Research Designs

- Demonstrations vs. Comparisons
- Experimental & Non-Experimental Designs
- "IVs" and "DVs"
- Between Group vs. Within-Group Designs

There are two basic ways of **providing evidence to support a RH:** -- a "demonstration" and a "comparison"

- a demonstration involves using the treatment and showing that the results are "good"
- a comparison (an experiment) involves showing the difference between the results of the treatment and a "control"
- most commercials use demonstrations
  - They tell you how wonderful is the product!!
- The evidence from a demonstration usually meets with the response -- "Compared to what ??"
- a single demonstration is a "implicit" comparison
  - "doesn't this wash look better then yours ?"
  - "did you last heartburn improve this fast ?"
  - "didn't your last headache last longer than this ?"
- explicit comparisons are preferred (unless they suck, are false comparisons, confounded, or other damn lies we've seen) !!!

When testing causal RH: we must have a "fair comparison" or a "well-run Experiment" that provides

- init eq of subject variables & ongoing eq of procedural variables
- For example what if our experiment intended to show that Tide works better compared…

Really dirty light-colored clothes washed in a small amount of cold water for 5 minutes with a single rinse -- using Brand-X	vs.	Barely dirty dark-colored clothes washed in a large amount of hot water for 25 minutes with a double rinse -- using Tide

What is supposed to be the "causal variable" that produces the difference in the cleanness of the two loads of clothes?

Can you separate the initial and ongoing equivalence confounds ?

Initial Equivalence confounds

- "dirtyness" of clothes
- color of clothes

Ongoing Equivalence confounds

- amount of water
- water temperature
- length of washing
- single vs. double rinse

Research Designs

True Experiments

If "well-done," can be used to test causal RH: -- alternative hyp. are ruled out because there are no confounds !!!

True Experiment

- random assignment of individual participants by researcher before IV manip (provides initial equivalence - subject variables - internal validity)
- treatment/manipulation performed by researcher (provides temporal precedence & ongoing equivalence - internal validity)
- good control of procedural variables during task completion & DV measurement (provides ongoing equivalence - internal validity)

Non-Experiments

No version can be used to test causal RH: -- can't rule out alternative hyp. Because there are confounds !!

Quasi-Experiment

- no random assignment of individuals (but perhaps random assignment of intact groups)
- treatment/manipulation performed by researcher
- poor or no control of procedural variables during task, etc.

Natural Groups Design also called Concomitant Measures or Correlational Design

- no random assignment of individuals (already in "IV groups")
- no treatment manipulation performed by researcher (all variables are measured) -- a comparison among participants already in groups
- no control of procedural variables during task, etc.

Words of Caution About the terms "IVs", "DVs" & causal RH:s ...

You might have noticed that we've not yet used these terms..

- Instead we've talked about "causal variables" and "effect variables" -- as you probably remember..
  - the Independent Variable (IV) is the "causal variable"
  - the Dependent Variable (DV) is the "effect variable"

- However, from the last slide, you have know that we can only say the IV ***causes*** the DV if we have ***a true experiment (and the internal validity it provides)***
  - initial equivalence (control of subject variables)
    - random assignment of participants
  - ongoing equivalence (control of procedural variables)
    - experimenter manipulates IV, measures DV and controls all other procedural variables

The problem seems to come from there being at least three different meanings or uses of the term "IV" ...

1 "the variable manipulated by the researcher"

- it's the "IV" because it is "independent" of any naturally occurring contingencies or relationships between behaviors
- the researcher, and the researcher alone, determines the value of the IV for each participant

2 "the grouping, condition, or treatment variable"

3 "the presumed causal variable in the cause-effect relationship"

In these last two both the "IV" & "DV" might be measured !!! So…

- you don't have a True Experiment ...
  - no IV manipulation to provide temporal precedence
  - no random assignment to provide init. eq. for subject vars
  - no "control" to provide onging eq. for procedural variables
- … and can't test a causal RH:

This is important stuff -- so here's a different approach...

It is impossible to have sufficient internal validity to infer cause when studying some IV-DV relationships

Say we wanted to test the idea that attending private colleges CAUSES people to be more politically conservative than does attending public universities.

- We wouldn't be able to randomly assign folks to the type of college they attend (no initial eq.)
- We wouldn't be able to control all the other things that happen during those 4 years (no ongoing equivalence)

Here are some other categories of "IV"s with the same problem…

- orientation, age, # siblings
- ethnic background, race, neighborhood
- characteristics/behaviors of your parents
- things that happened earlier in your life

IVs "vs" Confounds

Both IVs and Confounds are "causal variables" !!!

- variables that may cause (influence, etc.) scores on the DVs

What's the difference ???

The IV is the intended causal variable in the study! We are trying to study if & how & how much the IV influences the DV !

A confound interferes with our ability to study the causal relationship between the IV & the DV, because it is another causal variable that might be influencing the DV.

If the IV difference between the conditions is confounded, then if there is a DV difference between the conditions, we don't know if that difference was caused by the IV, the confound or a combination of both !!!!

Between Groups vs. Within-Groups Designs

Between Groups

- also called Between Subjects or Cross-sectional
- each participant is in one (& only one) of the treatments/conditions
- different groups of participants are in each treatment/condition
- typically used to study "differences" -- when, in application, a participant will usually be in one treatment/condition or another

Within-Groups Designs

- also called Within-Subjects, Repeated Measures, or Longitudinal
- each participant is in all (every one) of the treatment/conditions
- one group of participants, each one in every treatment/condition
- typically used to study "changes" -- when, in application, a participant will usually be moving from one condition to another

Between Groups Design

Experimental Tx	Traditional Tx
Pat	Glen
Sam	Sally
Kim	Kishon
Lou	Phil
Todd	Rae
Bill	Kris

Different participants in each treatment/condition

Within-Groups Design

Experimental Tx	Traditional Tx
Pat	Pat
Sam	Sam
Kim	Kim
Lou	Lou
Todd	Todd
Bill	Bill

All participants in each treatment/condition

Research Designs

Putting this all together -- here's a summary of the four types of designs we'll be working with ...

	True Experiment • w/ "proper" RA/CB - init eqiv • manip of IV by researcher	Non-experiment • no or poor RA/CB • may have IV manip
Between Groups (dif parts. in each IV condition)	Results <u>might</u> be causally interpreted -- if good ongoing equivalence	Results <u>can not</u> be causally interpreted
Within-Groups (each part. in all IV conditions)	Results <u>might</u> be causally interpreted -- if good ongoing equivalence	Results <u>can not</u> be causally interpreted

Four versions of the same study … which is which?

- Each participant in our "object identification study" was asked to select whether they wanted to complete the "visual" or the "auditory" condition. — BG Non
- Each participant in our "object identification study" completed both the "visual" and the "auditory" conditions in a randomly chosen order for each participant. — WG Exp
- Each participant in our "object identification study" was randomly assigned to complete either the "visual" or the "auditory" condition. — BG Exp.
- Each participant in our "object identification study" completed first the "visual" and the the "auditory" condition. — WG Non

So, you gotta have a True Experiment for the results to be causally interpretable?

But, does running a "True Experiment" guarantee that the results will be causally interpretable?

What are the elements of a True Experiment??

Random Assignment if Individuals to IV conditions by the researcher before manipulation of the IV → Supposed to give us initial equivalence of measured/subject variables.

Manipulation of the IV by the researcher → Supposed to give us temporal precedence & help control ongoing equivalence of manipulated/procedural variables

Please note: A "true experiment" is defined by these two elements!
BUT → there is "an asymmetry" between "true exp" and "causal interp"
Huh? True Exp is necessary, but not sufficient, for causal interpretability!

What could possibly go wrong …. ???

Random Assignment "might not take"

- RA is a "probabilistic process" → there's no guarantee that the groups will be equivalent on all subject variables!

Might introduce a confound when doing the IV manipulation

- might treat the conditions differently other than the IV

May "miss" or even "cause" other ongoing equivalence confounds

- often, especially for younger researchers or newer research topics, we don't really know what to "control"
- we may know what to control and just not get it done…

If only True Experiments can be causally interpreted, why even bother running non-experiments?

1st Remember that we can't always run a true experiment !

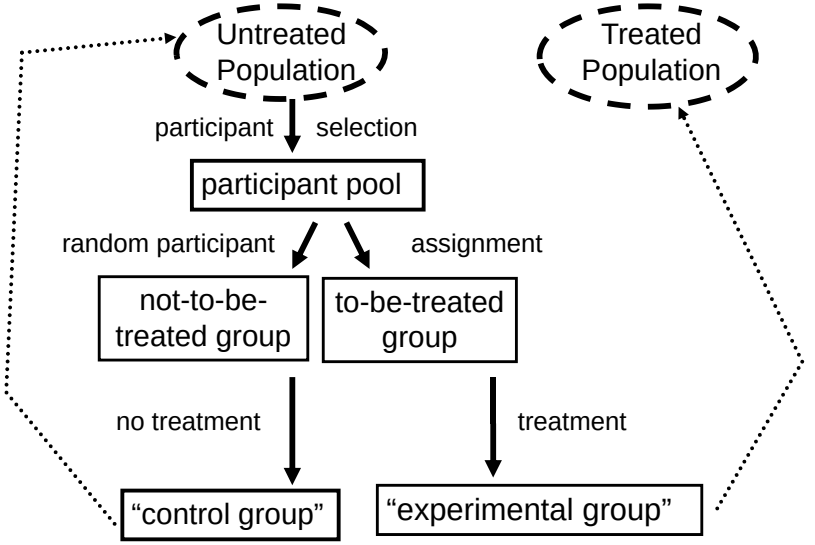
- Lots of variables we care about can't be RA & manip – gender, family background, histories and experiences, personality, etc.
- Even if we can RA & manip, lots of studies require long-term or field research that makes ongoing equivalence (also required for causal interp) very difficult or impossible.
- We would greatly limit the information we could learn about how variables are related to each other if we only ran studies that could be causally interpreted.

If only True Experiments can be causally interpreted, why even bother running non-experiments? Cont…

2nd We get very useful information from non-experiments !

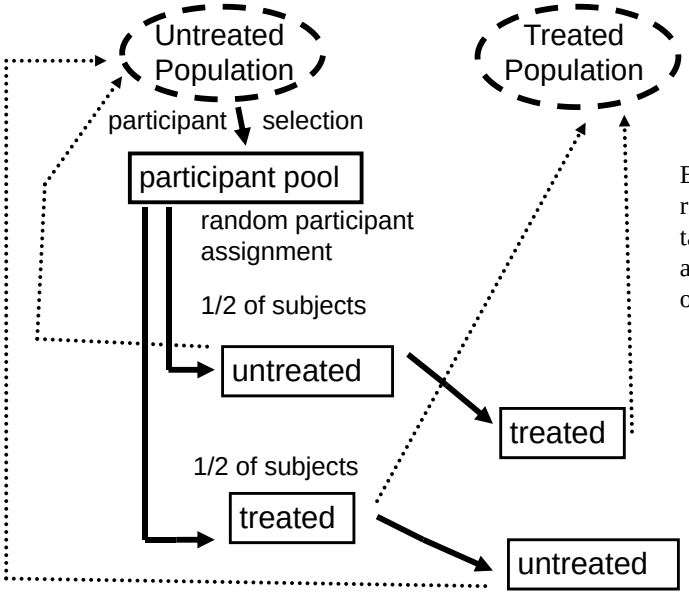
- True, if we don't run a True Experiment, we are limited to learning predictive information and testing associative RH:
- But associative information is the core of our understanding about what variables relate to each other and how they relate
- Most of the information we use in science, medicine, education, politics, and everyday decisions are based on only associative information – and things go pretty well!
- Also, designing and conducting True Experiments is made easier if we have a rich understanding of what variables are potential causes and confounds of the behavior we are studying

Between Groups True Experiment



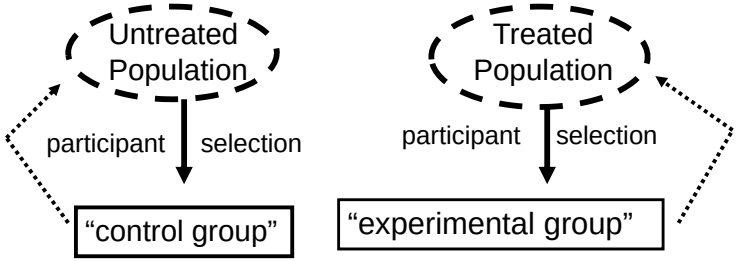
Rem -- samples & "groups" are intended to represent populatioins

Within-Groups True Experiment



Each participant represents each target population, in a counterbalanced order

Between Groups Non-experiment

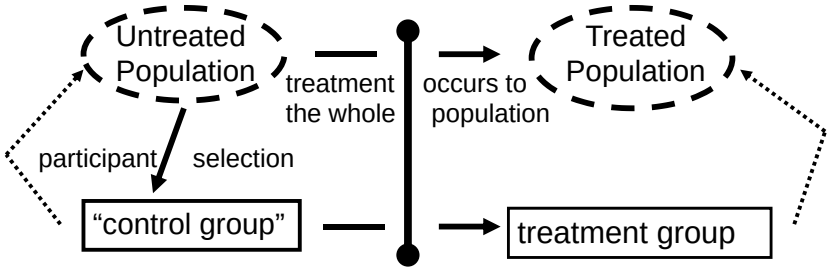


The design has the external validity advantage that each subject REALLY is a member of the population of interest (but we still need a representative sample)

The design has the internal validity disadvantages that ...

- we don't know how participants "end up" in the populations
  - no random participant assignment (no initial equivalence)
- we don't know how the populations differ in addition to the treatment per se
  - no control of procedural variables (no ongoing equivalence)

Within-Groups Non-experiment



The design has the external validity advantage that each subject REALLY is a member of each population of interest (but we still need a representative sample)

The design has the internal validity disadvantages that ...

- we don't know how the populations differ in addition to the treatment per se
  - no control of procedural variables (no ongoing equivalence)

There is always "just one more thing" ...

Sometimes there is no counterbalancing in a Within-groups design, but there can still be causal interpretation…

- A good example is when the IV is "amount of practice" with "10 practice" and a "50 practice" conditions.
  - There is no way a person can be in the 50 practice condition, and then be in the 10 practice condition
- Under these conditions (called a "seriated IV"), what matters is whether or not we can maintain "ongoing equivalence" so that the only reason for a change in performance would be the increased practice
  - The length of time involved is usually a very important consideration

Which of these would you be more comfortable giving a causal interpretation?

- When we gave folks an initial test, 10 practice and then the test again, we found that at their performance went up!
- When we gave folks an initial assessment, 6 months of once-a-week therapy and then the assessment again, their depression went down!