What defines an Experiment?

Independent Variable – a situational, task, or instructional variable that is completely under the control of the experimenter;

it is the potential cause of interest; some theory predicts that it does (or doesn't) determine the value of some other variable (the DV);

must take on at least two values, creating the conditions of the experiment

What defines an Experiment?

Dependent Variable – a labile, measured variable that some theory predicts to depend (or not) on the value of some other variable (the IV);

it is the effect of interest;

if at all possible, it should be quantitative

What defines an Experiment?

Extraneous Variable – any observable variable that might both influence the DV, and covary with the IV;

if both are true, then the EV is a confound;

to be "safe" we assume that all EVs can affect all DVs (unless or until there is evidence against this)

Confound – an EV that changes in parallel with the IV

Internal Validity

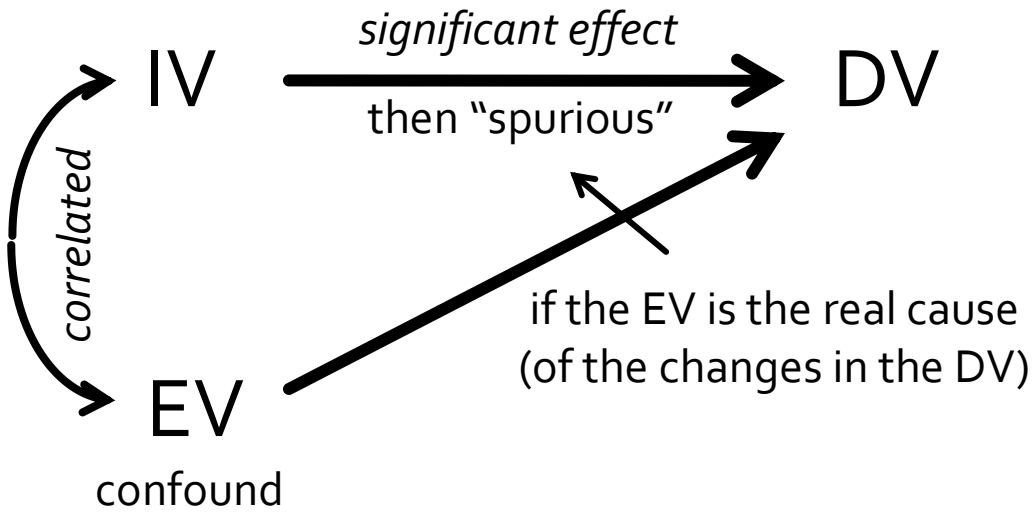
- Internal Validity (expts only) – *the extent to which the only observable & objective difference between the conditions of an experiment is that which defines the levels of the IV*

- Internal Validity (expts only) – *the extent to which there are no confounds*

- **please don't use these … they are just warm-ups**

Internal Validity

- Internal Validity (expts) – *the extent to which a significant (IV-DV) relationship is causal and not spurious*



Internal Validity

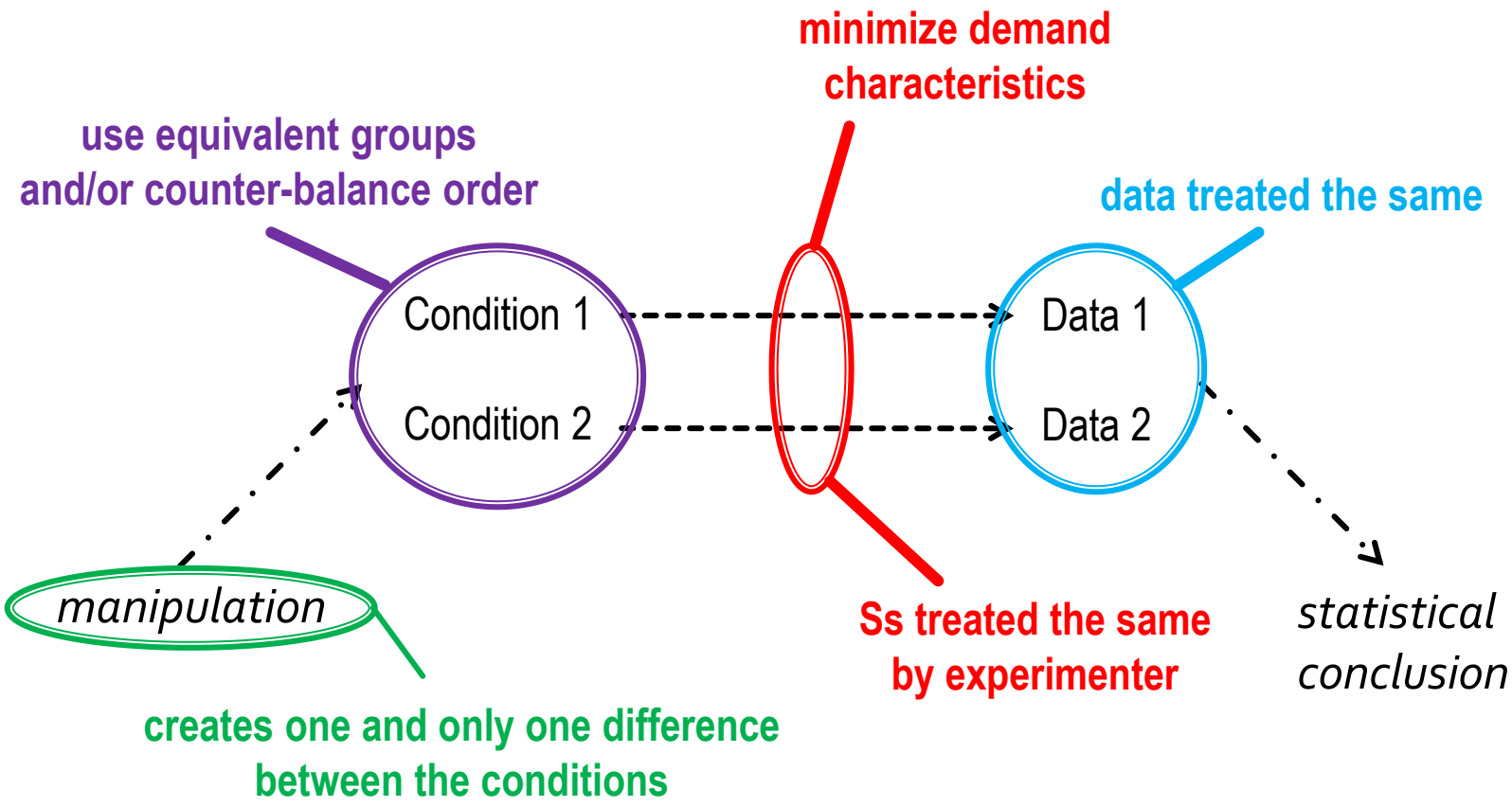
- Internal Validity (new; general) – *the extent to which the observed relationship between (two) variables is due to the causal relationship between the variables*

- **acceptable alternative**

Maintaining Internal Validity

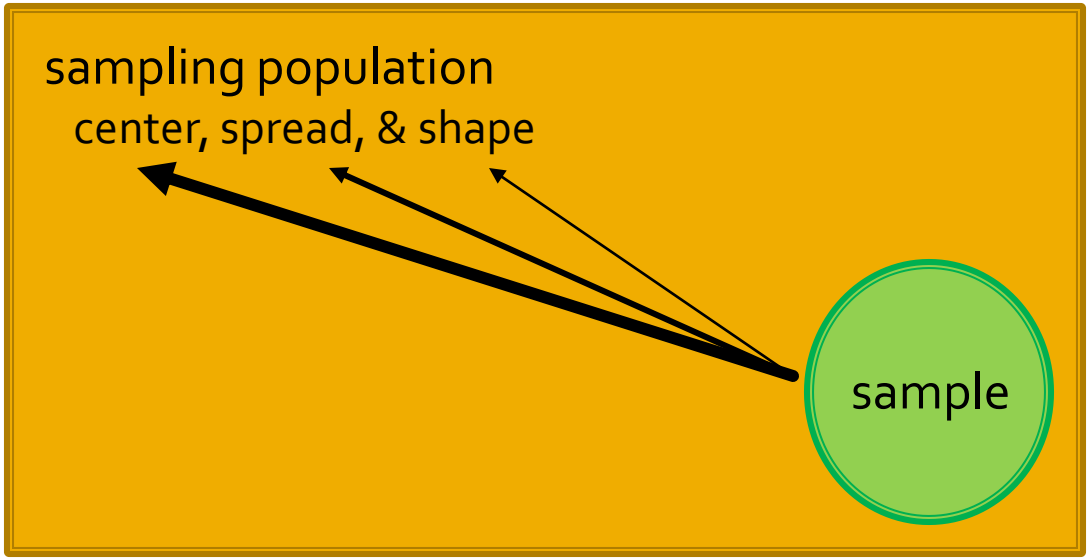
- all threats to IntVal are confounds (of some sort)
- the best way to approach this is in terms of what confounds arise at each step in the process
- and the best way to think about the steps in the process is in chronological order
- I'll assume you have already selected the IV and DV (which come from the theory being tested)
- I'll also assume that you have selected two levels of the IV for your experiment

Maintaining Internal Validity



StatsCon Validity

- Statistical Conclusion Validity – *the extent to which inferences <u>about the sampling population</u>, based on a sample, are accurate*



StatsCon Validity

- Sampling Population – *the set of people from whom the sample was taken; the set of people who could have been in the experiment*

  cf.

- Population – *everyone*

- StatsConVal only concerns the "jump" from the sample to the sampling population; a different kind of validity concerns the second jump from the sampling population to the entire population

StatsCon Validity

- all inferential stats start by making a "best guess"
- you should always attach an estimate of wrong you might be to any best guess

  $se = sd / \sqrt{N}$

- the above is also an estimate (or best guess); the quality of this estimate is given by the df
- in the second step, you boil this down to a yes-or-no answer to the question of whether the conditions produced significantly different results; yes = an effect of the IV on the DV *in the sampling pop*.

StatsCon Validity

what we conclude	what is true: No - the pop means are the same	what is true: Yes - the pop means are different
NO - the pop means are the same		"miss" (lack of) power β **Type II error**
YES - the pop means are different	**Type I error** "false-alarm" risk α	

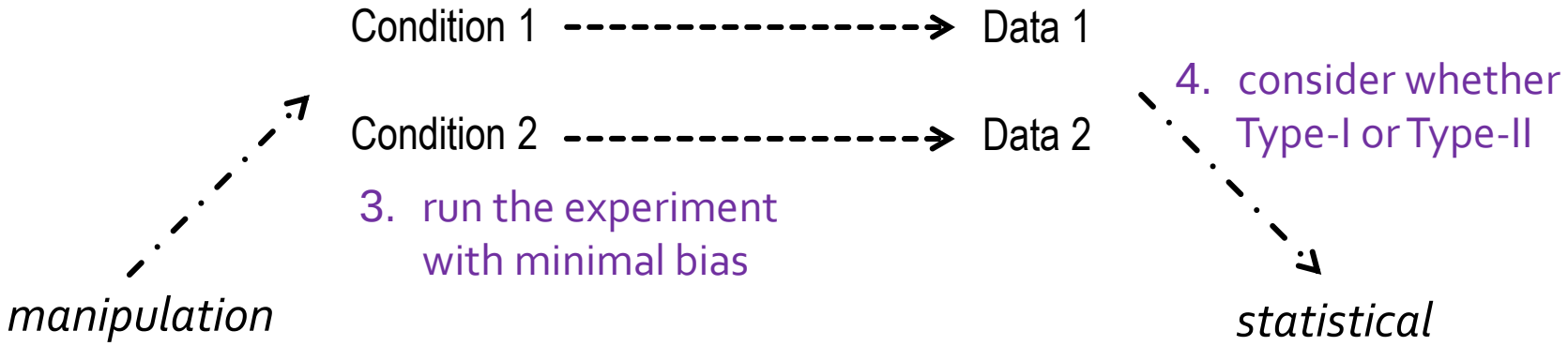
StatsCon Validity

- Type-I Error – *concluding in favor of a difference (in the sampling pop.) when none exists – aka "false alarm"*
- "Risk" – (α) *the probability of making a Type-I error (assuming that no difference actually exists)*
- Type-II Error – *concluding in favor of no difference (in the sampling pop.) when one exists – aka "miss"*
- "Power" – (1 – β) *the probability of **not** making a Type-II error (assuming that a difference actually exists)*

Planning & Running an Experiment

0. choose the IV and DV (which may require operationalizing the theory) and then choose (two or more) particular levels of the IV for conditions

2. choose a design type trading internal with stats-con validity



1. choose a method for creating the levels of the IV following the "control hierarchy" for each EV

1. Control Hierarchy (for IntVal)

- exert **control** – *do not allow the potential confound to vary at all*

- **pre-equalize** – *force the potential confound to be equal, on average, across conditions*

- **post-equalize** – *remove the effects of the potential confound after-the-fact*

- run a **control experiment** – *test the potential confound in a separate experiment*

2. Within- vs Between-Subjects

Favors Within-Subject

- **small effect**
- **unreliable measure**
- heterogeneous subjects; fear of failure of random assignment

When using, need to:

- counter-balance order

  balanced Latin Square

Favors Between-Subjects

- non-repeatable measure
- long-lasting manipulation
- need "vanilla" control cond
- **fear of demand chars**

When using, need to:

- create equivalent groups

  pseudo-random assignment + covariates, if still worried, or matching

3. Bias

- Experimenter Bias – *when beliefs and/or expectancies (conscious or otherwise) of the experimenter influence the results*
- Participant Bias – *when beliefs (conscious or otherwise) of the participant concerning how they <u>should</u> behave influence the results*
- Demand Characteristic – *any aspect of the experiment that indicates the purpose of the experiment*
- Evaluation Apprehension – *an internal state that causes subjects to alter their behavior so that they will be viewed more positively by other people*

4. Power (one last time)

- power depends on:

  the (absence of) "noise" in the measure

  the number of subjects run

  the design type employed

  the size of the actual effect

- think things through in many ways, looping back around to the beginning, etc.

Last-minute Questions

- 10 pm on Wed evening:

  http://www.justin.tv/directory/science_tech

  look for “Uipsymeth” stream

  if it asks for password: “exam2”