Recitation 6

Oct 12, 2021
Rachel Lee

Internal Validity

- **Selection bias -** When participants self-select (or are otherwise non-randomly allocated) into treatment conditions, differences across groups on the outcome cannot be unambiguously attributed to the treatment conditions.
- **Attrition -** Even in a randomized study, bias can result if participants drop out at different rates from treatment arms, and, especially if drop out is related to other factors.
- **Testing -** Measuring/testing may have an effect on the outcome that is independent of any effect of the manipulated treatment condition.
- **Regression (to the mean) -** Over time, those with extreme scores tend to return toward the mean when repeatedly measured. Thus, if a group is selected at baseline because of their extreme scores and then moves closer to the mean at follow-up, the cause is confounded with regression.
- **Maturation -** Changes in participants may be observed due to natural maturation as opposed to any effect of treatment. This is particularly hard to tease out without a well-designed comparison group.
- **History -** Natural events in the lives of participants that occur during the duration of a longitudinal experiment may influence outcomes apart from the treatment conditions. Again, this is hard to isolate without a good comparison group.

Internally Valid Example

Overweight volunteers agreed to participate in an experimental low maintenance weight loss program. On the basis of a coin flip, each participant was assigned to either the Record-Before Group or the Control Group. Subjects in the Record-Before Group were instructed to write down everything they ate at any given sitting immediately prior to eating it; subjects in the Control Group were not given this instruction. All subjects were weighed on the first day, and many expressed shock when told their true starting weight. For the next month, all subjects had weekly one-on-one sessions with a therapist who provided empathic listening. One month later, subjects were weighed again. Subjects in the Record-Before Group lost a mean of 10 pounds, while subjects in the Control Group, on average, did not lose any weight. We can conclude that self-monitoring in the form of writing down everything one soon plans to eat was effective with respect to weight loss.

Construct Validity

- The word "construct" typically refers to a latent (i.e., not directly measurable; not observable) outcome variable that is of primary interest in a study. This might be, for example, intelligence, depression, political power, self-efficacy, anxiety, etc.
  - The word "construct" is used because latent variables are typically measured by using a scale that is *constructed* of many single items, which, taken together, can be claimed to be a reliable measure of the underlying latent trait. Consider an IQ test of intelligence as an example.
- Poorly defined constructs are not useful. As MD note, a "major pitfall" related to construct validity is "*inadequate preoperational explication* of the construct" (p. 67).
  - In particular, a construct must be multifaceted, as needed, so as not to over-simplify the trait it means to represent.
  - Further, a construct must contain items that distinguish it from other, related constructs. Consider anxiety and depression. They are highly correlated but different.

External Validity

- External validity is related to the ability to generalize findings from an experiment across populations, settings, or time.
  - External validity is supported by design when random sampling from a defined population is used. However, this is almost never done.
  - Most typically, experimenters prioritize internal validity over external validity; experimenters tend work with the participants they have access to.
    - In such cases, claims to generalizability must be based on logical arguments by pointing out characteristics of the sample that are similar across the experimental group and the group to which generalization of findings is desired.
  - It is also possible to design field experiments with generalization in mind. See Beth Tipton's website for some examples.

Internally Valid Example

Overweight volunteers agreed to participate in an experimental low maintenance weight loss program. On the basis of a coin flip, each participant was assigned to either the Record-Before Group or the Control Group. Subjects in the Record-Before Group were instructed to write down everything they ate at any given sitting immediately prior to eating it; subjects in the Control Group were not given this instruction. All subjects were weighed on the first day, and many expressed shock when told their true starting weight. For the next month, all subjects had weekly one-on-one sessions with a therapist who provided empathic listening. One month later, subjects were weighed again. Subjects in the Record-Before Group lost a mean of 10 pounds, while subjects in the Control Group, on average, did not lose any weight. We can conclude that self-monitoring in the form of writing down everything one soon plans to eat was effective with respect to weight loss.

Assignment 1 - KEY

Instructions:

Please answer the following questions. Your answers should be typed using 12pt Times New Roman font, double-spaced, with one-inch margins on all sides. Include a page number, the assignment number, and your name at the top of each page. Assignments should be **submitted through Canvas no later than 6PM on the day of our Week 5 class**.

Your answers should be in sentence form and self-contained. If the questions asks "Was the study described by Stenglemaier (1900) an experiment in the sense of the term as used by Campbell and Stanley (1963)?" a response of "Yes" is not sufficient. Rather, please write something like the following:

> Stenglemaier (1900) was interested in whether paying workers more increased their production. He took a group of Ironton, Ohio, foundry workers who had been making an average of 500 railroad ties a day, at wages of $2 per day,

A doctor has a habit of measuring blood pressure twice. She notices that patients who are unusually high on the first reading tend to have somewhat lower on the second reading. She concludes that patients are more relaxed on the second reading. A colleague disagrees, pointing out that patients who are unusually low on the first reading tend to have somewhat higher second reading, suggesting they get more nervous. Which doctor is right? Or perhaps both are wrong? Explain briefly. **[33 pts]**

A doctor has a habit of measuring blood pressure twice. She notices that patients who are unusually high on the first reading tend to have somewhat lower on the second reading. She concludes that patients are more relaxed on the second reading. A colleague disagrees, pointing out that patients who are unusually low on the first reading tend to have somewhat higher second reading, suggesting they get more nervous. Which doctor is right? Or perhaps both are wrong? Explain briefly. **[33 pts]**

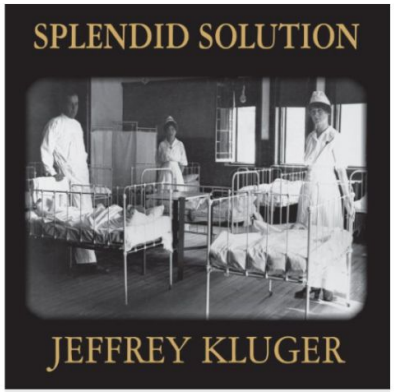
– Both doctors are wrong. Defining that the treatment as the first blood pressure reading, it becomes evident that both doctors are exclusively looking at observed outcomes under treatment, and at none under control.

A doctor has a habit of measuring blood pressure twice. She notices that patients who are unusually high on the first reading tend to have somewhat lower on the second reading. She concludes that patients are more relaxed on the second reading. A colleague disagrees, pointing out that patients who are unusually low on the first reading tend to have somewhat higher second reading, suggesting they get more nervous. Which doctor is right? Or perhaps both are wrong? Explain briefly. **[33 pts]**

- Both doctors are wrong. Defining that the treatment as the first blood pressure reading, it becomes evident that both doctors are exclusively looking at observed outcomes under treatment, and at none under control.
- In other words, both doctors are wrong because with this "design" all units are treated and there are no controls. So it is impossible to estimate any causal effect.

A doctor has a habit of measuring blood pressure twice. She notices that patients who are unusually high on the first reading tend to have somewhat lower on the second reading. She concludes that patients are more relaxed on the second reading. A colleague disagrees, pointing out that patients who are unusually low on the first reading tend to have somewhat higher second reading, suggesting they get more nervous. Which doctor is right? Or perhaps both are wrong? Explain briefly. **[33 pts]**

- Both doctors are wrong. Defining that the treatment as the first blood pressure reading, it becomes evident that both doctors are exclusively looking at observed outcomes under treatment, and at none under control.
- In other words, both doctors are wrong because with this "design" all units are treated and there are no controls. So it is impossible to estimate any causal effect.
- Comparing the same unit with two measurements at different points in time does not solve the problem, since we would need to know how each unit would have behaved *if it not had been treated.* The second measurement is only the *observed outcome conditional on having been treated before*, but we still lack the counterfactual.



Sample

Splendid Solution

Jonas Salk and the Conquest of Polio

By: Jeffrey Kluger
Narrated by: Michael Prichard
Length: 13 hrs and 12 mins
Unabridged Audiobook
Release date: 02-03-06
Language: English
Publisher: Tantor Audio
4.4 (234 ratings)
Whispersync for Voice-ready

Play
In your Library
More options