A Beginner's Guide to Behavior Research Methods

Lauren S. Seifert.

Part I: Terms and Definitions

Data: Pieces of information that are collected in a systematic fashion.

Empirical (an adjective): a term applied to research that involves measurement of observable behaviors or effects. This adjective refers to those measurement techniques that attempt to uncover universal truths or principles by way of objective, unbiased observations.

Grounded Theory: Perhaps, a term dating to the 1960's, grounded theories are derived *inductively* from data. Refer to Strauss and Corbin (1991) for a description of the systematic way that meaning can be derived from data in order to build theory.

Human Research: According to the Dept. of Health & Human Services, USA, it is collection of data about a person or persons in an effort to draw generalizable conclusions. See 45 CFR 46 of the US Federal Code .

The Scientific Method: a general set of principles that is followed by traditional sciences (e.g., physics, biology, chemistry, sociology, psychology). The guiding ideals of the scientific method are that there are basic, underlying universal principles and that those principles/laws can be discovered via precise and unbiased measurement techniques. Such principles are objective truths. The Method listed below is a "long" version. For an abbreviated Method with comparisons to action research, see the Table that follows this list.

Dr. Seifert's Summary of the Scientific Method (in "steps"):

Step 1: Decide on a topic area.

Step 2: Define the problem through review of pertinent literature, and develop a research question.

Step 3: Define key variables and develop one or more predictions about them. These predictions are called "hypotheses."

Step 4: Design one or more research studies that will help you to answer your research question and address your predictions.

Step 5: Design specific research measures and pilot test them.

Step 6: Refine research measures based on the results of your pilot testing.

Step 7: Run your study. Collect data.

Step 8: Code, compile, and analyze your research data into results that can be interpreted.

Step 9: Interpret research results and draw conclusions.

Step 10: Revisit your research hypotheses, build new predictions, and develop additional studies to address old/new research questions.

Comparing Traditional Science and Action Research

The Scientific Method	Methods in Action Research
1a) Identify a problem area.	1) Coghlan & Bougie's (2010) "Pre-Step": Defining Context and Purpose.
1b) Define the problem –a problem statement via objective analysis of the existing literature.	2) "Constructing" the problem space through dialogic activity among stakeholders.
2) State hypotheses/predictions.	2) continue constructing the problem space; and 3) Begin a plan for action.
3) Design an empirical study: one that uses objective (third-person) measures.	3) continue planning for action. Plans for action include personal (first-person), collaborative (second-person), and consultative (third-person) views.
4) Run the study/Collect Data.	4) Act. Note that actions are monitored and observed for their effects. The investigator: pays attention, acts intelligently, uses good judgment, and is practical.
5) Code* and analyze the data. * To code typically means to render data into a form that can be analyzed with statistics. 6) Draw conclusions from data analyses which relate back to theories and hypotheses.	5) Evaluate the action and its impact.

What is "action research"? It is a set of research methods that provides an alternative to traditional, positivist science. It has gained some popularity in such disciplines as education, the arts, business/management, social work, counseling, and nursing. An excellent guide to action research is authored by Coghlan and Bougie (2010).

To "code" data (a verb): To convert pieces of information, which have been collected in a research study, to a form that can be more readily analyzed. Often, data coding involves transformation into quantities that can be analyzed with statistics.

Ethology: a somewhat antiquated term, it refers to the study of animal behavior.

Stimulus (plural, stimuli): simply, "a thing"; anything that has the potential to elicit a response.

Response: a behavior

Variable: any of a number of factors that might be manipulated, controlled, and measured in a research study. Two basic types of variables are "Independent Variables" and "Dependent Variables."

Subject: the person, animal, or object being observed by a researcher.

Participant: considered a synonym of "subject" by some researchers, this term is generally considered to be more "politically sensitive" and is used to acknowledge a subject's right to choose to take part or to refuse to take part in a study. "Participant" is much more often used as a term in survey research—when those who take part are actively participating in giving data. The term "subject" is much more likely to be used in studies of objects, animals, children, or of adults whose behaviors are being actively manipulated (as in a true experiment).

Respondent: another common term for "participant" that is commonly used in survey research.

Population: a universal set; all persons, animals, or objects that share a predetermined set of characteristics. Examples: "All polar bears on the earth" is the population of polar bears. "All students currently enrolled at Malone University" is the population of Malone University students. A list of all members of a population is called a “sampling frame”.

Sample: a selective set; all persons, animals, objects or stimuli that has been selected for observation. Examples: 200 Malone University students who are approached as they walk out of the dining hall, who are asked to participate, and who agree to take part in a survey about their career goals; 30 polar bears who are observed in the Arctic during a scientist's exhibition there.

Sampling (a verb): the acts of deciding who will be asked and then soliciting persons to take part in a study; if animals or objects are to be studied, then, the act of deciding which animals or objects will be scrutinized.

Select Types of Sampling:

Simple Random sampling: occurs when all members of a population are given an equal chance of being selected to take part in a research study. This is a “probabilistic” method of sampling.

Representative sampling: attempts to utilize the techniques of random sampling, while adding an algorithm that attempts to bring in participants from population sub-groups in equal proportions to their representation in the population. Example: In sampling residents of New York City, a researcher uses census data to determine various sub-groups, like African-American persons, Asian-American persons, Latino-American persons, non-American citizens, etc. Then, the researcher attempts to build a study sample that includes those sub-groups in the same ratios that they represent in the city's actual population. This is a “probabilistic” method of sampling.

Convenience sampling: occurs when a researcher is selective in sampling—selecting participants on the basis of their availability, proximity, and/or the ease by which they might be studied. This is a non-random method of sampling. This is a “non-probabilistic” technique.

Snowball sampling: a type of convenience sampling that identifies members of a population that one wishes to study. Once those members are identified, they are asked to help identify other potential research participants—typically, who are like them. Example: Finding participants for a survey about homelessness by visiting a shelter and asking homeless persons to tell their friends 'on the street' about the study. This is a “non-probabilistic” technique.

Sampling in "Waves": a basic term, this technique can be used with just about any sampling method. It consists of sampling at distinct points in time in order to increase the validity of one's sampling method. For example, a researcher who desires to collect data from a representative sample of Malone University undergraduates might begin by using a random number table applied to the student directory. S/he would call potential participants who had been randomly selected. This would be the "first wave." After conducting this first round of phone calls, the researcher would carefully assess which student sub-populations have been missed in the first wave (e.g., perhaps noticing that commuters had been reached in higher numbers than their actual proportion in the student population). A second wave of sampling could then attempt to correct the disproportion by 'target-sampling' non-commuters.

Select Types of Research:

Naturalistic observation: a "non-experimental" research study; Generally speaking, there is no manipulation of variables by the researcher. S/he merely watches, listens, & observes behaviors and records information about them. In one type of naturalistic observation, called "observation of public behavior" the researcher watches and/or listens to behaviors that occur in public. In “participant observation” the researcher takes part and documents his/her experiences and observations as an “insider”.

Case study: the behaviors of one person or of a small group are studied. Case studies can be non-experimental, quasi-experimental, or experimental—depending upon how the researcher manipulates (or does not manipulate) the research environment.

Survey: any of a number of techniques, including paper-&-pencil questionnaires, face-to-face interviews, phone interviews, and focus group interviews, online surveys, and computerized balloting. Surveys aim to collect information directly from persons by asking for their responses to questions or reactions to items. See "Selected advice for composing better surveys" and "Basic types of survey items" below.

Quasi-experiment (also, "pseudo-experiment"): like a true experiment, this type of research study utilizes naturally occurring variables, like age, weight, height, gender, etc., to categorize subjects. Then, the potential effects of naturally occurring variables on other behaviors/traits (like depression, spending behavior, reading aptitude, spirituality, general health) might be studied. Quasi-experiments include less overall control than do true experiments.

True experiment: a researcher controls and manipulates various aspects of the environment (called Independent Variables) in order to find out how it will affect a person's or an animal's behavior (the behavior that is influenced is called the "Dependent Variable").

Reliability: a general term, this refers to the likelihood that a specific result can be repeated or observed again.

Replication: a repeated study (often conducted in order to determine the reliability of research results).

Limited replication: a repeated study in which some conditions or variables are altered from the original study. Often, such replications attempt to repeat a study "with improvements."

Conceptual replication: the theory or model that underpins a study's design is used to devise another method for testing the phenomenon in question. In this type of replication, the methods and/or research design might be quite different than were those in the original study. Yet, the goal is to find further evidence for the phenomenon and—often—to find further support for the foundational theory or model.

Converging evidence, or converging operations: (after Garner, Hake, & Eriksen, 1956) this term refers to the attempts of researchers to determine underlying principles or truths about

something by using several different research methods or measures in order to *converge* on those truths.

Validity: often confused with "reliability" this is a different term that refers to the general ability of a research study to uncover truth. When applied to a specific measurement technique, its validity is its ability to measure what it is designed to measure.

Face Validity: whether a measurement technique appears to make sense, whether it appears to measure that which it is designed to measure. Counterexample: As a measure of intelligence, one's hair color appears to have *poor* face validity.

***
Construct validity: the degree to which something can be concluded about a higher-level or more abstract concept from an objective measure. For example, how much about a student's knowledge of calculus (a higher-level construct) can be concluded from his/her final exam score in a calculus class (with the exam performance being a direct, objective measure)?

Content and Criterion-based validity are often assessed during an examination of a test's construct validity. They are:

Content validity: whether the measure adequately covers the breadth of content that is to be tested. For example, if an employer wants to be sure that his/her accountants have adequate knowledge to accomplish all the tasks of their job, then the employer will want to be sure to assess all aspects of their accounting skills sufficiently to judge whether s/he can do the job.

Criterion-based validity: Also, called Criterion-related validity--a standard ("criterion") is established for measuring a phenomenon. Then, new measures are compared to the standard in order to verify their value as measures of the construct. For example, a new personality test might be compared to the NEO-PI in order to verify (or "validate") the new test's ability to measure the Big Five Factors of personality.
***

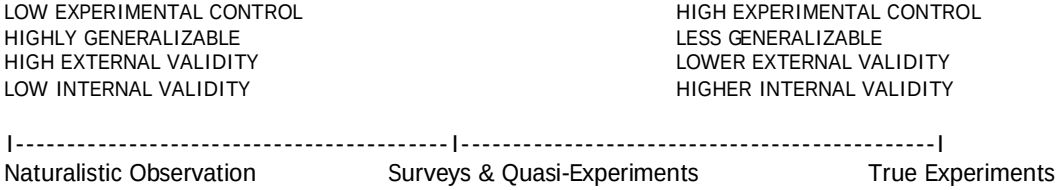
External Validity: the extent to which a conceptual model predicts "real-world" phenomena (see Cook & Campbell, 1979). Also used to refer to the extent to which the results of a research study reflect real-world phenomena.

Ecological Validity: thematically related to "external validity", it refers to the ways in which a research study and its results resemble real-world principles, environments, and occurrences (see Bronfenbrenner, 1977).

Internal Validity (related to "control"): the degree to which the results of a study are actually due to the researcher's manipulation of variable(s); the extent to which changes in the dependent variable(s) are owing to the manipulation of one or more independent variables.

Conceptualizing "Control" and "Generalizability" in Research:

A Graphic Representation of a "Continuum of Control"



Part II: Types of Research and Selected Terms

Observational Research (e.g., Naturalistic Observation)

Selected methods for sampling behaviors:

*By subject or case: An individual is identified and his/her behavior is documented.

*By interval: A time span is specified (e.g., 8:00AM – 10:00AM, daily) and only the behaviors occurring during that interval are documented.

*By duration: A specified behavior (e.g., grooming in albino Sprague-Dawley rats) is observed and its duration is documented (e.g., 10 sec grooming; 30 sec grooming; 40 sec grooming; etc.).

*By frequency: Generally, documenting with "hash" marks, the number of times a specific behavior occurs.

*By behavior category: The researcher documents all categories of behavior that are observed, e.g., grooming occurred, drinking occurred, mating occurred. Verbal/textual descriptions of the behavior categories are usually given in this type of observation.

Surveys

Selected types of survey research:

*By paper-and-pencil questionnaire: Respondents are asked to complete a paper version of the survey and return it to the researcher.

*By computer form or web-survey: Respondents are asked to complete a computer form that is on a dedicated computer (e.g., the researcher's laboratory computer) or that is on-line (e.g., a web survey through "Survey Monkey" that is accessible via the internet).

*By phone: Phone surveys usually involve a phone call from one person to another (researcher to respondent or vice versa) for the purpose of conducting a survey by voice, by phone.

*By PDA, pager, and/or "text messaging": More recently, some researchers conduct simple surveys (e.g., of patient compliance by paging the respondent and asking, "It's 10AM. Have U taken UR medication?) by page, text, or PDA. These allow researchers to obtain snippets of critical information about behaviors, *in vivo*, i.e., as they occur in everyday life.

*By face-to-face interview: An interviewer asks for responses to items/questions that s/he reads aloud or gestures (e.g., in ASL) to the respondent.

Basic types of survey items:

1) Open-ended: The participant can fashion the format of his/her response by writing, drawing, etc. Open-ended items often lead to richer data, but they are generally more difficult to "code." Coding open-ended responses into data that can be analyzed is often time consuming and difficult.

2) Closed-ended (or "fixed" or "multiple" choice): The survey item specifies response choices and, usually, asks the respondent to select one, rank the options, or select a fixed sub-set of options as his/her preferred choices.

Surveys (continued)

Examples of Open- and Closed-ended Survey Items about University Student Sleep:

1) Open-ended:

Please, describe your typical sleep pattern. How well do you sleep and how often? (Please, write in the space provided.)

2) Closed-ended:

*Using the following scale, how would you describe your nightly sleep?

1	2	3	4	5
very irregular	irregular	somewhat regular	regular	very regular

*Of the following, which one BEST describes the quality of your nightly sleep pattern? (Please, circle one response.)

A. very satisfying
B. satisfying
C. somewhat satisfying
D. dissatisfying
E. very dissatisfying

*Of the following, which one BEST describes the amount of your nightly sleep? (Please, circle one response.)

A. Far too little
B. Too little
C. Enough
D. Too much
E. Far to much

Selected advice for composing better surveys:

1) Use simple language. To assume no better than a 6th-grade reading level will provide comprehension of survey items by a broad, adult audience.

2) Ask about one issue per item. Don't combine issues within a single question or item. This can confuse respondents and muddy the data.

3) Use understandable response options and be sure that options are mutually exclusive, yet exhaustive.

Surveys (continued)

Selected advice for composing better surveys (continued):

4) Pilot test your draft survey with persons who are like the persons you will be surveying. This consists of giving the draft survey to a few people who can help identify problems with wording, clarity, length, duration, and subject matter.

5) Consider the ways you will code data. For example, in the third sample of a closed-ended survey item (previous page), it might have been more helpful to code the response options numerically (e.g., "1" = far to little sleep, to "5" = far to much sleep). That numeric response scale would lend itself more easily to a statistical analysis of the respondents' perceived amount of nightly sleep.

6) Be wary of sexist, racist, or otherwise prejudicial language. If you would like to describe a behavior vignette or scenario, be sure your language is inclusive, so that it does not alienate or ostracize research participants. [For example: Notice the use of "his/her" in the instructions to subjects, below. This is appropriately, non-sexist language in the instructions of an open-ended, vignette-type survey item.]

EXAMPLE: Please, read the following story and decide whether the person involved was justified in his/her behavior.

"Nino has studied all night for the big statistics exam. Upon reaching the classroom, he noticed that it was empty. Marta, another student in the class, stood nearby the door and asked, "Where is everyone?" It was 5-minutes before the class hour and Nino and Marta were the only persons at the classroom. As time ticked by, nobody else arrived. Finally, Marta said, "I'm going to call the professor's office, because something is just not right!" She called the office and while she spoke to someone on the other end of the phone, Nino said, "Oh, I'm just gonna go. We're obviously not having class or the test today." And Nino walked away.

Given that Nino and Marta each expected to come to the classroom and take an exam on that day and at that time, how would you explain their different behaviors?

Quasi-Experiments

Quasi- or pseudo-Experiments are like true experiments, but they are not authentically experimental in nature. Thus, the terminology is "pseudo" or "quasi"—to resemble a thing, yet *not* to be that thing. Quasi- or pseudo-experiments seem like true experiments, but the former lack true control of independent variables. Thus, one might utilize designs that are *like* the designs of true experiments. However, researchers are unable to manipulate the IV's of a quasi-experiment in the ways that they can manipulate them in true experiments.

Independent variable (Abbreviated IV): things the researcher manipulates in a study in order to produce change. In quasi-experiments, IV's cannot be manipulated. They are "pseudo" IV's (like a subject's age or gender).

Dependent variable (Abbreviated DV): those things that are affected by independent variables. A researcher measures them in order to find out whether they are influenced by the IV.

Quasi-Experiments (continued)

An example:

In a study about memory and Alzheimer's disease, a researcher might utilize a between-subjects design in order to assess the influence of extra testing on participant performance. Performance at the second (or "real") test is compared across the two groups in order to assess the effects of a practice test.

GROUP	held constant	DV
with probable Alzheimer's disease	study, then practice test	second, "real" test
without Alzheimer's disease	study, then practice test	second, "real" test

Notice that the experimenter can control study and test, and that s/he holds that constant at one level. That is, both groups get study time, followed by a practice test. Here, the IV of interest is "group" (with Alzheimer's versus without Alzheimer's). However, the experimenter has no control over who is assigned to groups. Persons come into the study with the disease or without it. Thus, the researcher has no true control over the IV of interest. It is a "quasi" manipulation, using subject characteristics to create the between-groups comparison.

True Experiments

Independent variable (Abbreviated IV): things the researcher manipulates in a study in order to produce change.

Dependent variable (Abbreviated DV): those things that are affected by independent variables. A researcher measures them in order to find out whether they are influenced by the IV.

Hypothesis: an informed prediction about what will happen in a research study. Example: "Caffeine will increase anxiety among students who take caffeine."

Null hypothesis: a prediction stated as "null." In science, we assume that we can NEVER prove something with 100% certainty. Instead, we are humble. We always assume that there is error in our methods. Because of this error, we state our predictions as if they will not be proven. Example: If I believe that caffeine will induce anxiety (as above), then I state the hypothesis as null, as follows, that "Caffeine will not change the anxiety levels of students who ingest it."

Statistical rejection of the null hypothesis: In the analysis of data from my research study, I will desire to reject the null hypothesis and show that my experiment supports the idea that caffeine does affect students' anxiety levels.

True Experiments (continued)

The caffeine study as a true experiment:

Randomly select 40 students from among those who attend Malone University. Then, randomly assign the students to the "No Caffeine" and "Caffeine" groups.

The research design:

A simple between-groups design with one independent variable

(NOTE: the IV has 2 "levels" of caffeine intake: no caffeine versus some caffeine. Each group receives just one level of the IV.)

Group	IV	DV
Control Group	water w/o caffeine	anxiety questionnaire
Experimental Group	water with caffeine	(same) anxiety questionnaire

Another research design:

A simple within-subjects design with one independent variable

[NOTE: the IV has 2 "levels" of caffeine intake: no caffeine versus some caffeine. Unlike the Between-Subjects Design (above) all subjects in the Within-Subjects Experiment (below) receive all levels of the IV]

TIME	IV	DV
1 (all subjects at Time 1)	"Baseline" = no caffeine	anxiety questionnaire
2 (all subjects at Time 2)	"Experimental" = caffeine	(same) anxiety questionnaire

Artifact: a global term for an error in research measurement, method, computation, or technique. Artifacts can originate from numerous sources, such as the ones listed below.

Some sources of ERROR in research:

Attrition: loss of subjects from a study during the interval over which they are studied. In longitudinal studies, participants might be lost in-between test sessions. In cross-sectional or single-time experiments, participants might be lost during the test session. Sources of subject loss can include loss due to: Subject election to leave/not to continue; subject illness or death; subject inability to return to the test site; etc. Loss of subjects is generally non-random. Thus, it

can contribute to errors in the results & conclusions from a study. Loss of subjects due to death is a type of attrition called “mortality”.

Experimenter bias: the researcher's own imperfection leads him/her to perceive, record, and/or analyze data in a manner that introduces error. Example: reading a scale incorrectly.

Participant bias: a subject's own imperfection leads him/her to answer a question or perceive a situation in ways that lead to mistaken behavior. Example: a participant in a phone survey doesn't understand a question in a survey and answers in a way that does not represent his/her true opinion.

Effects of history: life events occur that change a study's results in ways that do not represent truth. Example: US residents born in 1925 appear to have greater intellectual strengths in math than do those born in 1945. The result is more likely due to the greater emphasis on math skills in grade-school from 1930-1935 than from 1950-1955. The result is probably NOT due to a difference in innate intellectual abilities. It is a result "created" by history.

Sampling bias: something has gone wrong in a researcher's sampling method and it has created the research result.

Demand characteristics: features of a study that cue participants about how to act or about how the researcher "wants" them to act.

Confounds: uncontrolled variables that affect the dependent variable, thereby making it seem as if the independent variable has produced the effect. Indeed, careful study of and control for potential confounds is critical for sound assessment of an IV's effect on the DV.

Hawthorne Effect: historically, thought of as the tendencies of research subjects to behave differently, merely because they are being observed.

Experimental Error

In experimental research, a null hypothesis is set up as a prediction to be refuted by statistical analyses of the data. Because human researchers, human research participants, and the experiment process are all imperfect, error can enter research and lead to erroneous conclusions. The following diagram reflects ways in which error can lead to false conclusions about the real-world (as in Type I and Type II errors).

	Conditions in the Real World ("The Truth")	
Experimental Results	NULL IS TRUE	NULL IS FALSE
Null Result (Researcher's Conclusion: The null hypothesis is accepted, and the researcher concludes that the IV does not affect the DV.)	Correct acceptance of the null hypothesis. The IV manipulation does not cause a change in the DV.	INCORRECT acceptance of the null hypothesis. This is called a TYPE II ERROR.
Null rejection (Researcher's Conclusion: The null hypothesis is rejected, and the researcher concludes that the IV does affect the DV.)	INCORRECT rejection of the null hypothesis. This is called a TYPE I ERROR.	Correct rejection of the null hypothesis. The IV manipulation does cause a change in the DV.

Recommended Reading

American Psychological Association. (1982). *Ethical principles in the conduct of research with human participants.* Washington, DC: Author.

American Psychological Association. (2009). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

Baron, R.M., & Kenny, D.A. (1986). The moderator-to-mediator variable relationship in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173-1182.

Bordens, K.S., & Abbott, B.B. (2005). *Research design and methods: A process Approach* (6th ed.). New York, NY: McGraw-Hill.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32,* 513-531.

Campbell, D.T., & Stanley, J.C. (1963). *Experimental and quasi-experimental designs for research.* Boston, MA: Houghton Mifflin.

Christensen, L.B. (2007). *Experimental methodology* (10th ed.). Boston, MA: Pearson.

Coghlan, D., & Brannick, T. (2010). *Doing action research in your own organization* (3rd ed.). Los Angeles, CA: Sage.

Cook, T.D., & Campbell, D.T. (1979). *Quasi-experimentation: Design & analysis issues for field settings.* Boston, MA: Houghton Mifflin.

Garner, W.R., Hake, H.W., & Eriksen, C.W. (1956). Operationism and the concept of perception. *Psychological Review, 63,* 149-159.

Grimm, L.G., & Yarnold, P.R. (Eds.). (1995). *Reading and understanding multivariate statistics.* Washington, DC: American Psychological Association.

Kuhn, T.S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago, IL: University of Chicago Press.

Lynd-Stevenson, R.M. (2007). Concerns regarding the traditional paradigm for causal research: The unified paradigm and causal research in scientific psychology. *Review of General*

Psychology, 11, 286-304.

Popper, K.R. (1968). *The logic of scientific discovery*. London, UK: Hutchinson.

Sekaran, U., & Bougie, R. (2010). *Research methods for business: A skill building approach* (5th ed.). West Sussex, UK: Wiley.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory and procedures.* London, UK: Sage.