Step 8: Considering Validity and Discussing Limitations
Written and Compiled by Amanda J. Rockinson-Szapkiw and Anita Knight

Introduction

It is important to think about threats to validity prior to planning all of the details of your study so that you can be proactive in designing a study that controls for threats to validity. Then, in the limitation section of your research plan, identify your threats to both internal and external validity and discuss controls. Begin by thinking about your student and identifying a list of threats to internal and external validity. Think about weaknesses of the study, the design (e.g., lack of randomization, bias), the analysis, the instruments, and the sample (e.g. gender, age, ethnicity, geographical location). Once you identify the limitations, label them. In your manuscript.

- Identify the type of threat and describe it in terms of your present study.
- Discuss how the limitation could potentially influence the study.
- Discuss the steps taken to limit the threat, if applicable.

For example, if you were studying college students and non-ignorable nonresponse was an issue, you may say (Szapkiw, 2009, p.237):

"Results do not account for students at universities who chose not to participate due to multi-institutional research polices that prohibited participation or universities or who chose not to participate for other reasons. Since this study used survey data, responses made by students who did not respond to the survey or who dropped out of the online courses were not accounted for. This subjected the study to unit nonresponse and the issue of non-ignorable nonresponse. Within the realm of non-ignorable nonresponse issues, item nonresponse was not a problem in this study; however, the problem of unit nonresponse needs to be noted as a limitation when applying and making inferences based on part one of this study (King, Honaker, Joseph, & Shever, 1998). Since the data analysis did not use statistical controls to address the issue of non-ignorable nonresponse, findings cannot be applied to the students who did not respond. Thus, care should be taken not to make invalid inferences based on the results (Hausman & Wise, 1979)."

Topical Discussion

Define Validity

Validity, in reference to research, refers to the accuracy of the research results. There are 4 types of validity (Cook & Campbell, 1979; Kazdin, 2003):
1. **Internal validity**
2. **External validity**

3. **Construct validity**
4. **Statistical conclusion validity**

Internal Validity

Internal validity has been defined as "the confidence one can have in inferring a causal relationship among variables while simultaneously eliminating rival hypotheses" (Hepner, Kivilghan, & Wampold, 1999). It refers to the extent that a researcher has control over extraneous variables and can say that the intervention accounted for the results. Threats to internal validity can be thought of as rival hypotheses or alternative explanations for the effect within a sample.

To evaluate whether a study has internal validity, and the degree of the internal validity the study possesses the researcher must ask the following questions, "Is the change in the independent variable responsible for the differences observed in the dependent variable?" And, "To what degree can alternative explanations of changes in the dependent variable be ruled out?" As a researcher asks this question, the researcher is evaluating the internal validity of the study at hand. However, since this question may not be answered directly, researchers can conclusively respond to the question, by ruling out alternative explanations for the changes in the dependent variable (Hepner, Kivilghan, & Wampold, 1999).

And, threats to internal validity, these confounding or extraneous variables, need to be controlled. Several procedures may be undertaken to control for internal validity. These include the following (Hepner, Kivilghan, & Wampold): a) randomly assigning participants to groups, b) randomly selecting participants, c) determination of measurement times, and d) manipulation of the independent variable. Let's take a look at specific threats to internal validity and how they can each be controlled for either statistically or by design.

Table 1 identifies threats to internal validity. Each threat is labeled, defined, and illustrated. Possible actions a researcher can take to control for the threats are also identified.

Table 1. Threats to Internal Validity (Cook & Campbell, 1979; Kazdin, 2003)

Type of Threat	Definition	Illustration	Possible Control
History	An event outside of the experiment that affects the dependent variable	During the course of a diversity tolerance treatment aimed at improving tolerance of undergraduate students, the university has a multicultural awareness week with numerous events. This event hosted by the university may affect the tolerance attitudes of the participants, especially if a survey is used as surveys are particularly subject to this influence	Reduction of time between the pretest & posttest Control Group
Maturation	Change or growth over time (e.g. older, stronger,	In a treatment for depression, the process of "spontaneous remission" is wrongly attributed to a treatment effect.	Reduction of time between the pretest &

			posttest Control Group Randomizati on/ True experimental design
Testing	Presence of the pretest or posttest (e.g. familiarity with the test may cause improvement)	A group of adolescents take the Beck Depression Inventory (BDI) before and after treatment. Familiarity with the instrument in the post testing influences performance eon the instrument.	Parallel test form True experimental design to eliminate pretest Solomon 4-group design
Instrumentation	Changes in the instrument or the environment in which the measurement is being given	A graduate student in a counseling program is interested in exploring the similarities and differences in matters of motivation in parenting between mothers and fathers. She plans to collect data from 150 parents taking a class using an instrument she developed based on her experience. She plans to review the parents' responses to open-ended questions and to detect any generalities or trends without any system for data analysis.	Standardized testing manual; trained implementers and observers ; inter-rater reliability Valid and reliable instrument Objective observation instruments Minimized experimenter interaction with the participants
Statistical regression	The tendency for extreme score to regress toward the mean	A therapeutic intervention based in Beck's work was developed to decrease depression. If the intervention is implemented with 25% of the most depressed clients (determined by scores on the BDI) receiving services at a treatment center, The BDI scores will be assured improvement from pre to posttest by statistical regression alone.	Control group Random Assignment

Selection bias	Differences between groups prior to implementation of the experiment. This is often present in research that uses intact groups or self-selection to groups.	Two classes are exposed to different counseling skills teaching methods and both classes post tested on their skills. If the class A scores better than class B, the difference could be attributed to the fact that class A had better counseling skills prior to the teaching intervention.	Random Assignment Homogeneity of groups Statistical control (e.g. pretest, covariates) Matching
Attrition/ Mortality	Loss of participants	Clients who sign up to participate in a study drop out before the completion of the study.	Shorten duration of research Compare dropouts & non dropouts
Diffusion of treatment	The treatment is accidentally dispersed to both groups. This may occur for many reasons including the control and experimental group becoming aware of both treatments .	A teacher who is teaching the control group is trained in the methods the experimental group is receiving and integrates the method with the control group by nature of her training.	Minimize or Eliminate) contact between groups Use systematic procedures for interventions
Reaction to controls	Due to awareness about participation in the study, participants behave different. This can include the compensatory equalization of treatment, in which one treatment is seen as more desirable than the other by the implementer, and, thus, the implementation is biased. This also	Two classes are exposed to different counseling skills teaching methods and both classes post tested on their skills. Class B the comparison group sees Class A receiving "special treatment" and feels resentful. This resentment affects their performance on the post test.	Minimize or Eliminate) contact between groups Use systematic procedures for interventions

includes resentful demoralization of the control group, in which the control group perceives less benefit and this perception influences the outcomes.	

*These threats are not mutually exclusive. Many overlap and are considered in combination with one another. Additionally, it is important to note that these threats only pertain to experimental studies and not correlational studies. Correlational studies, particularly predictive correlation studies have a unique set of threats to validity. Also, this is not a comprehensive list.

> **Researcher application:** A researcher should strive to eliminate and minimize as many threats to internal validity as possible, while still recognizing that it is not possible to minimize all threats. In minimizing and seeking high internal validity, the researcher ensures highly creditable research. When treats cannot be fully controlled or eliminated, they need to be listed as limitations in the research report.

Internal validity is concerned with what is, and external validity is concerned with how the findings can be applied. As Mook (1983) purports meaningful research begins with first understanding the phenomenon, which may require little attention to external validity. For often as internal validity increase, external validity decreases. Internal validity is of greater concern than external validity, for without internal validity generalization is meaningless.

External Validity

External validity refers to the extent in which results can be generalized. It answers the question, To what extent (i.e. population, setting, etc) can the results be applied? There are a variety of ways that a researcher can increase external validity. These include, but are not limited to the following: (a) achieve representation of the population through strategies such as random selection, (b) use heterogeneous groups, (c) use nonreactive measures , and (d) use precise description to allow for study replication or replicate study across different populations, settings, etc.

Table 2 outlines threats to external validity.

Table 2. Threats to External Validity. This figure has been developed and modified from primarily from Rovai, Baker & Ponton, 2013 with additional information from: Cook & Campbell, 1979; Kazdin, 2003; Bracht & Glass, 2013).

Type of Threat	Definition	Illustration	Possible Control
.			

Population Validity	The extent to which the sample used was representative of the target population	If the population of a study is all College Students at a Southern University, then researchers could plan to survey a specified number of freshmen, sophomores, juniors, seniors, and graduate students so the sample is similar to the target population.	Identify target populations for generalizing & plan so that the sample clearly exemplifies the target population. Employ random selection to ensure selecting participants from the target group.
Reaction Effects of Experimental Arrangements (Setting)	The novelty of a treatment or the setting in which it takes place may impact the effects of the experimental treatment.	If participants are assigned a playlist of music to listen to as a method for regulation affect, and one group listens in comfortable chairs in a lounge and the other in a classroom the experimental conditions may impact the effects of the experiment.	Experimenter settings approximates population setting Use a naturalistic setting with non-invasive measures Extend the length of the treatment so that novelty can wear off
Reactive Effects of Testing	The pre-test may expose the participants to relevant information influencing them/becoming a component of the treatment. This may lead to less sensitization on the post-test.	Counseling students' emotional intelligence is surveyed before students participate in counseling skills training and then again after training. Students may learn information regarding emotional intelligence on the pre-test that may impact performance on the post-test.	Select a pre-test that does not alert participants to treatment targets. Consider eliminating the pre-test.
Experimenter Effect (Rosenthal Effect)	Researcher unintentionally changes participant behaviors with either verbal or non-verbal cues.	A counselor believes that the CBT treatment he provides is more effective than the EMDR treatment he provides and so his attitude impacts his	Implement a double blind study.\n\nBuild in a replication study that is simultaneous.

		delivery method.	
Interaction of History and Treatment Effects	Depending on the events going on in history at the time of the treatment- the current events may impact the efficacy of the treatment.	A counselor educator studying the impact of CBT on mild depression in NYC in November of 2011 may have different results due to the 9-11 historical events.	Build in replication studies that occur simultaneously Avoid instances where history and treatment interact.
Measurement of the Dependent Variable	All assessment information must be articulated clearly for the purposes of replication.	Instruments should be identified clearly with citations and psychometric properties. This information is necessary for others to replicate the study.	Maintain copies of assessment forms Maintain copies of instructions for coding and decoding.
Interaction of Time of Measurement and Treatment Effects	The time of day the treatment is provided may interact with the treatment effectiveness	A therapist is working with a personal trainer to refer clients diagnosed with mild depression for exercise therapy. The clients are asked to report to the gym at 5:30am each day. This time of day may impact the treatment myself.	Maintain consistent conditions for the study sample and the target population sample

*These threats are not mutually exclusive. Many overlap and are considered in combination with one another. Also, this is not a comprehensive list.

Construct and Statistical Conclusion Validity

In addition to internal and external validity, a research must also be concerned with construct and statistical conclusion validity:

Construct validity refers to the degree to which inferences can be made from operationalizations to the constructs upon which the operationalizations were made. Kazdin (2003) states that construct validity asks the following question, "Given the intervention was responsible for change, what specific aspect of the intervention or arrangement was the causal agent, that is, what is the conceptual basis underlying the effect?" (p. 23)

Statistical conclusion validity refers to how well an effect can be detected. According to Cook and Campbell (1979), "statistical conclusion validity refers to inferences about whether it is

reasonable to presume covariation given a specified alpha level and the obtained variances" (p. 41).

The Case of Charlie
The selection threat to validity is a major concern for Charlie.

- ▶ Since the randomization is not possible Charlie had the opportunity to adopt three types of control procedures, usually adopted to help achieve equality of groups and limit the selection threat.

    - ▶ **Matching**: If Charlie matches on gender, then a participant in one group who is a female would be matched with a female in the other group.

    - ▶ **Comparing homogeneous groups or subgroups:** If gender was an identified as a extraneous variable, Charlie might limit groups to contain only female participants. Or, each group might be divided into male and female subgroups.

    - ▶ **Analysis of covariance (ANCOVA):** A statistical method can be used to equate groups on one or more variables. If Charlie wanted to statistically equate the groups on gender, gender would be the covariate in the ANCOVA procedure.

Application: Developing My Research Plan
Answer the following questions:

- What are your threats to validity ? Think about your study and identifying a list of threats to internal and external validity. Think about weaknesses of the study, the design (e.g., lack of randomization, bias), the analysis, the instruments, and the sample (e.g. gender, age, ethnicity, geographical location).

_____
_____
_____
_____
_____
_____

- Identify the type of threat and describe it in terms of your present study.

_____
_____
_____
_____
_____
_____

- Discuss how the limitation could potentially influence the study.

_____

_____

_____

_____

_____

_____

- Discuss the steps taken to limit the threat, if applicable.

_____

_____

_____

_____

_____

_____