Addressing Causality

James H. Steiger

Department of Psychology and Human Development
Vanderbilt University

Addressing Causality

Introduction

- When Russ Whitehurst took over the IES at its inception in 2002, it quickly became apparent that administrators, teachers, and polcymakers were preoccupied with issues of causality.
- All other things being equal, would a particular administrative action or policy change or "change in input" result in a change in outcome?

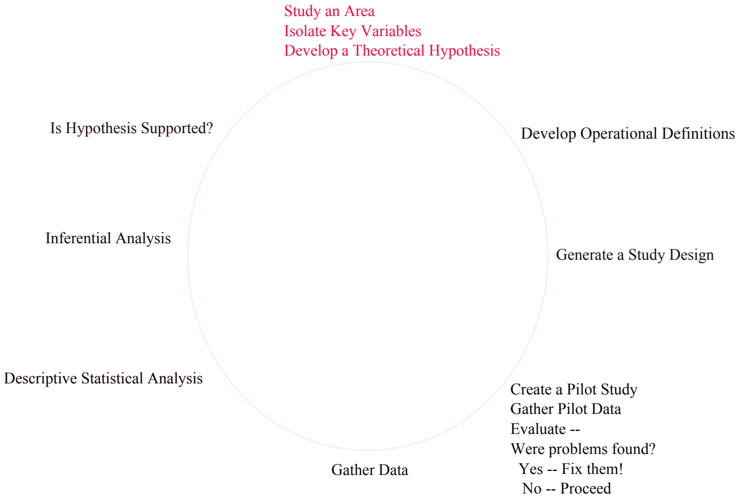
Conditions to Strive For

- A clear statement of the research question and the theory that drives it.
- A clear statement of the population of interest.

The Circle of Science

- Scientists have made tremendous progress in the last 3 centuries by employing a systematic approach to the development of knowledge.
- This approach is loosely referred to as "the scientific method" in many textbooks.
- Let's review fundamental aspects of this approach.

The Circle of Science



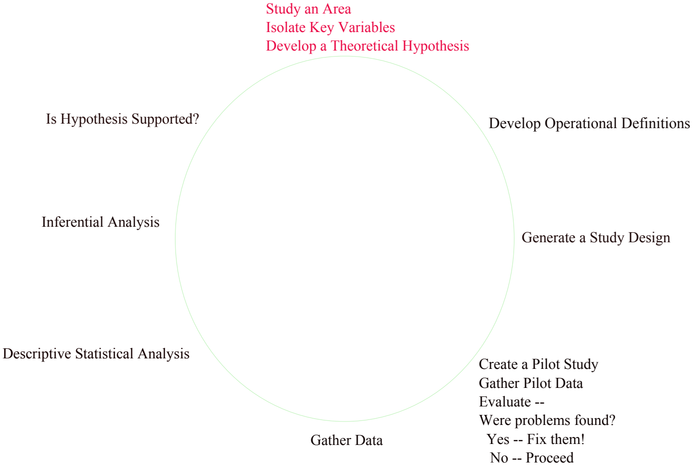
The Circle of Science

Hypothesis Formulation

- In this stage, the scientist picks an area of inquiry and studies what work has gone before.
- A critical analysis of previous work in the field can raise important theoretical insights into new possible avenues for research.
- The scientist develops an hypothesis about the relationship between scientific variables.
- Independent (or "exogenous") variables are viewed as affecting dependent (or "endogenous") variables.
- A simple kind of hypothesis is that "variable X affects variable Y."
- What are some examples of such a simple, two-variable hypothesis?

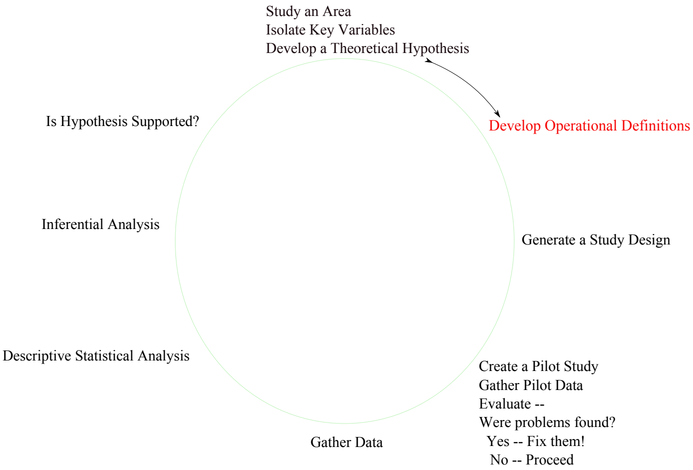
The Circle of Science

Hypothesis Formulation



The Circle of Science

Hypothesis Formulation



The Circle of Science

Operationalization

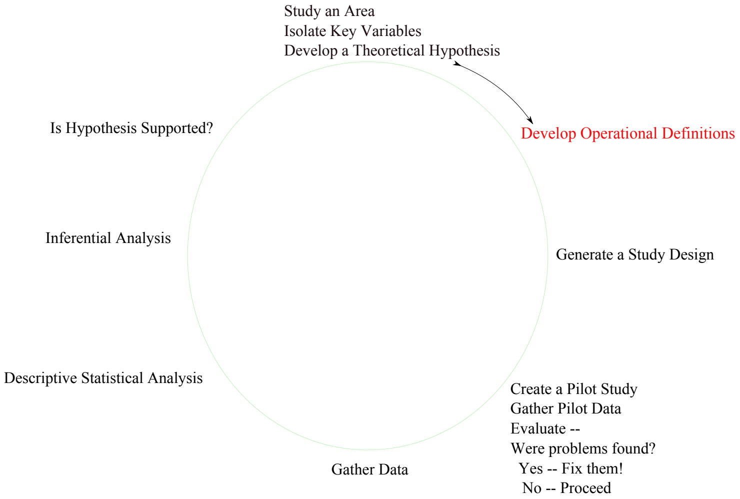
- When the scientist starts to plan a study, one of the first steps is *operationalization* of the key variables in the study.
- For example, suppose you were studying the impact of “intelligence” on income in a large organization.
- You would have to develop *operational definitions* for each of those variables.
- For example you might operationalize intelligence as the IQ score achieved by each employee on a well-known standardized IQ test.

The Circle of Science

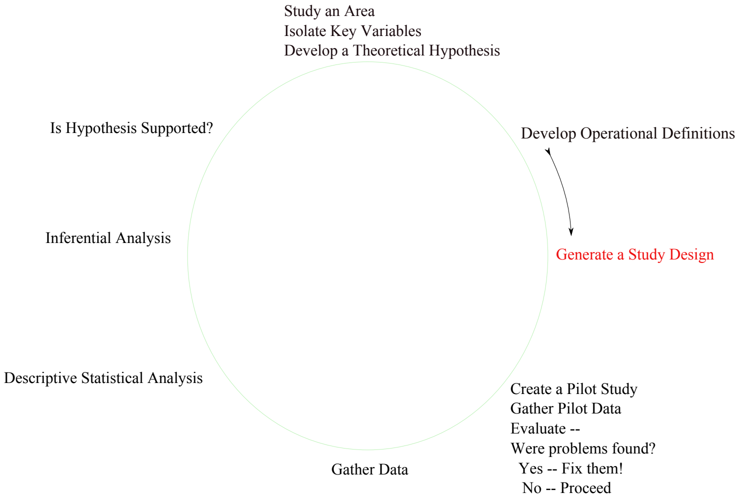
Operationalization

- Operationalization involves tradeoffs between generality, simplicity, and replicability.
- A simple, well-defined measure makes a study easier to do, and easier to replicate.
- However, if the measure is too simple, it may fail to capture important aspects of the construct being measured.
- Many people are not convinced that IQ scores adequately measure intelligence, for example. But they are effective at predicting performance within large organizations, and it seems they do correlate with what is commonly thought of as intelligence.

The Circle of Science



The Circle of Science



The Circle of Science

Choosing a Design

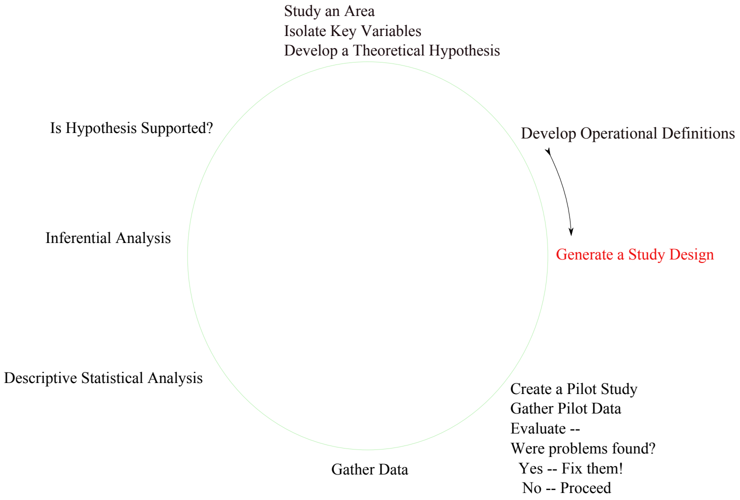
- One basic design choice is between *manipulative* ("Experimental") research and *non-manipulative* ("Observational," "Correlational," or "Quasi-Experimental") research.
- In manipulative experimentation, the experimenter manipulates the independent variable X and observes its effect on the dependent variable Y while trying to hold other variables constant.
- In non-manipulative research, the experimenter observes X and Y, and but does not directly manipulate either.

The Circle of Science

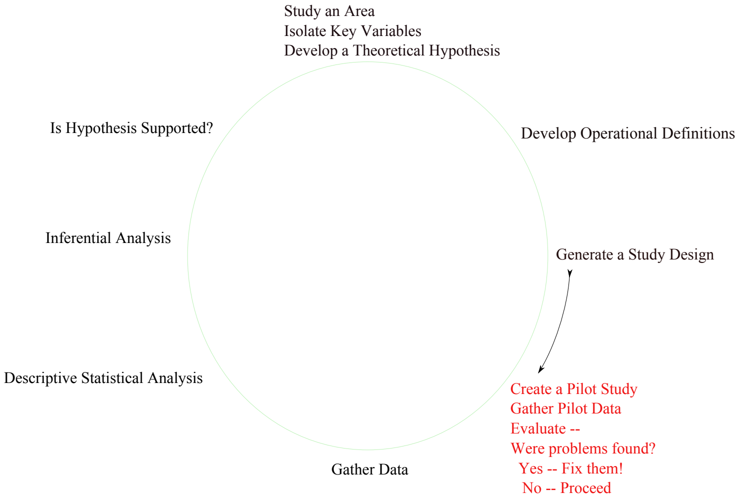
Choosing a Design

- One of the best known manipulative designs is the two-group, Experimental-Control group design.
- Both groups are treated exactly the same, except the Experimental group receives the variable of interest.

The Circle of Science



The Circle of Science



The Circle of Science

Pilot Studies

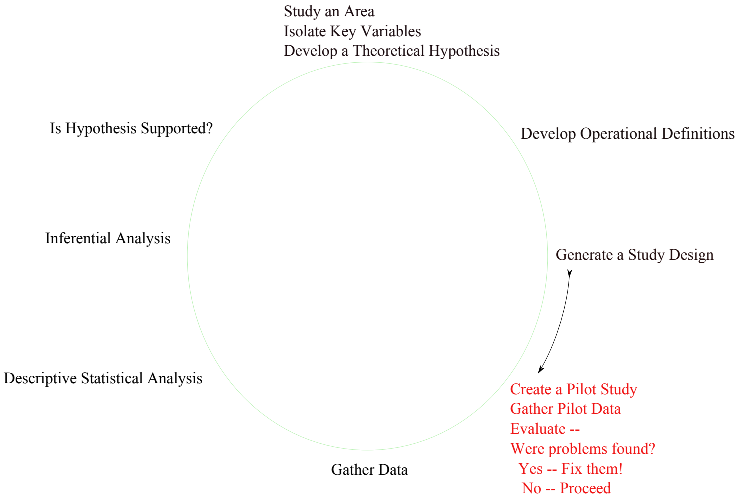
- Of course, many details have to be worked out before an experiment can get fully underway.
- A "pilot study" is often run on a small group of participants, in order to detect any obvious errors in the experimental setup.
- Problems often emerge during the pilot study stage. For example, instructions may be clear to the experimenters, but ambiguous to the participants. Or, the lighting in the room may turn out to be inadequate.

The Circle of Science

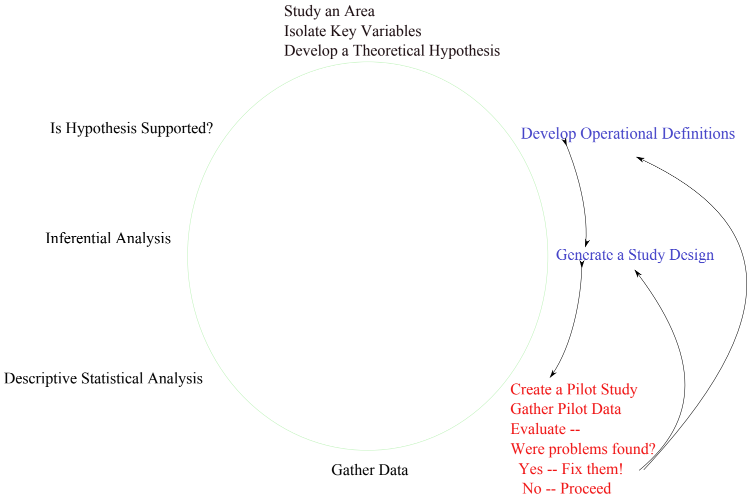
Pilot Studies

- The experimental "paradigm," consisting of the design setup, operational definitions, and specific procedures, may be refined successively through a series of pilot studies.

The Circle of Science



The Circle of Science

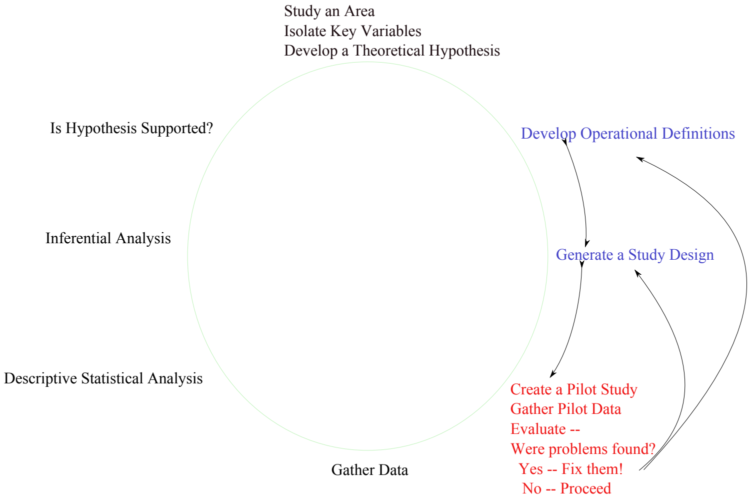


The Circle of Science

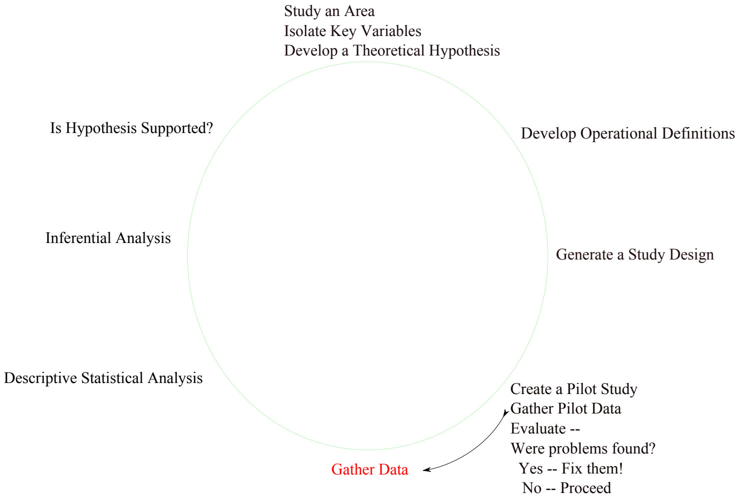
Data Gathering

- Once the paradigm has been refined to a point where there are no longer any obvious problems, the experimenters will perform *sample size analysis* in order to assess how many participants they should use in order to optimize the likelihood of making correct inferences.
- At that point, subjects are recruited (if they haven't been already) and the experiment commences.
- Weeks (or months) later, the experimenters have lists of numbers representing the performance of the subjects in the experiment. At that point, statistical analysis commences.

The Circle of Science



The Circle of Science

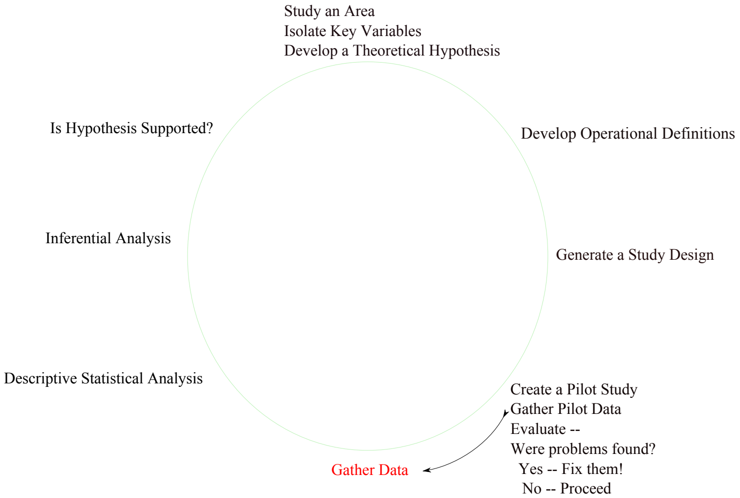


The Circle of Science

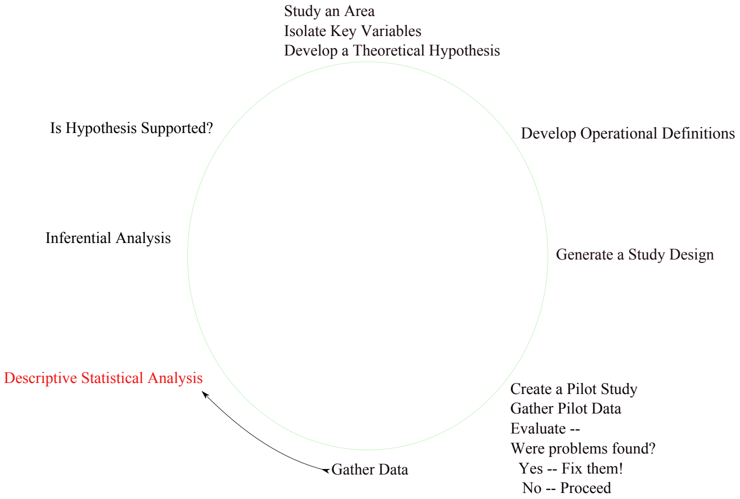
Descriptive Statistical Analysis

- Descriptive statistical analysis takes the lists of numbers and tries to describe what, if anything, is "there" in the way of important results.
- A wide variety of numerical and graphical methods are employed, and many of them have been developed only recently.

The Circle of Science



The Circle of Science

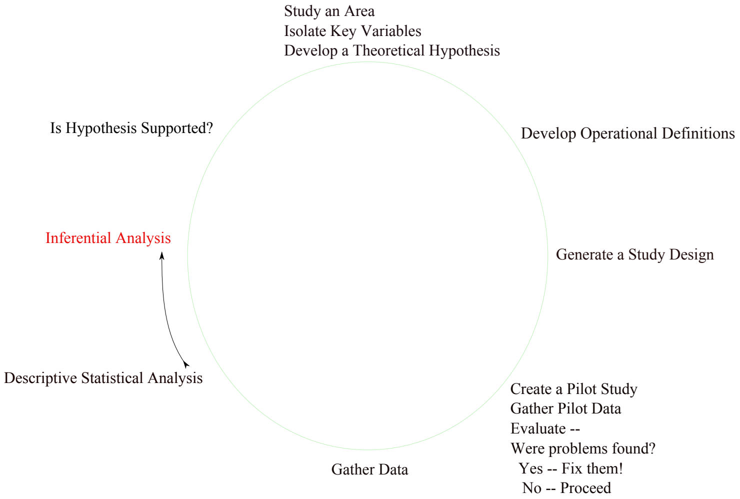


The Circle of Science

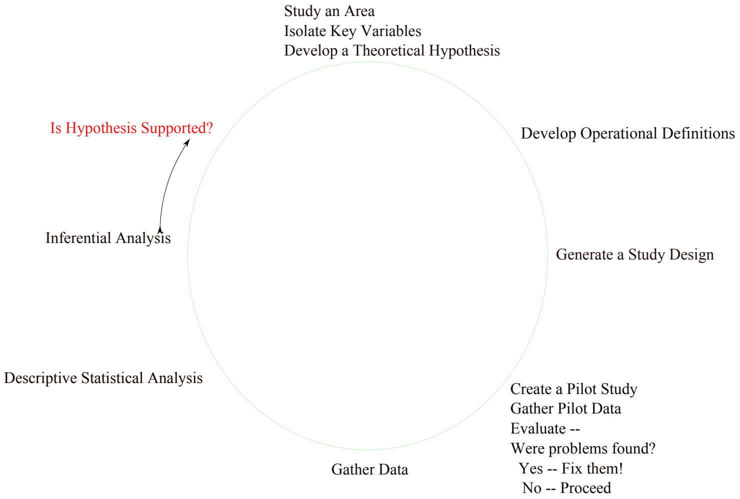
Inferential Statistical Analysis

- Suppose the descriptive analysis seems to indicate that "something is there," that is, that the manipulation had an effect.
- Ever-present is the possibility that what "is there" is just a mirage, created by "the luck of the draw."
- That is where the science of *inferential statistics* comes into play.
- Inferential statistics are techniques that attempt to apply mathematical and statistical theory to assess the likelihood that a result was caused by luck, as opposed to being indicative of a real scientific phenomenon.

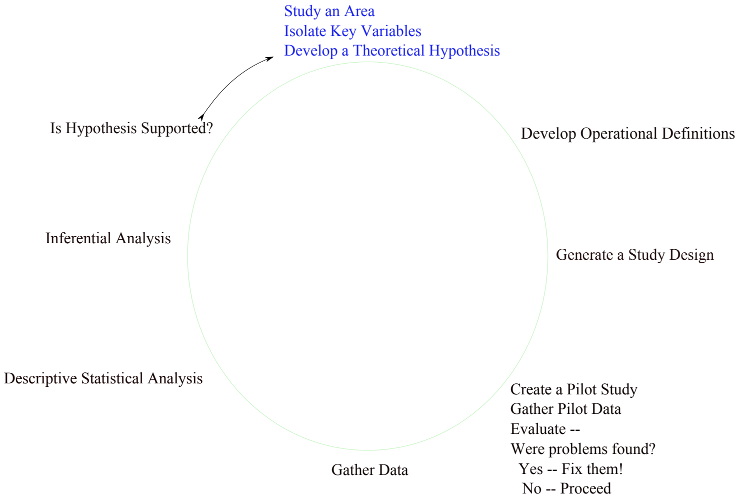
The Circle of Science



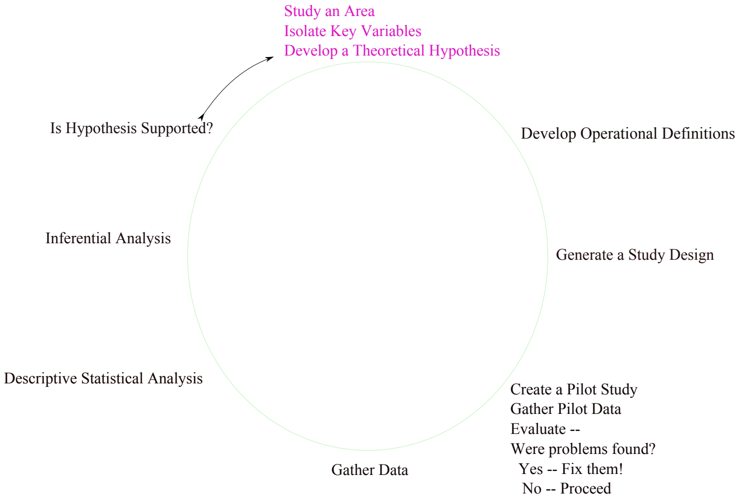
The Circle of Science



The Circle of Science



The Circle of Science



Conditions for Inferring Causality

- Time Precedence. The hypothesized cause must precede the hypothesized effect in time.
- Systematic Covariation. The levels of the hypothesized causal agent should differ in some systematic way, and there should be observed a corresponding variation in the hypothesized effect.
- Other Explanations must be Discountable. For example, suppose our data in a large observational study reveal that, as class sizes decrease, achievement in the classes (on average) shows an increase . . .

Conditions for Inferring Causality

Give an ALTERNATIVE EXPLANATION.

Randomized Experiments

- Perhaps the best way to achieve conditions for causal inference is to conduct a randomized experiment.

Definition (Randomized Experiment)

A randomized experiment is an empirical investigation in which the levels of a potential independent variable (cause) are manipulated by an outside agent functioning independently of the subjects in the research by randomly assigning subjects to different levels of the independent variable, and after which the consequences for an important dependent variable, believed to be affected causally by the independent variable, are measured.

Quasi-Experiments

- Quasi-Experiments are experiments in which the units are not assigned randomly to experimental conditions.

Example (A Hypothetical Quasi-Experiment)

Students entering the 6th grade in the city of Princeton, NJ, choose to attend one of three grammar schools for grades 6–8, two of which are public schools and one of which is a privately run Catholic parochial school. The outcome of standardized reading tests administered to 8th graders shows a significant difference between the 3 schools.

The Age of Regression

- Until recently, unfortunately, most educational researchers did not adopt creative approaches to analyzing causality.
- Rather, they relied on large scale observational studies in which many covariates were measured in an attempt to discount them, via regression analysis, as potential explanations for observed relationships between variables of central interest.
- In such observational studies, there are always countless alternative explanations for observed effects.

The Age of Regression

Example (Observational Studies of Class Size)

Many studies have examined relationships between class size and performance. But without random assignment, there are countless alternative explanations of any observed improvement in performance that might be associated with reduced class size. For example, saavy parents might move their residence to areas with smaller classes and better schools. Moreover, these same parents might give their children more attention at home.

The Key Challenge of Causal Research

- Assessment of causality in experimental design is often analyzed within the potential outcomes framework.
- Consider all the subjects who receive a treatment in a study and are assigned to the Treatment group. Their observed score on the dependent variable might be signified y_i^1, because they are in the Treatment (T) group and in fact received the treatment (as indicated by a superscript of 1).
- Then consider all the subjects who did not receive the treatment, and were assigned to the control group. Their observed scores might be denoted y_i^0.

The Key Challenge of Causal Research

- Notice that, for each subject, we might consider another, counterfactual score that in principle would have been observed had they been assigned to the other group.
- So, for example, the ith subject in the experimental group would have received a score of y_i^0 had they been in the control group.
- For the ith subject (or, if you prefer, experimental unit), the treatment effect can be defined as

$$\text{treatment effect for unit } i = y_i^1 - y_i^0 .$$

- The fundamental problem of causal inference is that at most one of the two potential outcomes can be observed for each experimental unit.

The Key Challenge of Causal Research

Hypothetical Complete Data

- In the first table on the next slide (adapted from Gelman & Hill, Figure 9.3), we see what results would be obtained if it were possible to observe each experimental unit in both the treatment and control conditions. The treatment indicator T is equal to 1 if the unit is in the experimental condition, 0 in the control condition.
- X_1, X_2, and X_3 are observed pre-treatment measures.

The Key Challenge of Causal Research

Hypothetical Complete Data

The outcome that is actually observed for a given unit is shown in boldface.

Unit i	X_1	X_2	X_3	T	y_i^1	y_i^0	$y_i^1 - y_i^0$
1	2	1	50	0	75	**69**	+6
2	3	1	98	0	108	**111**	−3
3	2	2	80	1	**102**	92	+10
4	3	1	98	1	**111**	112	−1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
100	4	1	104	1	**114**	111	+3

Below is what would actually be observed by the researcher in the same situation:

Unit i	X_1	X_2	X_3	T	y_i^1	y_i^0	$y_i^1 - y_i^0$
1	2	1	50	0	?	69	?
2	3	1	98	0	?	111	?
3	2	2	80	1	102	?	?
4	3	1	98	1	111	?	?
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
100	4	1	104	1	114	?	?

The Key Challenge of Causal Research

Causal inference = missing data analysis