ISS 305:002
Evaluating Evidence:
Becoming a Smart Research Consumer

9. Establishing Causal Relationships

Reminder: Turn on your I<CLICKER

http://www.factcheck.org/2016/11/how-to-spot-fake-news/

Types of relationships

- We' ve considered several types of relationships:
  - (predictive/probabilistic) relationships
  - perfect relationships
  - weak vs. strong relationships
  - statistically significant relationships
  - practically significant relationships
- We' ll now consider another important type of relationship– a causal relationship
- What' s your intuitive understanding of the statement X causes Y (e.g., smoking causes lung cancer)?
  - implies ability to control Y by creating differences in or **manipulating** X

Establishing Causal relationships

I. Defining causal relationships
II. Illustrating that correlation does not establish causation
III. Correlation & causation
IV. The experimental method
   A. Defining a true experiment
   B. Between vs. Within
      1. Internal and External Validity
   C. Other common threats

Causal relationships

- There is a **CAUSAL RELATIONSHIP** between variables A and B (or **A CAUSES B**) when **all three** of the following conditions can be met:
  - there is a relationship between A and B
  - changes in A precede (in time) changes in B
  - there is no other third variable, C, which covaries with A
- Thus, just having the first of these **is not sufficient to establish a causal relationship**, or
- **Not all relationships are causal relationships!**
- **Covariation/Correlation does not imply causation!**
- **One way to grasp this is to consider a number of non-causal relationships**
  - **they satisfy the first, but not both of the other two conditions**
  - **some will be obvious and even silly; some are not; all illustrate the same fundamental point**
  - See Huff Ch. 8 and Stanovich Ch. 5 for more examples

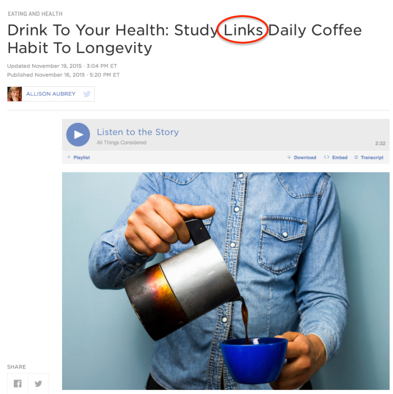
Methods of Research

- Two general approaches to studying the relationships between variables

1. Nonexperimental (correlational) methods
   - Relationships are studied by making observations or measures of the variables of interest
2. Experimental methods
   - Relationships are studied by direct manipulation and control of variables

Correlational Research

- Measure the variables as they already exist in nature
- A relationship is established by finding the degree to which two variables covary
  - As one variable changes, what does the other variable do?
- If you are interested in the relationship between alcohol use and GPA
  - For a group of people measure how much they drink and what their GPA is
  - Look for a correlation



May predict

Is correlated with

Is tied to

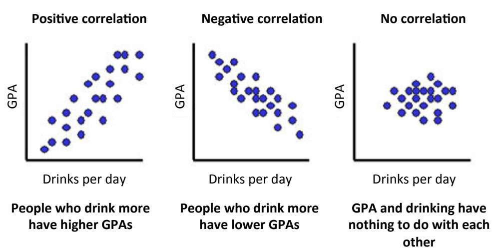
Is associated with

Correlational Research

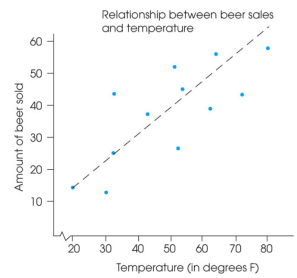
- We often want to know if two variables go up and down together or if one goes up while the other goes down.
  - Is alcohol consumption (Y) negatively related to heart disease (X)?
  - Does the accuracy of performance (Y) decrease as speed of response (X) increases?
  - Are people who are high in extraversion also high in openness?

Direction of correlation

...each circle is one person

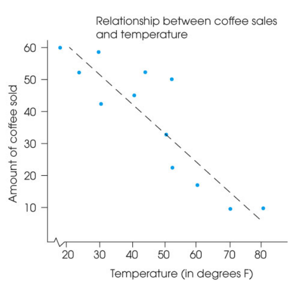


Positive Correlation Example



- The relationship between beer sales and temperature is positive
- As the temperature increases (X), beer sales (Y) increase
- As the temperature decreases (X), beer sales (Y) decrease

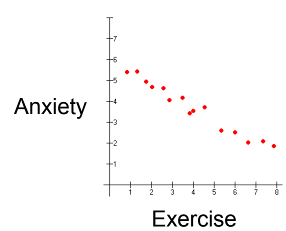
Negative Correlation Example



- The relationship between coffee sales and temperature is negative (or inverse)
- As the temperature increases (X), coffee sales (Y) decrease
- As the temperature decreases (X), coffee sales (Y) increase

Correlation Research Example

- Suppose you are interested in anxiety
- You find a negative correlation between exercise and anxiety
- What does a negative correlation mean?
- Does Exercise reduce anxiety?
- **Correlation is not causation**



As the amount of exercise increases, anxiety decreases.

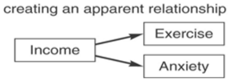
As the amount of exercise decreases, anxiety increases.

Correlation Research Example

1. The first problem is direction
- Exercise → Anxiety
- Anxiety → Exercise
- Can't be sure without establishing temporal precedence

2. The second problem is known as the **third-variable problem**
- Maybe exercise has no effect on anxiety
- Maybe anxiety has no effect on exercise
- Maybe income causes exercise & anxiety

A third variable such as income is associated with both variables, creating an apparent relationship between exercise and anxiety.



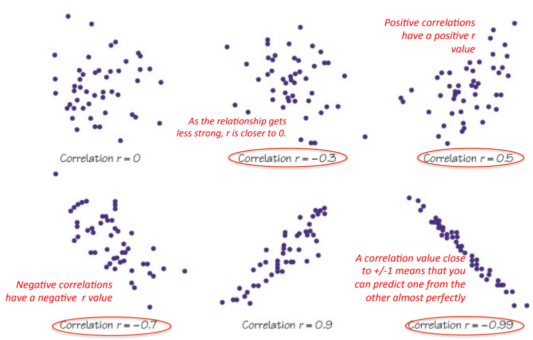
Strength of correlation

- How well the points form a line gives us the **correlation coefficient**
- Statistical index of the strength of the relationship between two variables
- Often represented with the symbol, *r*
- *r* varies between -1 and 1
- In psychology, we have some guidelines to evaluate how strong an association is

An *r* of approximately	Effect size
.10 (or -.10)	Small / weak
.30 (or -0.30)	Medium / moderate
.50 (or -0.50)	Large / strong

Cohen, 1992

Strength of correlation



Correlation does not imply causation

- Correlation does not imply causation!
  - A correlation between X and Y does not imply a cause and effect relationship between X and Y.
  - Correlation is not sufficient to establish causality.
  - http://tylervigen.com/
  - http://www.bbc.com/news/health-24710089
  - http://abcnews.go.com/Health/WellnessNews/butt-thigh-fat-make-healthier/story?id=9534982
- In order to arrive at causality we must conduct an experiment.

Why does correlation not imply causation?

-

-

-

  -

-

Why does correlation not imply causation?

The more ice cream consumed →

Correlation ≠ Causation

- Example #1
  - In Taiwan, a positive relationship has been found between owning a toaster (A) and low birth rates (B)
  - Does this mean that family planners in Taiwan should buy young couples toasters to reduce the birth rate?
    - Does A ➔ B?
  - Why not?
  - Which of the other two necessary conditions for a causal relationship is not met?

Correlation ≠ Causation

- Example #2
  - Statistics show that people who live in Arizona (A) are more likely to die of lung disease (B) than residents of any other state.
  - Does this mean that the air in Arizona is especially polluted, causing more lung disease (A➔B)?
- Why not?
- Which of the other two necessary conditions for a causal relationship is not met?
- Most likely, the real causal chain<u>s</u> are...

Correlation ≠ Causation

- (Non obvious) Example #3
  - It was discovered that those who were poor readers (B) scanned text with their eyes in irregular ways (A)
    - instead of scanning left to right, line by line, their eyes would wander all over the page
  - It was concluded that correcting irregular eye movements (A) would improve reading skills (B) (i.e., A➔B)
    - and many expensive remedial eye-movement programs were begun
  - They turned out to result in no improvement in reading.
- Why not? What was wrong with the inference that A➔B?
- Which of the other two necessary conditions for a causal relationship is not met?
- Most likely, the real causal chain is...

Correlation ≠ Causation

- (Non obvious) Example #4
  - Early public health studies showed that when parents had tuberculosis (A), the probability was higher that their children also would get the disease (B)
  - Is it reasonable to conclude that something about the parents (e.g., the genes they pass on to their children) make their children more vulnerable to the disease?
    - that parent's vulnerability ➔ children's vulnerability?
- Why not?
- Which of the other two necessary conditions for a causal relationship is not met?
  - Probably
- Most likely, the real causal chain is...

Correlation ≠ Causation

- (Non obvious) Example #5
  - Pellagra was once a serious disease, especially in the American South and in South America
    - Symptoms include sensitivity to sunlight, diarrhea, insomnia, and red skin lesions
  - The poorest people, who also had very poor sanitation (A) were the most likely to contract pellegra (B)
  - Is it then reasonable to conclude that improving sanitation will reduce the level of pellegra (i.e., that A➔B)? Why or why not?
- Which of the other two necessary conditions for a causal relationship is not met?
- Subsequent research has shown that the actual causal chain is...

Correlation ≠ Causation

- (non obvious) Example #6
  - A study shows a <u>negative</u> correlation between high school teacher salaries (A) and the mean SAT score of students from their high schools (B)
    - students from schools with higher paid teachers got lower SAT scores
  - Does it make sense to lower teacher salaries to improve SAT scores (i.e., to conclude that A➔B)? Why or why not?
- Which of the other two necessary conditions for a causal relationship is not met?
- Probably the actual causal chain is...

Correlation ≠ Causation

- (Non obvious) Example #7
  - Many studies show that students who attend private schools (A) tend to do better in college (B) than students who attend public schools.
  - Does this mean that the extra expense of sending your child to private school pays off in better college performance (and better opportunities thereafter)? (i.e., that A➔B)?
- Why not? Which of the other two necessary conditions for a causal relationship is not met?
- A likely causal chain is...

How and when can one get from (cor)relation to causation?

1. One can measure and check the influence of other, potential causal variables
   - The correlation between type of school and academic achievement vanishes when one <u>controls for</u> (i.e., holds constant) family economic background (e.g., parental education, socioeconomic status, number of books in home, etc.) and student academic aptitude (e.g., IQ)

- This is a way of showing that A ~~➔~~ B, not that A ➔ B
  - **by eliminating one, or two, or many possible "third variables", one does <u>not</u> guarantee that there are not other, unexamined causal variables**
  - an inability to think of a plausible "third variable" does **not** mean that it doesn't exist
    - There's no obvious "third variable" for the correlation of living by power lines and childhood leukemia, but maybe it is non-obvious

How and when can one get from (cor)relation to causation?

- It is the use of a particular type of method (the "correlational" method, where many variables are uncontrolled) that gives rise to this ambiguity, not whether or not alternative causes are apparent
- Need to establish the relationship using methods which guarantee all three necessary conditions-- **the experimental method**
- **Moral:**

Causal Heuristics

- The Spectacular Explanation Fallacy
  - Extraordinary events do not require extraordinary causes
    - "There are more things in heaven and earth, Horatio, Than are dreamt of in your philosophy."
      - Hamlet, Act I, Scene 5 (William Shakespeare)
    - "Shit happens." – Anonymous
    - Example – "Hot" streaks in gambling, but are expected by chance alone. They do not require extraordinary causes (rituals, lucky rabbit's feet, etc.).
- The "Jim twins" vs. the "Patricia Ann Campbell's"
- Dr. Weaver will present more examples that he will post as blue notes.

Causal Heuristics

- The Intervention-Causation Fallacy
  - The cure doesn't prove the cause
    - We may be able to surgically remove a cancerous tumor, but our ability to do so does not, in any way, explain what originally caused the creation and development of the cancerous tumor.
    - Aspirin and headache example
- The Insight Fallacy
  - To understand something isn't necessarily to change it
  - "It is naïve to expect that, by telling people what we think we see they are doing, we will enable them to stop doing it." – R. D. Laing

More complex causal possibilities

(see Ch. 9, Stanovich):

1. A➔B or B➔A
2. A ⬅➔ B (two-way or reciprocal causal relationship)
   - What are examples?
     - **Social isolation & Aggression**
     - **Gun ownership & Crime**
3. A➔B when C is at one level, but not when C is at another. (C **moderates** the relation between A & B)
   - Example?
     - **spark (A) will cause combustion (B) of gasoline IF oxygen (C) present; spark will not cause combustion of gasoline IF oxygen absent**
       - **oxygen is a *necessary* cause (but not a *sufficient* cause) of fire**

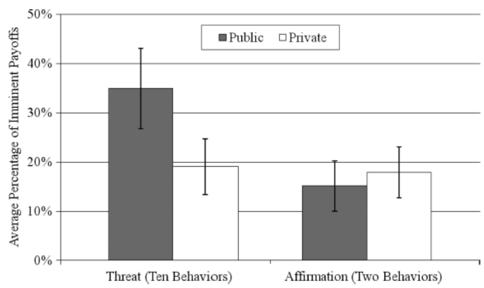


Figure 3a. Average percentage of imminent financial payoffs selected as a function of gender threat and audience (Experiment 2). Error bars represent standard errors of the mean.

More complex causal possibilities

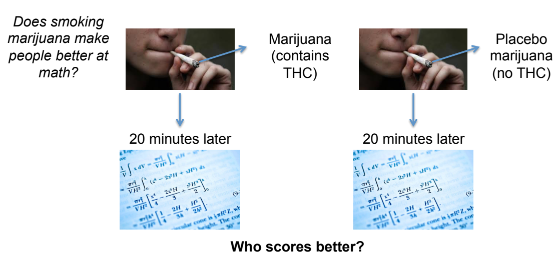
(see Ch. 9, Stanovich):

4. A & C & D & … ➔ B (multiple causation)
   – What are examples?
     - If B is being diagnosed with HIV?
       – **causes include unsafe sex with an HIV-positive person, sharing needles, getting transfused with infected blood, etc**
     - If B is prejudice?
       – **causes include personality traits, conflict between groups, beliefs about outgroup members, etc.**

Experiment

Marijuana condition	
Contains THC	**No THC**
Average math score after smoking real marijuana	Average math score after smoking fake marijuana

Manipulate one variable and control for extraneous variables



Variables

- Independent variable (IV)
  – Variable that is **manipulated** by experimenter.
- Dependent variable (DV)
  – Variable that is **measured/observed/recorded** by experimenter.
  – DV "depends upon" the independent variable
  – In other words, this is your **outcome** variable
- Control Variable (CV)
  – Held constant during an experiment. WHY?
    - Time of day, temperature, time of last meal, knowledge of subject

Key components of experiments

1. Manipulated independent variable (IV) - The factor of interest to the experimenter, the one that is being studied to see if it will influence behavior.
   A. *At least* 2 conditions.
   B. Control of extraneous variables
2. Statistical testing of hypotheses
   A. Most often, using Analysis of Variance (ANOVA)
   B. Probability that differences are due to chance is < .05.

Key components of experiments

1. Manipulated Independent Variable

A. At least 2 conditions.
– Usually at least ONE experimental condition and ONE control (or placebo) condition.
  - In the experimental condition, the IV is present.
  - In the control condition, the IV is absent.
– Not all experiments have a true "control" condition.
– In a Quasi-experiment, called the grouping variable (GV).

Key components of experiments

1. Manipulated Independent Variable

B. Control of extraneous variables

- The experimental and control conditions are *exactly the same except for* the presence/ absence of the IV. (Time of day, temperature, time of last meal, knowledge of subject)
  - **Extraneous variables** = any variable that the experimenter wants to hold constant on purpose to eliminate alternate explanations for our data.
- Also includes **random assignment to condition**

Random Assignment to Condition

- It means that each participant has an equal chance of being in any particular condition.
  - Conditions are independent of the participant.
  - Random assignment is ***necessary*** to establish causality in a true experiment.
  - **HELPS** Controls for extraneous / confounding variables
    - Hopefully equals everything out and decreases the probability of selection effects/biases
  - However, you can conduct an experiment without *true* random assignment.

Experiment example

•**Independent and Dependent Variables**

•**An example:** Does breastfeeding children have an impact on their intelligence later in life?

•The **independent variable** is whether mothers are assigned to an experimental condition in which their children breast feed, or to a control condition in which they feed their children formula.

•**Random assignment –** assigning participants to experimental and control groups by chance, thus minimizing preexisting differences between the groups that could affect the dependent variable (e.g., SES).

Experiment example

•**An example:** Does breastfeeding children have an impact on their intelligence later in life?

•The **dependent variable** is the child's score on an intelligence test at age 8. (This is the measureable aspect of psychological functioning that we think may be influenced by the independent variable.)

Condition	Independent variable	Dependent variable
Experimental	Breast milk	Intelligence score, age 8
Control	Formula	Intelligence score, age 8

Benefits of Experiments

1.

2.

3.

Limitations of Experiments

1.

2.

3.

4.

5.

Quasi-Experiments

- "Look" like real experiments
- Independent variable is a pre-existing **grouping variable (GV),** and as such is not manipulated
  - Gender, Ethnicity
  - Pre-existing attitudes
    - e.g., Democrats & Republicans
  - Personality or trait variables
    - e.g., introverts and extroverts

Types of Experimental Designs

- Between-Subjects
  - Each participant participates in one and only one condition
  - Each participant produces only one score per DV.
  - Random assignment to condition
- Within-Subjects
  - Each participant participates in all conditions
  - Each participant produces 1 score per DV for each condition
    - number of scores produced = number of conditions
  - Randomly assign *order* of condition

Key Components of Experimental Design: Internal Validity

- **Internal validity**
  - Is the study internally consistent,
    - i.e. is the study logic solid?
    - Are there no other alternate explanations for the results?
  - Internal validity allows for **inferences of causality**
    - i.e., X caused Y; IV causes DV.
  - Achieved by
    - **random assignment to condition &/or order**
    - **control of extraneous variables**

Key Components of Experimental Design: Internal Validity

- **Some Threats to Internal validity**
  - Ambiguous Temporal Precedence
    - Does the causal variable come *before* the effect variable in time?
    - Does A precede B or vice versa?
  - Selection (cohort effect)
    - Differences between groups from the start
    - Maybe you accidentally assign one type of person to one condition and one to the other (e.g. people who register early or late for SONA credit)
    - To avoid this problem, use random assignment
  - History
    - Events outside of the experiment may affect responses
  - Maturation
    - Naturally occurring changes over time
  - Regression
    - Phenomenon that if a variable is extreme on its first measurement, it will tend to be closer to the average on its second measurement

Key Components of Experimental Design: Internal Validity

- **Some Threats to Internal validity**
  - Attrition/Mortality
    - Some participants fail to complete the DVs
    - Problem arises when certain *types* of people drop out of the study before completing the post-test
  - Testing
    - Repeated testing might influence scores
  - Instrumentation
    - Change in a measuring instrument
    - Including a comparison group helps to check the instrument
  - Additive and Interactive Effects of Threats to Internal Validity
    - More than one threat occurring (Selection-history threat and Selection-attrition threat)

Key Components of Experimental Design: External Validity

- **External validity**
  - To what populations, settings, independent variables, and dependent variables can this effect be generalized?
  - Also known as **generalizability**.
  - The ability to apply the results of your experiment to other ***people*** or ***situations***.
  - External validity allows you to generalize your findings.
    1. **Population generalizability (people)**
    2. **Situational generalizability (situations)**

Population (*people*) Generalizability

- To ***what people*** can you generalize?
- Allows you to assume that the results of your study are true for some larger group of **people**
- Two types
  1. **Statistical** generalizability
     - You must have random sampling.
     - Allows you to generalize to the population from which you randomly selected.
  2. **Practical** generalizability
     - The results may hold for similar individuals.
     - You can generalize to similar individuals (i.e., college students).

Situational Generalizability

- To ***what other situations*** can you generalize?
  – Can findings from the lab be applied to real life?
- What are the limitations of the experimental setting compared to real life?
  1. How is the **research setting** different from other settings?
  2. How were the variables (IV's & DV's) **operationalized**?

Operational Definitions

- Definitions of theoretical constructs that are stated in terms of concrete, observable procedures.
- The goal is to connect *unobservable* traits or experiences (e.g., fear, attention) to things that can be *observed* and *measured* (e.g., self-report, behavior, physiological responses.)
- Often verbed: Operationalize.

Threats to External Validity

- Some already covered
- Interaction of the Causal Relationship (I.C.R) with Units (a.k.a. populations)
  – Does the effect found in one population hold for another population?
- I.C.R. Over Treatment Variations
  – Effect found for one treatment may not hold with other variations
- I.C.R. with Outcomes
  – Does the treatment hold across different DVs?
- I.C.R. with Settings
  – Does the treatment hold across different settings?

Other Common Threats

- Many common threats to the quality of experiments. We'll consider 4:
  – Treatment confounds
  – Demand characteristics
  – Experimenter effects
  – Weak/invalid manipulations

Look out for: Treatment confounds

- Where the experimenter has created differences not just on the independent variable, but other variables too
- Generally, how are confounding variables avoided?
- **By anticipating and controlling them**
- **Ideal experiment has no confounding variables; good real experiments have no plausible confounding variables (that could plausibly affect the DV)**
- **As a critical consumer, you need to ask not just what has the experimenter tried to manipulate, but what are all the variables that they may have (unintentionally) actually manipulated**

Look out for: Demand Characteristics

- Cues available to participants in an experiment that enable them to figure out what is expected the experimenter. May alter their natural behavior.

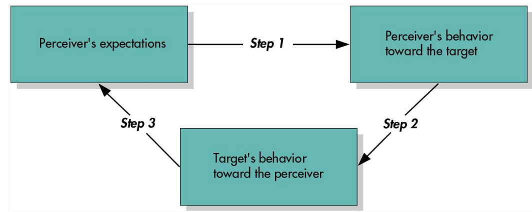
- How can such problems be avoided?
  - **by making the true purpose of the experiment unclear**

Look out for: Experimenter effects

- Treating Ps differently depending on the condition they are in.
  - Self-fulfilling prophecy
- How can such experimenter effects be avoided?
  - **Do "double blind" experiments, where**
    - **the participant is "blind" as to which condition they are in**
    - **investigator/experimenter is "blind" (uninformed) as to which condition each participant is in**
  - **Also have experimenters blind/uninformed as to the hypothesis being tested**

Look out for: Experimenter effects

- Treating Ps differently depending on the condition they are in.
  - Self-fulfilling prophecy



Self-Fulfilling Prophecies

- Word, Zanna, & Cooper (1974)
  - Experiment 1
    - White participants played the part of an interviewer
    - Interviewed an applicant - either a White or Black confederate
    - Results showed that the Black applicant received:
    -
    -
    -

Self-Fulfilling Prophecies

- Word, Zanna, & Cooper (1974)
  - Experiment 2
    - White participants were interviewed by a confederate interviewer
    - White applicants received behaviors (treated like) that were similar to those given to either the Black or White applicants from Exp. 1
    - Results showed that those who were treated like the Black job applicants were treated in Exp. 1:
    -
    -
    -

When the experiment produces null results (no relationship), look out for:

- **Floor/ceiling effects.**
  - If scores on DV in one condition are already as high (low) as they can go, then no manipulation can make them any higher (lower)
- **Weak manipulation ("restriction of range" again)**
  - If only tiny differences have been created on the IV, then we're not surprised at no differences on the DV.
- **Insensitive measure of DV**
  - Not able to detect the differences based on your measure.
- **Invalid manipulation of IV**
  - Unless one has **really** created differences on IV, don't expect effect on DV (even if A --> B)
  - How can this be avoided?
  - Requires a ***manipulation check***—a direct measure of the IV

Generalizability Example

The Ford Motor Company would like to re-evaluate their selection program for auto mechanics at Ford dealers. An industrial/organizational psychologist develops a mechanical aptitude test designed to predict job performance. To test this potential selection criterion, all of the names of mechanics employed by Ford for at least the last three years were entered into a random drawing. Based on this drawing, 200 of the mechanics were chosen to take the aptitude test. Supervisory ratings on a scale from 1 to 10 (where 10 = best possible performance) from the most recent annual evaluation were used as a measure of job performance. The psychologist divided the 174 mechanics that completed the aptitude test by the due date into two groups based on their aptitude test score: one group of high scorers (those who scored at or above the 50th percentile) and one group of low scorers (those who scored below the 50th percentile). She then compared the supervisory ratings of these two groups.

- Is this an experiment? ___________
- Statistical Population Generalizability:
  – To what population of people can you statistically generalize?
- Practical Population Generalizability:
  – To what population of people can you practically generalize?
- Situational Generalizability:
  – To what other situations can you generalize?
    - Research setting?
    - Operationalization of variables?

Is this study internally valid?

Based on her own experience, Professor Neuberg strongly believed that study groups enhance student learning. During the very first class she taught, she encouraged her students to contact Joe, a Learning Skills Specialist, at Counseling Services. She told them that Joe and his colleagues would help them form groups, make arrangements for where and when to meet, and supervise the meetings. About half the students took Professor Neuberg's advice. They met once a week in their respective study groups for the duration of her course. Counseling Services kept a record of participating students, which Professor Neuberg used for her data analysis. Her dependent measure was the final exam grade, which was marked by her teaching assistant. She discovered that students who met weekly with study groups (Study Group condition) scored two letter grades higher than students who did not meet weekly with study groups (Control condition). She attributed the higher exams scores to the study group meetings.

Research report example #1:

Ed Gogek, upset that he got called out about his correlation study showing a relationship between marijuana use and lower IQ, decides to do his own study on the effects of marijuana on IQ. He comes to the first day of the Spring 2015 ISS305 class (not taught by Dr. Weaver) and gives everyone in the class last semester's Exam #4 (from Dr. Weaver's class). Then, because he would like to show that marijuana can hurt the IQ of even the poorest students, he identifies the 10 students who did worst on the exam. They got an average score of 15 items right (out of the 50 items on the exam). Then, every class afterwards, he put 10 small platters on a table in the front of the ISS305 classroom. On each platter was the name of one of the ten students, a large marijuana laced brownie, and a steaming cup of coffee. A sign over the table said "If you see your name on a platter, help yourself." At the end of the Spring 2015 semester he gets the scores of everyone on the regular Exam #2. Of the original ten students who had scored most poorly, the average of the 6 (who hadn't dropped the course) was 25. Ed reluctantly concluded that rather than hurting IQ, marijuana actually seemed to raise IQ. He said, "I guess I should have known that marijuana wasn't harmful since a majority of voters in Colorado and Washington just voted to make it legal."

Research report example #1:

Old issues:
- Empirical question?
- Reliable & valid measures?
- Misleading descriptions?
- Relationship?
- Fallacies?

New questions:
- Significant relationship?
- Any clear rival hypotheses?:
  – Threats to internal validity?
- Any likely problems with
  – Confounds with the IV?
  – Experimenter effects?
  – Demand characteristics?
  – Weak/Invalid manipulation?
  – Other?

Research report example #2:

Tobias Fünke, a famous actor and experimental psychologist, decided to conduct a study on helping behavior. He hypothesized that each sex would only help people of the opposite sex, not the same sex. Further, he suspected that help would only be given when the need is great. He hired two confederates, Ann Veal (an 18 year-old White Female undergraduate student) and Carl Weathers (a 65 year-old Black Male retiree). Participants were 1,040 White students. In the low-need condition one of the confederates would stagger out of a room with 5.2 ounces of ketchup on his/her shirt, screaming "Help!". In the high need condition, there was 6.2 ounces of ketchup displayed. Participants reported one at a time to the laboratory and took a seat in the waiting room. The sex of the confederate and his/her level of need was randomly determined. Resulting in 130 participants in each condition (2 [Sex of confederate] x 2 [Sex of P] x 2 [level of help]). The chosen confederate then staggered into the waiting room. The other confederate observed the scene through a one-way mirror and rated how vigorously the participant helped on an 11 point scale. These interactions were also videotaped and the confederates were asked to rerate all the scenes which they had rated before. The mean average test-retest correlation for the two confederates was .96. Fünke finds that Ps helped a male victim (mean = 7.5) significantly more than a female victim (mean = 7.4), but amount of help needed had no effect.

Research report example #2:

Old issues:
- Empirical question?
- Reliable & valid measures?
- Misleading descriptions?
- Relationship?
- Fallacies?

New questions:
- Significant relationship?
- Any clear rival hypotheses?:
  – Threats to internal validity?
- Any likely problems with
  – Confounds with the IV?
  – Experimenter effects?
  – Demand characteristics?
  – Weak/Invalid manipulation?
  – Other?