside entities) or no treatment at all (e.g. no similar services available through the study site or outside entities). <p>Some RCTs have multiple treatment or control groups, to test different aspects of the intervention in question. In a careful experimental study, all potential sources of bias are addressed.⁴ The key to experimental studies is random assignment.⁵</p>
Random assignment	<p>Random assignment is the process of dividing participants between control and treatment groups at random (by chance). This approach is used in experimental study designs to ensure that the treatment and control groups are alike and that each group is an unbiased representation of the target population (e.g. no self-selection). Randomly assigning a large number of individuals to one group or the other ensures</p>

diversitydatakids.org

data for a diverse and equitable future



		<p>that on average, the groups have similar characteristics and are comparable at the beginning of the trial. Therefore, in a trial using well-executed random assignment, researchers may infer that any differences that arise between the two groups are attributable to intervention or treatment impacts. Notably, this inference rests on the assumption that during the study period there are no other confounding events or factors that influence the treatment or control groups differently (apart from the intervention itself).</p>
<p>Non-experimental research designs</p>		<p>Research studies use non-experimental designs to observe and describe behaviors or characteristics of individuals or a population, or to associate a particular intervention with participant outcomes. In the latter case, due to the nature of non-experimental designs, it is not possible to definitively establish causality between treatment and outcomes even though a strong association may exist. Below are more details on non-experimental data.</p>
<p>Descriptive evaluation with no comparison group</p>		<p>In human intervention studies, descriptive data consist of information collected on the behaviors or characteristics of individuals or a population within a particular environment (such as a group participating in a program). This type of descriptive data is useful for understanding participant characteristics and needs. Examples of this type of data include observations, case studies, program administrative data, or survey data collected on program participants.⁶</p>
<p>Quasi-experimental evaluation with comparison groups</p>		<p>Quasi-experimental evaluations are non-randomized intervention studies that attempt to determine the impact of an intervention by comparing a treatment group to a comparison group without random assignment. Quasi-experimental evaluations may be used either because random assignment is considered unethical (e.g. if denial of the intervention is known to be harmful) or unfeasible (e.g. cost barriers, inability to randomize individuals or locations).⁷</p> <p>There are many types of quasi-experimental designs, depending on the data available (e.g. one group pre-test, post-test; two group post-test only, etc.). One common quasi-experimental design compares the outcomes of non-randomized treatment and comparison groups before and after an intervention (a two group pre-test post-test). However, because participants are not randomly assigned, those who participate in the intervention may be systematically different from those in the control group in both observed and unobserved ways. These systematic differences may influence the outcomes of the two groups such that any differences in key outcome variables cannot be attributed with certainty to the intervention in question.⁸ At a minimum, to address selection issues, the study should demonstrate that the treatment and comparison groups are similar on all observable measures at baseline.⁹</p> <p>To address the potential for bias in quasi-experimental studies, researchers may use</p>

diversitydatakids.org

data for a diverse and equitable future



different statistical methods to construct more valid comparison groups, either using propensity score matching, matched comparison groups, controls for pre-intervention characteristics, or other techniques. For example, in studies of Head Start (a federally funded early childhood education program for low-income children), some researchers have compared the outcomes of children who participated in Head Start to the outcomes of their siblings who did not participate in the program, thereby reducing selection bias.¹⁰ Even with these techniques, it is difficult to eliminate all potential for bias, and quasi-experimental research designs alone cannot lead to conclusive causal evidence on the impacts of an intervention.

Mixed methods research design

A mixed methods evaluation “involves the systematic integration of different kinds of data, usually drawn from different designs.” Mixed methods evaluations may use multiple designs such as randomized controlled trials and case-studies. These evaluations often use multiple types of data (quantitative and qualitative) and collect this data through various methods (including surveys, interviews, observations, and reviews of administrative data). Mixed methods evaluations are particularly useful when there are multiple evaluation questions that require different designs or data, or if researchers would like to answer the same question using different types of evidence, thereby increasing confidence in the evaluation findings.¹¹

Additional Terminology

Statistically significant

The “statistical significance” of an impact in a research study is the probability that the observed impact is occurring because of random chance. A “statistically significant” impact is one which researchers have calculated probably did not occur due to random chance alone. Many researchers consider an impact to be statistically significant if there is a five percent chance or less of it occurring due to random chance.¹² While this five percent level is a convention, some researchers may use thresholds that are more lenient (e.g. ten percent) or more stringent (e.g. one percent) to determine statistical significance.¹³ Researchers select thresholds for statistical significance based on several factors including their field of work, the measurement precision in the impact being considered, and whether it would cause more harm to declare an impact when there is none, or to declare no impact when there actually is one.

Sources & notes

- ¹ Higgins, J.P.T., Deeks, J.J., & Altman, D.G. (Eds.). (2008). Special topics in statistics. In J.P.T. Higgins & S. Green (Eds.). *Cochrane handbook for systematic reviews of interventions* (pp. 481-530). West Sussex: John Wiley & Sons Ltd.
- ² Ibid; Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., ..., Downer, J. (2012). *Third grade follow-up to the Head Start Impact Study final report*. (OPRE Report 2012-45). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. Retrieved from http://www.acf.hhs.gov/sites/default/files/opre/head_start_report.pdf.
- ³ For more information of calculating TOT effects, see Bloom, H.S. (2006). *The core analytics of randomized experiments for social research* (MDRC working papers on research methodology). New York, NY: MDRC. Retrieved from http://www.mdrc.org/sites/default/files/full_533.pdf. For additional discussion on crossovers, see U.S. Department of Health and Human Services, Administration for Children and Families. (2010). *Head Start Impact Study. Technical Report*. Washington, DC. Retrieved from

diversitydatakids.org

data for a diverse and equitable future



http://www.acf.hhs.gov/sites/default/files/opre/hs_impact_study_tech_rpt.pdf.

- ⁴ Higgins, J.P.T., Altman, D.G., & Sterne, J.A.C. (Eds.). (2011). Chapter 8: Assessing risk of bias in included studies. In J.P.T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions*, Version 5.1.0. The Cochrane Collaboration. Retrieved from http://handbook.cochrane.org/chapter_8/8_4_introduction_to_sources_of_bias_in_clinical_trials.htm.
- ⁵ Sibbald, B. & Roland, M. (1998). Understanding controlled trials: Why are randomised controlled trials important? *BMJ*, 316(201). doi: <http://dx.doi.org/10.1136/bmj.316.7126.201>. Retrieved from <http://www.bmj.com/content/316/7126/201>.
- ⁶ Clause, C. (n.d.). *Descriptive research design: Definition, example & types*. Education Portal. Retrieved from <http://education-portal.com/academy/lesson/descriptive-research-design-definition-example-types.html#lesson>.
- ⁷ Harris, A.D., Bradham, D.D., Baumgarten, M., Zuckerman, I.H., Fink, J.C., & Perencevich, E.N. (2004). The use and interpretation of quasi-experimental studies in infectious diseases. *Clinical Infectious Diseases*, 38, 1586–1591.
- ⁸ Harris, A.D., McGregor, J.C., Perencevich, E.N., Furuno, J.P., Zhu, J., Peterson, D.E., & Finkelstein, J. (2006). The use and interpretation of quasi-experimental studies in medical informatics. *Journal of the American Medical Informatics Association*, 13(1), 16-23. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1380192/>; Price, P. & Oswalk, K. (2008). *Quasi-experiments*. Department of Psychology, California State University, Fresno. Retrieved from <http://psych.csufresno.edu/psy144/Content/Design/Nonexperimental/quasi.html>.
- ⁹ Coalition for Evidence-Based Policy. (2014). Which comparison-group (“quasi-experimental”) study designs are most likely to produce valid estimates of a program’s impact?: A brief overview and sample review form. Retrieved from <http://coalition4evidence.org/wp-content/uploads/2014/01/Validity-of-comparison-group-designs-updated-January-2014.pdf>.
- ¹⁰ Currie, J. & Thomas, D. (1995). Does Head Start make a difference? *American Economic Review*, 85(3), 341-364; Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics*, 1(3), 111–134.
- ¹¹ USAID. (2013). *Technical note: Conducting mixed methods evaluation* (Monitoring and evaluation series). Retrieved from http://www.usaid.gov/sites/default/files/documents/1870/Mixed_Methods_Evaluations_Technical_Note.pdf.
- ¹² Effective Health Care Program. *Glossary of terms, Statistical significance*. U.S. Department of Health & Human Services. Retrieved from <http://effectivehealthcare.ahrq.gov/index.cfm/glossary-of-terms/?termid=67&pageaction=showterm>.
- ¹³ Schunemann, H.J., Oxman, A.D., Vist, G.E., Higgins, J.P.T., Deeks, J.J., Glasziou, P., & Guyatt, G.H. (2011). Chapter 8: Assessing risk of bias in included studies. In J.P.T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions*, Version 5.1.0. The Cochrane Collaboration. Retrieved from http://handbook.cochrane.org/chapter_12/12_4_2_p_values_and_statistical_significance.htm.