Evaluation Design: Strengths and Weaknesses[1]

Design (examples)	Strengths	Weaknesses
Experimental **Cluster Randomized Controlled Trial**	• Can be used when individual randomization is inappropriate or impossible • Can evaluate in real-world setting	• Requires recruiting more participants for detecting significant change (during data analysis) compared to individual RCTs • Requires complex data analysis methods
Description: An experimental cluster RCT design is when study participants are grouped (clustered) on the basis of some characteristic (e.g., city, zip code, age, travel mode) and then randomized as a group to either the intervention or control group.		
Example: The Advocacy for Pedestrian Safety Study was a mixed-methods, multi-center cluster randomized controlled trial that evaluated uptake of effective pedestrian safety interventions in high risk communities. This evaluation measured effects and factors related to the success or failure of the advocacy initiative aimed at improving pedestrian safety in high risk communities in the UK. The intervention was directed at local politicians who represented electoral wards housed within local authorities. Local authorities were randomized to intervention or control groups, stratified by study center area and local authority size.[2]		
Quasi-experimental **Pre-Post with Control Group**	• Can easily be used with surveys • Can account for the possibility that some other factor occurred at the same time as the intervention	• Can show only short-term changes • Requires a control group that's similar to the intervention group
Description: The quasi-experimental pre-post with control group design consists of both the intervention and control group receiving pre-test and post-test. The comparison group does not receive the intervention in between the two tests.		
Example: To evaluate a comprehensive pedestrian safety countermeasure program in Miami-Dade County, Florida, consisting of 16 education, enforcement, and engineering treatments deployed at four high-collision zones, pre- (1998-2001) and post-treatment (2002-2004) data were collected. Data was collected to compare local collision trend in neighboring counties for comparison.[2]		
Quasi-experimental **Pre-Post (without a control group)**	• Usually convenient and lower-cost • Is simple and quick to implement • Can easily be used with surveys • Can provide a reasonable estimate of the post-intervention change	• Can show only short-term changes • Cannot account for pre-existing trends • Cannot account for the possibility that some other factor occurred at the same time as the intervention (least rigorous in showing causal link between intervention and outcomes) • Is weak at ruling out other alternative explanations
Description: The quasi-experimental pre-post without a control group design consists of one intervention group that receives a pre-test followed by and intervention and then a post-test.		
Example: A quasi-experimental pre-post design was used to evaluate pedestrian and driver behavior at crosswalks in different types of locations and communities to determine whether effectiveness of Yield-to-Pedestrian Channelizing Devices was impacted by community type (urban, suburban, small city, college town); location type (intersection or mid-block); or intersection traffic control (all-way stop control, or partial stop control). Similar analyses were conducted for other sites in the same communities to examine the presence of spillover effects.[3]		
Quasi-experimental **Post-only**	• Is simple and straightforward • Can use with one group (no comparison group) or two groups (with a comparison group) of participants	• Can use when baseline equivalency is already established or assumed • Can be a low-cost way to assess differences between multiple versions of interventions
Description: The post-only design is one of the simplest designs. It consists of one or more groups completing a post-test after the intervention has been implemented.		
Example: To evaluate a new online version of an existing pedestrian safety education intervention, a quasi-experimental post-only design can be used to compare the effects of each version on mean knowledge scores from an online survey administered immediately after the intervention.		
Quasi-experimental	• Is easy to use with data routinely	• Is hard to use if special data collection methods, (e.g.,

[1] Adapted from: *University of Albany Center for Problem-Oriented Policing.* Assessing Responses to Problems: An Introductory Guide for Police Problem-Solver. *Appendix D: Summary of Evaluation Designs' Strengths and Weaknesses. Retrieved from:*

http://www.popcenter.org/tools/assessing_responses/7

[2] Lyons RA, Kendrick D, Towner EM, Coupland C, Hayes M, Christie N, et al. The advocacy for pedestrian safety study: cluster randomised trial evaluating a political advocacy approach to reduce pedestrian injuries in deprived communities. PLoS One. 2013;8(4):e60158. Available at: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0060158

[3] *Pennsylvania Department of Transportation Bureau of Planning and Research. C. Strong, M. Kumar. Western Transportation Institute College of Engineering Montana State University.* Safety Evaluation of Yield-to-Pedestrian Channelizing Devices: Final Report. *2006 Retrieved from:* https://www.dot.state.pa.us/public/pdf/YTPCDFinalReport.pdf

Interrupted Time Series	collected over many time periods • Can rule out pre-existing trends and alternative explanations	surveys), are used to measure the problem • Cannot account for the possibility that some other factor occurred at the same time as the Intervention • Takes a long time to establish results • Is hard to interpret when there are few events per time period before the intervention

Description: The simplest type of time series designs is the interrupted time series. This design is often used to evaluate the impact of a population-wide policy or intervention. It involves a single treatment group which is measured many times before and after the start of the intervention. It is called "interrupted" time series because the researcher graphs the data before and after the intervention, and looks for an interruption in the line or curve where the intervention was introduced.[4]

Example: A quasi-experimental interrupted time-series design was used to evaluate change in crash frequency across varying degrees of pedestrian countdown timer treatment density during 120 months of a large study in Detroit, Michigan. This evaluation design allowed evaluators to assess whether change in crash frequency depended upon of the degree to which the countdown timers penetrated the treatment unit. BE.[5]

Non-experimental Multiple Time Series	• Is easy to use with data routinely collected over many time periods • Can rule out preexisting trends and many other alternative explanations • Results easily presented (graphs) and interpreted (patterns) • Can forecast short-term trends	• Is difficult to use if data collection methods (e.g., surveys) change over time • Often needs qualitative data to explain • Takes a long time to establish results • Can be hard to interpret when there are few problem events per time-period before the intervention

Description: Multiple time series designs (or multiple baseline design) incorporates a baseline and an intervention condition across multiple participants, behaviors, or contexts. The greater the number of replications, the more confident one can be that the treatment produced the observed changes.

Example: A non-experimental multiple time series design can be used to evaluate trends in pedestrian injury rates in intervention areas before and after a community intervention. Data can be collected retrospectively for a five-year period, 2.5 years before intervention, and 2.5 years after the intervention, from state database records. Comparison of monthly pedestrian injury rates before and after the intervention can show any changes in injury trends.

Non-experimental Case Study	• Can provide a detailed description of problem or initiative, operations and contextual information, such as the historical perspective • Can be used to integrate data from multiple sources (e.g., documents, interviews, participation, observation, videos) • Can be used to report various characteristics (e.g., participation rates, representativeness, partnership influence, how resources are leveraged) • Can be both quantitative and qualitative data that details facilitators,	• Takes a long time • The amount of data can become unwieldy • Need skills in multiple data collection techniques • Insiders can be biased • Outsiders can be naïve

[4] *O'Connor, E., J. Bellamy, B. Spring.* Evidence-Based Behavioral Practice Online Training Course. *Critical Appraisal: Time Series Designs Retrieved from:* http://ebbp.org/course_outlines/critical_appraisal/#C

[5] *Huitema, RV Houten, H Manal.* Time-series intervention analysis of pedestrian countdown timer effects. *Accident Analysis & Prevention 2014.* 72:23–31. Retrieved from: http://europepmc.org/abstract/med/25003967

	barriers, and gaps	

Description: A non-experimental case study is an in-depth study of a particular person, group, program, policy, problem, decision, or organization over a period of time in order to provide a detailed description.

Application: Case studies are appropriate for assessing changes in public health capacity in sub-population groups. Case studies are especially applicable when the intervention is unique, when an existing intervention is used in a new setting, when a unique outcome is being assessed, or when an environment is unpredictable. Case studies can also allow for an exploration of community characteristics and how these may influence intervention implementation, as well as identifying barriers to and facilitators of change. Examples of case studies in pedestrian safety.[6]

---

[6] Compiled from Federal Highway Administration Office of Natural and Human Environment 2005. *Pedestrian and Bicycle Data Collection in United States Communities: Quantifying Use, Surveying Users, and Documenting Facility Extent*. Retrieved from: http://www.pedbikeinfo.org/pdf/PlanDesign_Tools_FHWACaseStudies.pdf