Internal Validity

- Measured & Manipulated Variables & Constants
- Causes, Effects, Controls & Confounds
- Components of Internal Validity
- "Creating" initial equivalence
- "Maintaining" ongoing equivalence
- Interrelationships between Internal Validity & External Validity

Internal Validity is about Causal Interpretability

Before we can discuss Internal Validity, we have to discuss different types of variables and review causal RH:s and the evidence needed to support them…

Every behavior/measure used in a research study is either a ...
Constant -- all the participants in the study have the same value on that behavior/measure
or a ...
Variable -- when at least some of the participants in the study have different values on that behavior/measure

and every behavior/measure is either …

Measured -- the value of that behavior/measure is obtained by observation or self-report of the participant (often called "subject constant/variable")

or it is …
Manipulated -- the value of that behavior/measure is controlled, delivered, determined, etc., by the researcher (often called "procedural constant/variable")

So, every behavior/measure in any study is one of four types….

	constant	variable
measured	measured (subject) constant	measured (subject) variable
manipulated	manipulated (procedural) constant	manipulated (procedural) variable

Identify each of the following (as one of the four above, duh!)…

- Participants reported practicing between 3 and 10 times
- All participants were given the same set of words to memorize
- Each participant reported they were a Psyc major
- Each participant was given either the "homicide" or the "self-defense" vignette to read

From before...

• Causal RH: -- differences in the amount or kind of one behavior cause/produce/create/change/etc. differences in amount or kind of the other behavior

Using our newly acquired language…

• Causal RH: -- the value of the variable manipulated by the researcher causes the value of the variable measured from the participant

In a causal research hypothesis…

• the manipulated variable = the "causal variable"

• the measured variable = the "effect variable," the "response variable" or the "outcome variable"

Be sure to notice -- The "**causal variable**" in the causal RH: absolutely must be **manipulated by the researcher** !!!!

Circle the manipulated/causal & underline measured/effect variable in each

• Practice improves performance.

• Treatment decreases depression.

• Schizophrenic symptomology is decreased by pharmacological intervention.

• Reading speed is improved by larger print size.

Try this one (you'll have to "figure out" what the manipulated variable is from the description of the different "conditions")

Completing the group therapy will lead to lower social anxiety scores than will completing the individual therapy.

manipulated variable --> Type of Therapy (group vs. individual)

measured variable --> Anxiety Score

Review of evidence required to support a causal research hypothesis …

Evidence needed to support a causal hypothesis...

- temporal precedence ("cause preceeds effect")
- demonstrate a statistical relationship
- elimination of alternative explanations (no other viable causes/explanations of the effect)

This identifies four different "roles" variables/constants might play in a study ….

Causal variable -- manipulated by the researcher -- the variable to which we want to attribute the effect

Effect variable -- measured from each participant after manipulation of causal variable by the researcher

Confounding variable(s) -- any variable (other than the one manipulated by the researcher) that might have caused the effect -- an alternative causal variable or explanation of the effect

Control -- any constant/variable that can't have caused the effect because it is "equivalent" across conditions

One of those things about “how we use words oddly”

We often talk about two kinds of variables – like this…

“Variables” – behaviors or characteristics of interest in the study

Variables – behaviors or characteristics for which different participants have different values

Constants – behaviors or characteristics for which all participants have the same value

Control Constants vs. Control Variables

Control Constants

- any behavior or characteristic for which all participants have the same value
- “a constant can’t be a confounding variable”

Control Variables

- any behavior or characteristic for which participants have different values, but for which the treatment or conditions are “balanced” or “equivalent” on that variable
- Examples
  - if ½ of the participants in each treatment/condition are male and ½ female, then gender is a control variable (note – don’t need a ½ - ½ split, only that the split is the same in each treatment/condition)
  - if the participants in each treatment/condition have the same average IQ, then IQ is a control variable

Control Constants, Control Variables & Confounds – some practice

80% of treatment group participants have prior experience with the task and 20% of the control group participants have prior task experience — confound

60% of treatment group participants have prior experience with the task and 60% of the control group participants have prior task experience — control variable

None of the participants in either group have prior task experience — control constants

All participants are 6 years old — control constants

The average age of the treatment group is 7 and the average age of the control group is 45. — confound

The average age of the treatment group is 7.1 and the average age of the control group is 7.2, — control variable

So, to summarize ...

Before the study begins...		After the study is over ...
Causal Variable	→	Causal Variable
Effect Variable	→	Effect Variable
Potential Confounds	→	(Control) Constants
	→	Control Variables
	→	Confounding Variables

We must take control of the potential confounds, so that they become controls and not confounds, if we are going to be able to causally interpret our research results.

Let's try using these terms …

RH: Computerized spelling practice leads to better performance than does paper & pencil practice.

Twenty English speaking 4th grade students were given 10 words and practiced them 5 times each on the computer. Twenty English speaking 2nd grade students were given the same 10 words and practiced them 3 times each using paper & pencil. When tested, the "computer practice" students did better than the "paper & pencil practice" students

What's the intended causal variable? Type of practice (comp.vs. pap&pen)

What's the intended effect variable? Test performance

Any control variables/constants? Is each measured or manipulated?
- English speaking – meas. const
- same words -- manip. const

Any confounds? Is each measured or manipulated ?
- grade -- measured
- # practices -- manipulated

So, can these results be used to support the causal RH: why or why not?

NO! We have temporal precedence, we have a statistical relationship, but we also have **confounds**, so we can't be sure what caused the effect

Here's another...

RH: Group therapy will lead to lower dep. scores than individual therapy

Five male & five female patients with no prior therapy completed a 24-session course of group therapy, meeting each time at the university psychiatric clinic. A different group of five male & five female patients patients, each of whom had previously received therapy for depression, completed a 10-session series of individual therapy, meeting at the same clinic. After the respective therapies, the group therapy patients had lower depression scores.

What's the intended causal variable? Type of therapy (grp vs. ind.)

What's the intended effect variable? Depression score

Any control variables/constants & is each measured or manipulated?
- Tx location -- manipulated const.
- gender -- measured var.

Any confounds & is each measured or manipulated ?
- # sessions -- manipulated
- prior therapy -- measured

So, can these results be used to support the causal RH: why or why not?

NO! We have temporal precedence, we have a statistical relationship, but we also have **confounds**, so we can't be sure what caused the effect

Components of Internal Validity

Initial Equivalence

- Prior to manipulation of the causal variable, participants in the different conditions are the same (on the average) on all measured/subject variables

Ongoing Equivalence

- during manipulation of the causal variable, completion of the task, and measurement of the effect variable, participants in the different conditions are the same (on the average) on all manipulated/procedural variables except the causal variable.

Practice with Types of Variables & Types of Equivalence

Tell the confounding variable, whether it is sub/msr or manip/proc and tell the type of equivalence that is at "risk" ...

I'm concerned that before the treatment began, those in the Drug Treatment group were more depressed than were those in the Therapy Treatment group.

Depression:
- Subject/Measured Variable
- Initial Equivalence

Are you sure that there was no problem allowing those in the Drug Treatment group to attend an extra 5 sessions ? Those in the Therapy Treatment group didn't have the extra sessions.

sessions:
- Manip./Procedural Variable
- Ongoing Equivalence

More practice ...

Tell the confounding variable, whether it is sub/msr or manip/proc and tell the type equivalence that is at "risk" ...

Because of the class schedule, those in the Computer Training Condition only had 20 minutes to take the test, while those in the Lecture Condition had 30 minutes.

Training time:
- Procedural Variable
- Ongoing Equivalence

To save time, only those who are familiar with computers were included in the Computer Training Condition, and everybody else was put in the Lecture Condition.

Familiarity:
- Subject Variable
- Initial Equivalence

From before -- using our new language

RH: Computerized spelling practice leads to better performance than does paper & pencil practice.

Twenty English speaking 4th grade students were given 10 words and practiced them 5 times each on the computer. Twenty English speaking 2nd grade students were given the same 10 words and practiced them 3 times each using paper & pencil. When tested the "computer practice" students did better than the "paper & pencil practice" students

We identified "grade" as a confound.

Does it mess up initial or ongoing equivalence & how do you know ??

initial equivalence -- it is a subject/measured variable

We identified "number of practices" as a confound.

Does it mess up initial or ongoing equivalence & how do you know ??

ongoing equivalence -- it is a manipulated/procedural variable

Another from before -- using our new language

RH: Group therapy will lead to lower dep. scores than individual therapy

Ten female patients with no prior therapy completed a 24-session course of group therapy, meeting each time at the university psychiatric clinic. Ten other female patients, each of whom had previously received therapy for depression, completed a 10-session series of individual therapy, meeting at the same clinic. After the respective therapies, the group therapy patients had lower depression scores.

We identified "# sessions" as a confound.

Does it mess up initial or ongoing equivalence & how do you know ??

ongoing equivalence -- it is a manipulated/procedural variable

We identified "prior therapy" as a confound.

Does it mess up initial or ongoing equivalence & how do you know ??

initial equivalence -- it is a subject/measured variable

How do we "produce" internal validity????

Important point -- we use different techniques to produce initial equivalence (of subject variables) and to produce ongoing equivalence (of procedural variables).

Initial equivalence of subject variables

- Random assignment of individual participants to treatment conditions before treatment begins

Ongoing equivalence of procedural variables

- Procedural standardization of manipulation, confound control, task completion, and performance measurement
- Darn it!!! There is no one "cure" for procedural confounds; they are avoided only by knowledge of their existence and diligent adherence to experimental procedures!

When are external and internal validity important???

External validity is obviously ALWAYS important! For any study we need to know to who, what, where & when it directly applies and "how far" it can be generalized!

You can find the argument that "internal validity is only important if you are testing causal RH:"… but consider this…

The more confounds you have, the less you learn from their being a statistical association between two variables, whether what you are trying to learn is associative or causal !!!

From which study will you learn more???

Study #1 Those who got more practices were also more motivated and were run during a different semester than those who got fewer practices

Study #2 Those who got more practices were equally motivated and were run during the same semester than those who got fewer practices

Whether you are testing a causal or an associative RH, the data from Study #2 is going to be easier to interpret!

The fewer confounds you have, the more you learn from their being a statistical association between two variables, whether what you are trying to learn is associative or causal !!!

Participant Assignment – "creating" initial equivalence

- "Who will be in what condition of the study, when?"
- Goal is to for participants in each condition of the study to be equivalent, on the average, before the manipulation of that condition begins
- Related type of validity is Internal validity - initial equivalence

Reminder:

- **Participant selection relates to the External Validity of the study (specifically → population validity)**

- **Participant assignment relates to the Internal Validity of the study (specifically → initial equivalence)**

Reminder about:
In Between Groups Designs
each participant will complete only one condition – assignment determines which condition for each participant

In Within-Groups Designs
each participant will complete all conditions -- assignment determines the condition order for each participant

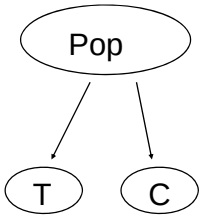
Acceptable Assignment Procedure for Causal RH:

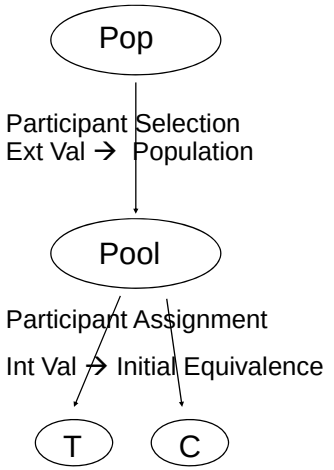
- Random Assignment of individuals by the researcher before manipulation of the IV
  - each participant has an equal chance of being in each condition of the study (BG) or each condition order (WG)
  - thus, all subject variables are "balanced" or "averaged out" across the conditions before manipulation begins
  - this is what gives us "initial equivalence" in a true experiment

Separating "Selection" & "Assignment"

A common representation of the participant acquisition process is shown below.

Folks are randomly chosen from the pop and placed into one of 2 groups.





What usually happens is shown above: First participants are selected into a "pool" and then are assigned into groups. Different aspects of validity are influenced by each step!!!

Unacceptable -- procedures that thwart testing Casual RH:

- Random assignment of groups (rather than individuals)
  - don't know that the groups were equivalent
- Arbitrary Assignment by the researcher
  - anything not using a "probabilistic" process -- might even be based on a "good idea" -- but isn't random
- Self Assignment by the participant
  - participant chooses what condition/order they will be in
- Administrative Assignment
  - non-random assignment determined by someone other than the researcher
- Non-Assignment or "Natural Assignment"
  - participant is already "in" conditions before they arrive at the study -- "causal variable" is really a subject variable

Problem with all of these?
For each of these there is a "reason" for why participants are in a particular condition/order -- that reason, and anything associated with it, produces a confounding of initial equivalence

Tell whether each is random, arbitrary, self, administrative or involves no assignment (were in "natural groups" before arriving to participate in the study)...

• after being presented with the options, each patient chose whether they would receive the "standard" or the "experimental" operation	Self
• the researcher decided that the first 20 participants would be assigned to the treatment condition, the rest would be assigned to the control	Arbitrary
• the Hospital Executive Committee determined that people who were over 60 years old would all receive the "standard" operation and all others would be randomly assigned to which operation they would receive	Admin
• medical records were examined to determine if the each participant had received the "standard" or "experimental" operation	None
• we randomly assigned 3 of the 6 laboratory sections to receive the "new" exam preparation assignment and the other 3 to receive the assignment we had been using for the last 2 years	RA of Intact groups
• whether each patient would receive the "standard" or "experimental" operation was determined by a coin-flip	RA

Random Assignment to Control Initial vs. Ongoing Equivalence

Randomly assigning individual participants to the conditions of a study (which condition for BG or condition order for WG) is used to control initial equivalence of subject variables.

- RA "ensures" that, on average, participants in the different conditions (BG) or different condition orders (WG) are the same "on average" on all subject variables

We also use random assignment to help control the ongoing equivalence of some procedural variables, for example…

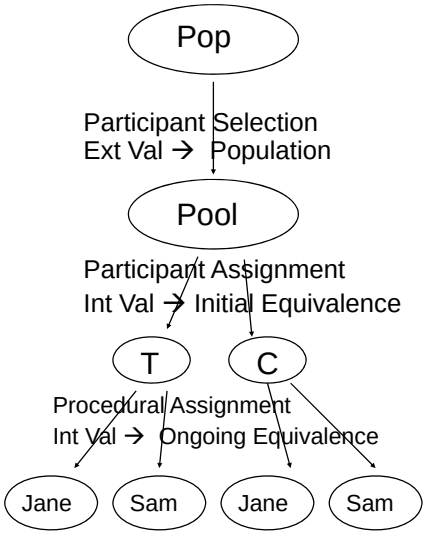
- if we have multiple research assistants – we should RA which research assistant runs each participant
  - researcher gender, age, appearance, race/ethnic & perceived comfort are all known to influence participant motivation, attention & performance !!!
- if we have multiple sets of instrumentation – we should RA which set is used for each participant

Separating Assignment for Initial & Ongoing Equivalence

So, the whole process often looks like this…

Multiple Procedural Assignment steps may be necessary:

Data collector, room, equipment, stimulus set, data coder, etc.



Tell whether each random assignment controls subject variables or procedural variables and whether the RA improves initial eq. or ongoing eq. …

IV → Type of Operation (experimental vs. standard)

•whether each patient would receive the "standard" or "experimental" operation was determined by a coin-flip — SV → initial

• we flipped another coin to decide which of four surgeons would perform the operation — PV → ongoing

IV → Modality (vision vs. touch)

•½ the participants were assigned to use the old stimulus set we've been using for years and ½ were assigned to use the new stimulus set we just had made this semester — PV → ongoing

• ½ the participants were randomly assigned to complete first the visual and then the touch conditions, while the other ½ completed the conditions in the opposite order — SV → initial

IV → Amount of practice (10 vs. 50)

•Jane ran a random ½ of the participants and Sam ran the other ½ — PV → ongoing

• whether the participant was given 10 practices or 50 practices was based the roll of a 6-sided die. — SV → initial

"Random assignment of individual participants by the researcher before manipulation of the causal variable" is the standard procedure to ensure initial equivalence of subject variables !

2 things to "look for" help you evaluate ongoing equivalence …

1. Research conducted in the field (outside the laboratory) is unlikely to have good control of ongoing equivalence
2. The longer the procedure (manipulation, task completion and data collection) the harder it is to maintain ongoing equivalence.

Which of the following is more likely to have good ongoing equivalence?

Laboratory study of practice effects comparing 5 & 10 practices.

Study of the effects of two different types of out-patient therapy.

Remember!! There is no "critical experiment"! A major reason for converging operations is because no one study will give us "sufficient" ongoing equivalence – we hope find similar results from multiple similar studies!

Finally ..The Relationship between Internal & External Validity

There are two different ways to think about the relationship between these two types of validity. Actually they are mutually exclusive, but we seem to alternate between using them both

- "Trade-off" characterization
  - it is impossible to promote both internal and external validity within a single study
  - the researcher must choose which will be emphasized in a particular study
    - internal validity (control)
    - external validity (representativeness)

- "Precursor" characterization
  - without causal interpretability (from having internal validity), what is there to generalize ???
  - focuses on causal information - suggesting associative information is not valuable