Experimental Studies

A. Alexander Beaujean, Ph.D.

Allison D. Gutierrez, Ed.S.

Baylor University

The concept of an *experiment* has undergone many changes in psychology's relatively short existence, and its essential features are still debated. Psychologists agree that it involves (a) systematically observing and recording phenomena; (b) explicitly describing the conditions under which the observations occurred, and (c) analyzing the resulting data to look for meaningful relations. The major area of disagreement is whether manipulation of phenomena is required.

Before the 1930s manipulation was desirable, but not necessary. Thus, comparing naturally-formed groups (e.g., adults vs. children) using a standardized psychological instrument was considered an experiment. In the 1930s, some influential psychologists began narrowing the definition so that it required manipulation. More specifically, it required manipulating the amount of a phenomenon (termed *independent variable*; IV) and then capturing some target phenomenon (termed *dependent variable*; DV) in order to determine if there was a functional (i.e., causal) relation between them. By the 1970s, this was the common view in psychology.

Early psychological experiments were largely of the kind now called *single-case* or *small-N*. A single participant is exposed to multiple IV conditions, one of which is usually a baseline condition that involves no manipulation. If there are meaningful differences in the DV across IV conditions, it means the IV is functionally related to the DV. More than one participant can be used in these designs, but each participant constitutes a self-contained experiment. The first participant is the original experiment and all others are replication experiments.

Francis Galton introduced the concept of comparing groups of participants who were alike on all important phenomena expect one. In this scenario, participants comprise samples from larger populations, and these samples are the primary focus —not individual participants. Since samples are the focus, data analysis requires examining aggregate trends via statistics instead of examining values for particular participants.

Galton's groups were naturally formed, so only allowed for descriptions. Causal inference required making group membership an IV, which meant researchers had to control membership into the groups. Moreover, they had to manipulate group membership in such a way as to rule out the possibility that other phenomena caused any differences observed in the DV. The cleanest way to do this is random assignment of participants to groups (i.e., IV's conditions). If the randomization process is properly carried out, then random assignment yields unbiased estimates of the IV on the DV in aggregate. Thus, randomization became the preferred method of controlling group membership.

There are two fundamental classes of designs used for experiments with groups: between-subjects and within-subjects. In *between-subjects* designs, participants are randomly assigned to one of the IV's conditions. In *within-subjects* designs, participants are exposed to multiple IV conditions at different points in time, with random assignment dictating the order of IV exposure for a particular participant. When there are more than two IV conditions, the number of unique ways to order presentation of IV conditions grows quickly. Thus, some systematic method for controlling order effects is typically implemented, such as counterbalancing or Latin square.

There are many variations of the between-subjects and within-subjects designs available. Repeated measures designs examine change in DV values across many data collection time points. Factorial designs examine both the independent and interactive effect of two or more IVs.

Mixed designs include aspects of both between- and within- subjects design, such as examining between-group differences in DV changes.

Further Reading

Appelbaum, M. I., & McCall, R. B. (1983). Design and analysis in developmental psychology. In P. H. Mussen (Ed.), *Handbook of child psychology* (4th ed., Vol. 1, pp. 415--476). New York, NY: Wiley.

Danziger, K. (1987). Statistical method and the historical development of research practice in American psychology. In L. Krüger, G. Gigerenzer & M. S. Morgan (Eds.), *The probabilistic revolution, Vol. 2: Ideas in sciences* (pp. 35--47). Cambridge, MA: The MIT Press.

Kirk, R. E. (2013). *Experimental design: Procedures for the behavioral sciences* (4th ed.). Thousand Oaks, CA: Sage.

Smith, J. D. (2012). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods, 17*, 510-550.

Winston, A. S. (1990). Robert Sessions Woodworth and the "Columbia Bible": How the psychological experiment was redefined. *The American Journal of Psychology, 103*, 391-401. doi:10.2307/1423217