Encyclopedia of Social Science Research Methods

Internal Validity

Contributors: Marilynn B. Brewer
Editors: Michael S. Lewis-Beck & Alan Bryman & Tim Futing Liao
Book Title: Encyclopedia of Social Science Research Methods
Chapter Title: "Internal Validity"
Pub. Date: 2004
Access Date: April 06, 2015
Publishing Company: SAGE Publications, Inc.
City: Thousand Oaks
Print ISBN: 9780761923633
Online ISBN: 9781412950589
DOI: http://dx.doi.org/10.4135/9781412950589.n436
Print pages: 503-505

http://dx.doi.org/10.4135/9781412950589.n436
Internal validity refers to the confidence with which researchers can make *causal* inferences from the results of a particular empirical study. In their influential paper on "Experimental and Quasi-Experimental Designs for Research," Campbell and Stanley (1963) characterized internal validity as the sine qua non of experimental research. This is because the very purpose of conducting an experiment is to test causal relationships between an INDEPENDENT VARIABLE and a DEPENDENT VARIABLE. Hence, it is appropriate that the primary criterion for evaluating the results of an experiment is whether valid causal conclusions can be drawn from its results. However, the concept of internal validity can also be applied to findings from correlational research anytime that causal inferences about the relationship between two variables are being drawn.

It is important to clarify that internal validity does not mean that a particular independent variable is the *only* cause of variation in the dependent variable—only that it has some independent causal role. For instance, variations among individuals in their attitudes toward a social program can be caused by many factors, including variations in personal history, self-interest, ideology, and so on. But if the mean attitude of a group of people that has been exposed to a particular persuasive message is more favorable than that of a **[p. 503 ↓]** group of people that did not receive the message, the question of interest is whether the message played a role in causing that difference. We can draw such a conclusion only if no other relevant causal factors were correlated with whether an individual received the message or not. In other words, there are no plausible alternative explanations for the covariation between attitude and receipt of the message. The existence of correlated rival factors is usually referred to as a CONFOUNDING of the experimental treatment because the potential effects of the variable under investigation cannot be separated from the effects of these other potential causal variables.

In this example, the most obvious challenge to causal inference would be any kind of self-selection to message conditions. If participants in the research were free to decide whether or not they would receive the persuasive message, then it is very likely that those who already had more favorable attitudes would end up in the message-receipt group. In that case, the difference in attitudes between the groups after the message was given may have been predetermined and had nothing to do with the message itself. This does not necessarily mean that the message did not have a causal influence, only

that we have a reasonable alternative that makes this conclusion uncertain. It is for this reason that RANDOM ASSIGNMENT is a criterial feature of good experimental design. Assigning participants randomly to different levels of the independent variable rules out self-selection as a threat to internal validity.

Basic Threats to Internal Validity

A research study loses internal validity when there is reason to believe that obtained differences in the dependent variable would have occurred even if exposure to the independent variable had not been manipulated. In addition to self-selection as a potential confounding factor, Campbell and Stanley (1963) described several other generic classes of possible threats to internal validity. These are factors that could be responsible for variation in the dependent variable, but they constitute threats to internal validity only if the research is conducted in such a way that variations in these extraneous factors become correlated with variation in the independent variable of interest. The types of potential confounds discussed by Campbell and Stanley (1963) included the following:

History. Differences in outcomes on the dependent variable measured at two different times may result from events other than the experimental variable that have occurred during the passage of time between measures. History is potentially problematic if the dependent variable is measured before and after exposure to the independent variable, where other intervening events could be a source of change.

Maturation. Another class of effects that may occur over the passage of time between measures on the dependent variable involves changes in the internal conditions of the participants in the study, such as growing older, becoming more tired, less interested, and so on. (These are termed *maturation effects* even though some representatives of this class, such as growing tired, are not typically thought of as being related to physical maturation.)

Testing. Participants' scores on a second administration of the dependent variable may be affected by the fact of their having been exposed to the measure previously. Thus,

testing itself could be a source of change in the dependent variable. This would also be a problem if some individuals in a study had received prior testing and others had not.

Instrumentation. Changes across time in dependent variable scores may be caused by changes in the nature of the measurement "instrument" (e.g., changes in attitudes of observers, increased sloppiness on the part of test scorers, etc.) rather than by changes in the participants being measured.

Statistical Regression. Unreliability, or error of measurement, will produce changes in scores on different measurement occasions, and these scores are subject to misinterpretation if participants are selected on the basis of extreme scores at their initial measurement session.

Experimental Mortality. If groups are being compared, any selection procedures or treatment differences that result in different proportions of participants dropping out of the experiment may account for any differences obtained between the groups in the final measurement.

Selection-History Interactions. If participants have been differentially selected for inclusion in comparison groups, these specially selected groups may experience differences in history, maturation, testing, and so on, which may produce differences in the final measurement on the dependent variable.

Again, it should be emphasized that the presence of any of these factors in a research study **[p. 504 ↓]** undermines internal validity if and only if it is differentially associated with variations in the independent variable of interest. Events other than the independent variable may occur that influence the outcomes on the dependent variable, but these will not reduce internal validity if they are not systematically correlated with the independent variable.

Marilynn B.Brewer

http://dx.doi.org/10.4135/9781412950589.n436
References

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching . In N. L. Gage (Ed.), Handbook of research on teaching (pp. 171–246) . Chicago: Rand McNally.

Crano, W. D., & Brewer, M. B. (2002). Principles and methods of social research (2nd ed.) . Mahwah, NJ: Lawrence Erlbaum.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). Experimental and quasi-experimental designs for generalized causal inference . Boston: Houghton Mifflin.