The SAGE Encyclopedia of Human Communication Sciences and Disorders

Experimental Research

Contributors: Vesna Mildner
Edited by: Jack S. Damico & Martin J. Ball
Book Title: The SAGE Encyclopedia of Human Communication Sciences and Disorders
Chapter Title: "Experimental Research"
Pub. Date: 2019
Access Date: May 8, 2019
Publishing Company: SAGE Publications, Inc.
City: Thousand Oaks,
Print ISBN: 9781483380834
Online ISBN: 9781483380810
DOI: http://dx.doi.org/10.4135/9781483380810.n242
Print pages: 728-732

Research methods are many and varied, and ideally, the choice is guided primarily by the nature of the problem to be studied. With respect to *participants*, research can be carried out on individuals (single-case studies) or groups. The latter may involve a single group or multiple groups (at least two: experimental and control). They can be further classified as between-subject/group or within-subject/group studies. With respect to *type of data*, there are quantitative (numerical data), qualitative (textual data), and mixed methods, which may be applied to nominal, ordinal, interval, ratio, or other type of scales. Number of *variables* will determine whether the study uses single-variable (one independent variable) or factorial (at least two independent variables) design. Based on the criterion of amount of *control* the design is categorized as nonexperimental (also referred to as pre-experimental), quasi-experimental, or true experimental. Nonexperimental studies allow the least amount of control because they deal with phenomena occurring in natural environments. Most of them are correlational: They try to determine correlations among relevant variables without manipulating them. With respect to phenomena studied, they can be classified into descriptive/exploratory survey studies and interrelationship/difference studies. Quasi-experimental design (most frequently used in applied research) allows more control over independent variable(s) and involves manipulation but lacks randomness in subject selection. (True) experimental design is characterized by manipulation of independent variable(s) and a high amount of control over subject selection and assignment.

The features discussed in this entry are incorporated into steps that are common to all types of research: choice of topic; survey of available literature; definition of research problem and formulation of hypothesis, including choice of variables; study design; subject selection; data collection and analysis; interpretation of results; and consideration of implications.

Features of Experimental Research

There are many definitions of experimental research, but they all share several obligatory features:

a. a dependent variable and at least one independent variable,
b. cause-and-effect relationship between these two variables while eliminating or controlling all other possible confounding factors,
c. an experimental group and a control group,
d. appropriate selection and assignment of subjects, and
e. replicability.

Dependent Variable and at Least One Independent Variable

The dependent (criterion) variable is the object of study and measurement. The independent (predictor) variable is selected and manipulated by the experimenter. The experiment reveals how the changes in the independent variable (IV) affect the dependent variable (DV). All other possible influences on DV, represented by variables referred to as extraneous/ confounding/spurious, need to be eliminated or controlled for. Factorial design involves two or more IVs and in that case the experimenter looks at main effects (how each of them affects DV) and interaction (how they affect it together). It is preferable to a series of single-variable studies. The variables that strengthen the relationship between two other variables are called moderator variables. Mediator (intervening) variables explain that relationship. They may be treatment- or subject-related. Control(led) variables are not independent. They are under the experimenter's control, but if their number increases or the experimenter fails to control them they turn into confounding variables and present serious problems for internal validity. There are slight nuances in definitions and descriptions of variables in literature, but rule of thumb is to eliminate all sources of uncontrollable influence on DV.

Cause-and-Effect Relationship

One of the first steps in experimental research is formulating a hypothesis about causality: The experimenter expects some effect (measured by DV) to be produced by a cause (contained in IV), eliminating all other sources of variation or influence (confounding variables) that might lead to rival hypotheses. A null hypothesis states that there is no significant difference between the studied groups (with respect to DV) that can be attributed to IV. Theoretically, this is the starting point, or the working hypothesis that the study sets out to refute. In practice, hypotheses are usually formulated as nondirectional (there is a difference between the groups) or directional (e.g., one of the groups will have a higher value); that is, the experimenter wants to test whether there is an effect that can be attributed to a cause (e.g., treatment) that can be expected on the basis of experience or available literature.

Experimental Group and Control Group

Experimental and control groups are comparable in all respects except in the application of treatment, and any difference in measurements that is found between them can be attributed to (and only to) treatment; it actually refers to any change in IV and can include active intervention, that is, addition to the stimulus or removal. In studies where both groups receive (different) treatment, they are called comparison groups.

Appropriate Selection and Assignment of Subjects

Subject selection and/or assignment to groups are generally regarded as the key feature that distinguishes true experimental designs from others. Ideally, subjects are randomly picked from the population and assigned to the experimental or control groups (probability sampling). Stratified sampling is a type of random sampling where the population is first divided into strata and then participants are randomly selected from each stratum. In a proportional stratified sample, the proportion of subjects from each stratum mirrors the proportion of the stratum in the population. If selection is not random it is called nonprobability sampling, because not all individuals have equal probability of being selected: Frequently, groups consist of subjects who happen to be at hand (convenience sampling) or those who have characteristics that the experimenter wants to study (purposive sampling). Even if selection itself is not random, the subjects may still be randomly assigned to different comparison groups. Another method of subject selection is matching: Subjects in the control group are chosen so as to match the subjects in the experimental group in all aspects related to DV. In most applied research, only nonprobabilistic sampling and matching are possible or appropriate.

Replicability

Replicability means that other researchers who carry out the experiment using the same methodology and comparable subject groups will get the same or similar results. To make this possible, methodology should be meticulously described when presenting the study. It is advisable that experimenters replicate their own work as part of strengthening external validity.

Reliability and Validity

In addition to the features listed above, every experimental design needs to be concerned with reliability and validity.

Reliability

Reliability refers to consistency. In research designs that involve evaluation by judges (raters), *interrater reliability* measures variation in scores across raters and *intrarater reliability* measures variation in evaluations of the same task or token by a single rater. *Test-retest reliability* primarily refers to variation in measurements of the same person on the same test on different occasions but under identical conditions. It is also related to intrarater reliability when it refers to variation in evaluation of the same task or token by the same rater at different times. *Intermethod reliability* is exhibited by good correlation between results obtained by different methods in the same subjects. *Parallel forms reliability* is an issue when test items are randomly assigned to (at least) two tests from a pool of items of similar difficulty. High correlations on these types of reliability are desirable because they indicate consistency of methods, instruments, tests, and sets of measurements. *Internal consistency* of a test is also manifested as reliability: A test is reliable if all of its items correlate with each other (0.60 – 0.90), that is, measure the same thing. However, correlations higher than 0.90 are indicative of items' redundancy.

Validity

Validity corresponds to truthfulness or appropriateness. Research that satisfies this requirement will yield results that are comparable to similar samples and will eventually lead to interpretations that can be applied to similar populations or situations. To achieve this, two sets of criteria that are grouped around *test validity* and *experimental validity* have to be considered. The former set includes questions about *construct validity* (does the test measure what it was designed to?), *content validity* (does the test measure the entire/representative sample of behavior that is the object of research?), and *criterion validity* (is the test related to the [predefined] criterion variable?). The latter set of criteria is concerned with the procedure itself and interpretation of results. Two major aspects are *internal validity* and *external validity*. As mentioned above, one of the salient features of experimental research is the existence of cause-and-effect relationship, which is covered by the notion of internal validity. If this causal relationship satisfies the requirements of temporal precedence, covariation, and non-spuriousness, internal validity is achieved: In an internally valid experiment, cause precedes effect, the two are related and that relationship (covariation) cannot be explained by any other hypothesis. In the process of designing and conducting research as well as during data analysis and interpretation, there are factors that may jeopardize internal validity (threats). They include the following:

- Attrition (mortality)—some participants fail to complete the study (needs to be accounted for in making inferences about effect of treatment; can be remedied by pretest/posttest);
- Ceiling/floor effect—participants reach maximum/minimum that the test allows (can be prevented by pilot study and improved test sensitivity and specificity);
- Compensatory rivalry / resentful demoralization—participants in study groups compete and reach performance levels they would not have in a natural setting or in the absence of the "rival" group / participants may become demoralized and don't see any point in doing their best (can be minimized by preventing contact among participants and awareness of group assignment; so-called blind design);
- Diffusion—treatment effects spread from one study group to the other (may lead to disappearance of between-group differences; can be minimized by blind design and preventing contact among participants);
- Experimenter bias / lack of implementation integrity / instrument change—experimenters behave inconsistently in different testing conditions or with different groups / treatment is not delivered as planned or its delivery is inconsistent / changes in any aspect of methodology during study (can be remedied by using double-blind design, perfecting instruments, training experimenters before testing, and avoiding changes once the experiment is under way);
- History—events that occur between pre- and posttest situations, unrelated to treatment, may affect

the outcome of the study (especially relevant in repeated-measures design and longitudinal studies; can be minimized by using control group);
- Maturation—participants undergo normal changes over time (developmental/maturational), which may affect the outcome (can be minimized by using control group and varying presentation conditions in within-subject design);
- Post hoc bias—the researcher is not objectively interpreting the data (can be minimized by peer review/comments);
- Regression to the mean (statistical regression) —participants who have performed extremely well or extremely poorly on the pretest tend to perform more like the average on the posttest (can be remedied by using control group and randomization);
- Selection bias—caused by inadequate subject selection (pretest differences between groups are unknown to or neglected by the researcher; can be minimized by using random subject sampling and pretest/posttest design);
- Selection-maturation interaction—participants in study groups develop/change differently over time (can be remedied by random subject selection); and
- Testing situation—possible practice effects (can be minimized by using control group and excluding subjects who frequently participate in tests).

External validity is what strengthens the purpose of research: An externally valid experiment will lead to conclusions that are transferable to similar contexts, thus satisfying the requirement of generalizability, that is, applicability to other individuals (population validity), circumstances (ecological validity), and times (temporal validity). To achieve this, the subjects must be representative of the population studied, the constructs adequate, and the procedure replicable. Most frequent threats to external validity include the following:

- Aptitude-treatment / selection-treatment interaction—selected participants are not representative (exhibit features that interact with IV in atypical ways);
- Interaction between treatment and testing—participants are affected by the awareness of being tested;
- Multiple treatment interference—participants are subjected to multiple treatments (simultaneously or at short intervals), which may lead to cumulative effects and/or interaction of treatments;
- Pretest and posttest effects—pretesting affects posttest scores (pretest sensitization);
- Reactivity—causal relationships would not occur in natural situations; and
- Situation—anything to do with the testing situation (e.g., time, place, distractions).

The best way to increase external validity is to repeat the experiment in slightly changed conditions and with different subjects.

Internal and external validity considerations are flanked by two other types that precede and follow the actual experiment. *Intentional validity* deals with the question of appropriateness of the chosen method, and is in that respect a variation of test validity. *Conclusion validity* may be defined as providing statistical evidence for the correlation between IV and DV based on the applied experimental design and methodology. Some of the threats to conclusion validity are statistical power, unsystematic variations in procedure, differences among subjects, unreliable DV, multiple comparisons, and violation of assumptions of statistical tests.

Ecological validity can be regarded as part of external validity (see above) or separately. If considered on its own, it refers to applicability of conclusions obtained in a laboratory setting to real-life situations. It may be jeopardized by high degree of control that is characteristic of studies with strong internal validity.

Obviously, it is impossible to achieve maximum strength of all types of validity at the same time. There are always trade-offs, particularly between external and internal validity. Priorities will depend somewhat on the

nature of the study and its expected practical implications. Internal validity commonly takes precedence, but in applied research external validity is very important. An experiment cannot be externally valid if it is not internally valid, but the opposite does not necessarily apply.

With the exception of internal validity, which is most closely related with experimental and quasi-experimental research, the rest of discussion on validity is pertinent for other types of research methods.

Types of Experimental Research

Two-group posttest-only design is the simplest experimental design, good for assessing causal relationships. It involves an experimental group and a control group formed by random selection or matching. Groups' means on DV are compared after treatment (applied to experimental group only). Used outside the laboratory it may be subject to internal validity threats caused by social interaction among the subjects. This design may be used when pretesting is impossible and the sample of subjects is large enough (> 30 per group). It may be conducted within groups.

Two-group pretest-posttest design involves two groups formed by random selection and assignment of subjects, pretest and posttest measurements in both groups, and treatment applied to experimental group only. It is highly controlled for threats to internal validity, but external validity may be jeopardized by the subjects' knowledge that they are being tested.

Solomon four-group design is a combination of pretest-posttest control group design and posttest-only group design, with good internal and external validity.

Experimental design that involves only one group of subjects is called *repeated measures design*: the same group of subjects is repeatedly measured for different DVs or for one DV at different times. It requires fewer subjects, eliminates the issue of nonequivalent groups, and is sensitive to finding statistical differences. Methodological concerns involve practice, fatigue, and carry-over effects. Order effects inherent in this type of study may be resolved by counterbalancing (e.g., Latin squares, randomized blocks) and increasing the time interval between treatments/measurements.

In a *combination of within- and between-subject pre- and posttest design*, all subjects are pretested and afterward randomly assigned to the experimental or control group. Posttest measures are taken as within-subject repeated measures design in the control group and group membership may be taken as a between-subject factor.

Ethical Considerations

There are two aspects of ethical issues to consider in experimental research. The first one is related to the experimental process itself and is often referred to as experimenter bias: An experimenter who is familiar with the hypotheses is more likely to conduct research in a way that will result in their confirmation. It is relatively easy to avoid by applying a double-blind design (both the experimenter and the participants are unaware of the hypotheses and/or to which group the subjects have been assigned). Further chance for the bias to jeopardize the objectivity of the study is interpretation of results, when the author may be tempted to report only data that support the hypotheses and ignore conflicting ones. This is also referred to as post hoc bias (threat to internal validity). There are no methodological safeguards against it—only researchers' integrity and rigorous review process by competent peers.

The second aspect is more complex and involves treatment of subjects and data obtained from them. The fundamental dilemma here is the cost-benefit ratio; that is, is the expected benefit (e.g., to science, humankind) greater than the possible harm inflicted on the subjects during research? Here the term *harm* is used in the broadest sense to cover a wide range of experimental situations in which subjects may find themselves, from mildly embarrassing and annoying to those where individuals are, due to random selection or group assignment, given less effective treatment or denied one altogether. Regardless of the level of possible harm, the subjects need to be informed about research and their role in it and give written consent. Such informed consent typically contains a general statement about the purpose of the study, description of what the participants will be expected to do, description of possible risks and their management, option to refuse to participate or withdraw at any time without penalization, information about protection of privacy and confidentiality, information about the contact person and the procedure of expressing concerns and asking questions, and information on availability of findings. Minors' participation must be approved by a parent or legal guardian. If deception was used, the subjects should be debriefed. Naturally, every research has its specificities that may be reflected in informed consent, but the above mentioned elements should be a minimum.

Advantages and Disadvantages of Experimental Research

Like any other method, experimental design has strengths and weaknesses. The most obvious advantages are control over IVs, straightforward determination of causal relationships, possibility of verifying results through repeatability/replicability, and the opportunity to create conditions that are not easily observed in natural settings or would take too long. Among the disadvantages are its unnaturalness (which may result in outcomes not otherwise found in natural situations or not to the same degree); difficulty to generalize, that is, apply the results to real-life situations (in part because the sample or situation may not be representative); and ethical considerations. Although good experimental design is a powerful tool, it cannot be applied to all types of research problems, and results may appear significant because of experimenter error or the inability to control for all extraneous variables. Finally, although experimental research reveals causality (i.e., there is an effect that can be attributed to a cause), it does not necessarily explain why this occurred.

Apart from the already mentioned steps that researchers can take to improve particular types of validity and take advantage of the experimental design strengths, it is advisable to use large enough groups, collect precise and detailed background subject data including appropriate standardized tests, and run pilot tests followed by introspection feedback from subjects. A training session (where appropriate) will help familiarize participants with the test, eliminate learning effects, and minimize initial differences among subjects. Where applicable, double-blind procedures should be preferred.

See also Ethics in Communication Disorders Research; Qualitative Research; Quantitative Research; Significance; Statistics: Descriptive; Statistics: Predictive

Vesna Mildner
http://dx.doi.org/10.4135/9781483380810.n242
10.4135/9781483380810.n242

Further Readings

Barnes, L, Hauser, J., Heikes, L., Hernandez, A. J., Richard, P. T., Ross, K., Yang, G. H., & Palmquist, M. (1994–2012). Experimental and quasi-experimental research [Writing guide]. Writing@CSU. Colorado State University. Retrieved January 20, 2017, from https://writing.colostate.edu/guides/guide.cfm?guideid=64
Blakstad, O. (2008, March 10). Research methodology. Retrieved October 3, 2017, from https://explorable.com/research-methodology
Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. Chicago, IL: Rand McNally College Publishing.

Creswell, J. W. (2009). Research design: Qualitative, quantitative, and mixed methods approaches (3rd ed.). Newbury Park, CA: SAGE.
Mildner, V. (2013). Experimental and quasi-experimental research in clinical linguistics and phonetics. In N. Müller & M. J. Ball (Eds.), Research methods in clinical linguistics and phonetics: A practical guide (pp. 28–47). Oxford, UK: Blackwell Publishing.
Trochim, W. M. K. (2006). Research methods knowledge base (
2nd ed
.). Retrieved from http://www.socialresearchmethods.net/kb/ (version current as of October 20, 2006).