GETTING STARTED ON YOUR RESEARCH PROPOSAL

- **Getting an Idea**
  - read the classics - ask your mentor
  - mentor's goal should be to teach you *how*, not *that*
  - apprentice model
  - current literature <u>reviews</u> from leaders in the field
  - start lit review <u>early</u> (like yesterday)
  - start broadly
  - what are the practical problems in your field?
  - grant agency / foundations Requests for Proposals (RFPs)
  - think it up yourself
  - find a niche – fill it

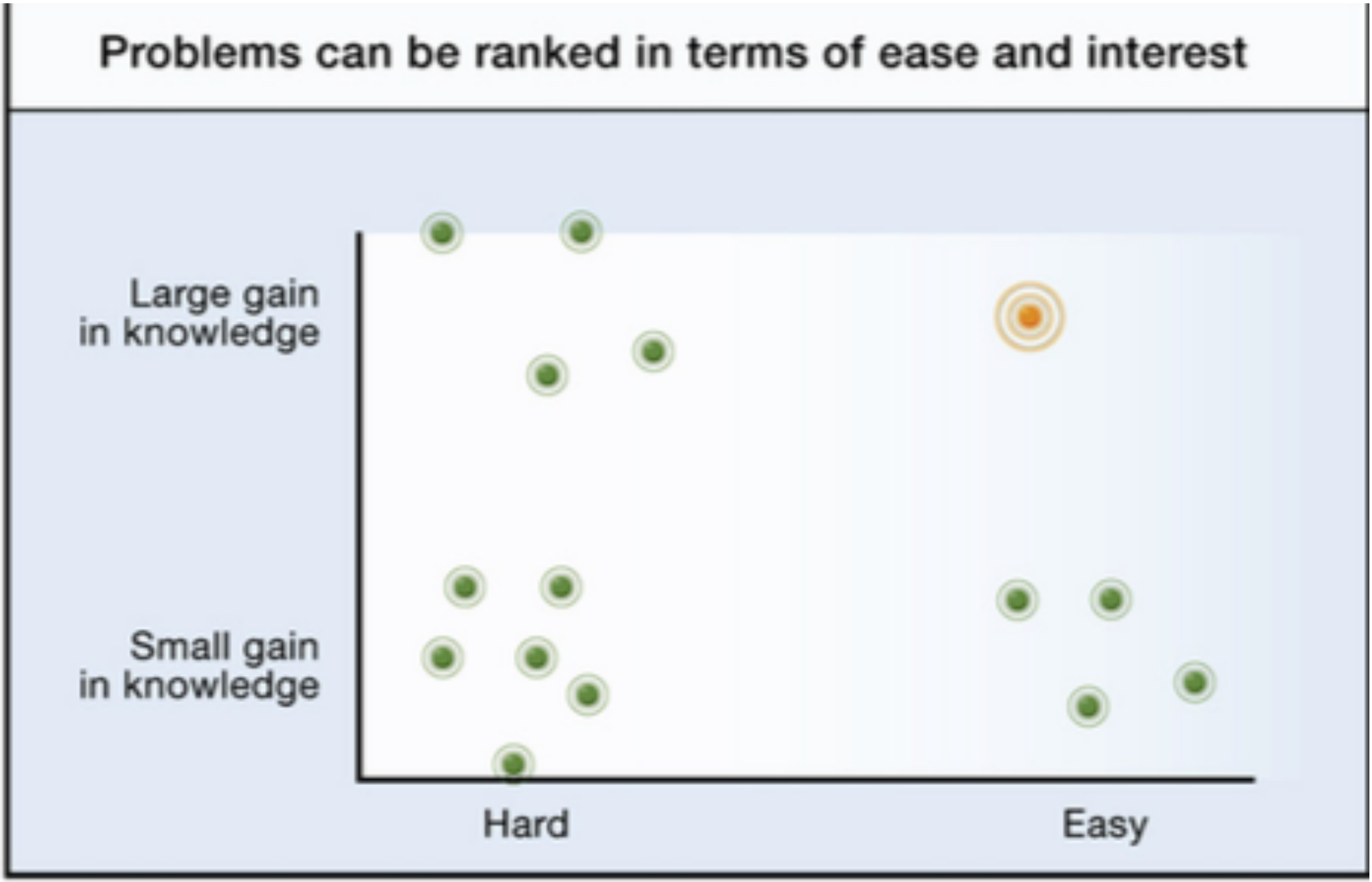
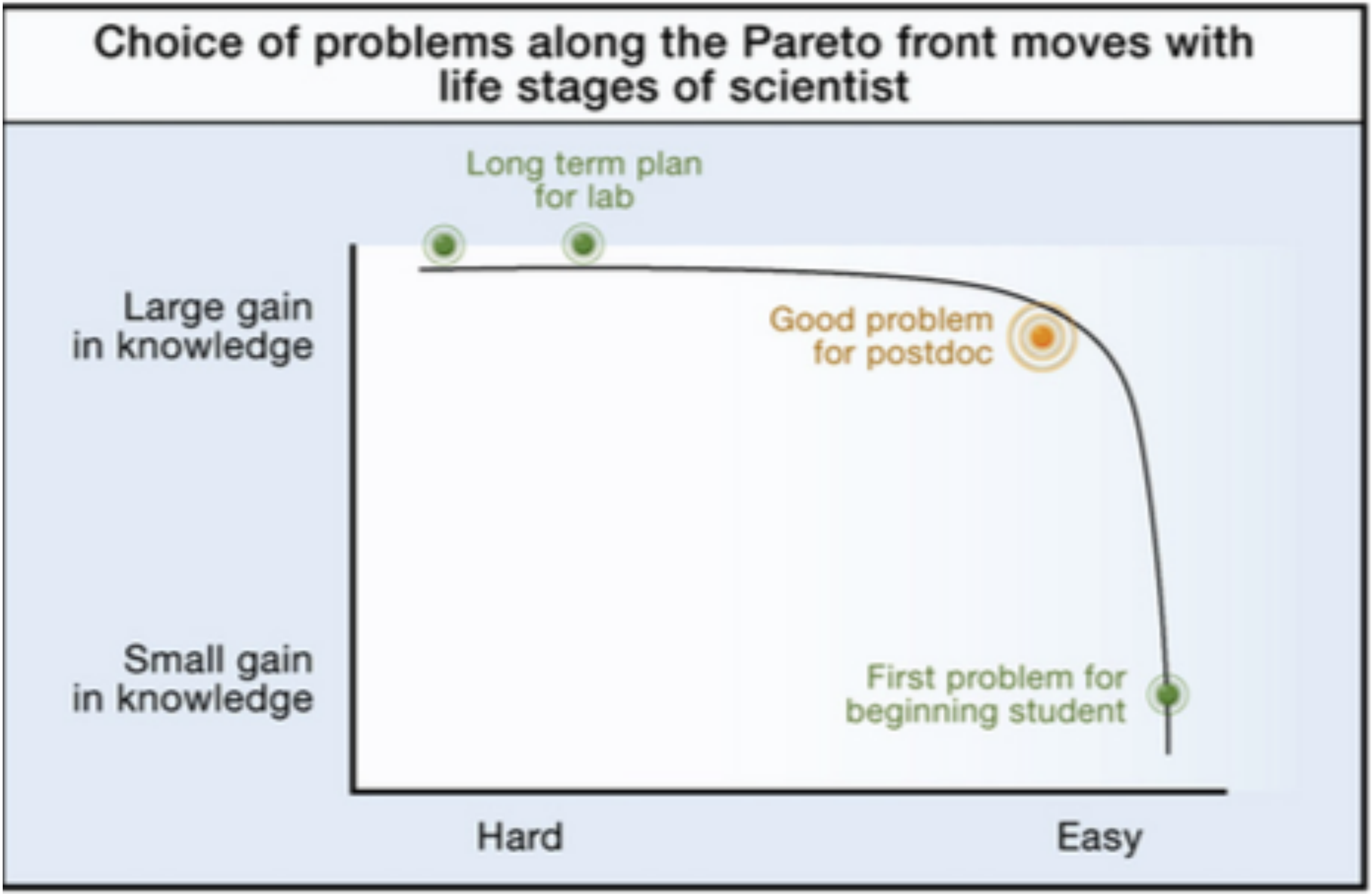


Figure 1. The Feasibility-Interest Diagram for Choosing a Project
Two axes for choosing scientific problems: feasibility and interest.


Choice of problems along the Pareto front moves with life stages of scientist
Long term plan for lab
Large gain in knowledge
Good problem for postdoc
Small gain in knowledge
First problem for beginning student
Hard
Easy

THE RESEARCH PROPOSAL

- **The Literature Review**

  - start lit review early and cast a broad net

  - use a reference manager

  - start with “review” papers written on topics similar to yours

  - read the interesting papers that the reviews cite

  - stay current - have weekly emails sent with current papers
    - journal websites
    - MyNCBI (pubmed.com)
    - societies
    - blogs

Generally, the purpose of a literature review is to analyze critically a segment of a published body of knowledge through summary, classification, and comparison of prior research studies, reviews of literature, and theoretical articles.

Advice for a Young Investigator

- no inquiry should be started without a thorough review of the literature
- how to read scientific papers:
  - 3 key pieces of information:
    - methods
    - data
    - unsolved problems
      - flee from abstracts as if from the plague
- we may learn a great deal from books, but we learn much more from the contemplation of nature

Writing the introduction

- Define the general topic, issue, or area of concern, providing appropriate context for the review
- Point out overall trends, or conflicts (theory, methodology, evidence, conclusions), or gaps in research, or a single problem / new perspective of immediate interest
- Establish the writer's point of view, explain the criteria to be used in analyzing and comparing literature and the organization of the review (sequence)

State why certain literature is or is not included (scope).

Writing the body

- Group studies according to common denominators such as qualitative versus quantitative approaches, conclusions of authors, specific purpose or objective, chronology, etc.
- Summarize individual studies or articles with as much or as little detail as each merits according to its comparative importance in the literature, remembering that space (length) denotes significance
- Provide the reader with strong "umbrella" sentences at beginnings of paragraphs, "signposts" throughout, and brief "so what" summary sentences at intermediate points in the review to aid in understanding comparisons and analyses

Writing the conclusion

- Summarize major contributions of significant studies to the body of knowledge under review, maintaining the focus established in the introduction.
- Evaluate the current "state of the art" for the body of knowledge reviewed, pointing out major methodological flaws or gaps in research, inconsistencies in theory and findings, and areas or issues pertinent to future study.
- Conclude by providing some insight into the relationship between the central topic of the literature review and a larger area of study such as a discipline, a scientific endeavor, or a profession

THE RESEARCH PROPOSAL

- **Planning the experiment:**

  - stay organized

  - get a lab notebook and write <u>everything</u> in it with pen (or have an electronic notebook with multiple backups)

  - keep the big picture in mind, but do small experiments

  - try to consider each experiment as a <u>publishable</u> project

  - develop an early reputation as a conscientious and thorough investigator

  - think critically about the studies of mentors, friends, co-workers and yourself

THE RESEARCH PROPOSAL

- **Planning the experiment:**

  - define the question
  - address a hypothesis by answering 1 or 2 specific questions

  - define your variables (“operationalize” your question)
    - independent variable - “groups”
    - dependent variable - “measurements”
  - controls – ideally for every variable
    - positive / negative

  - determine statistical analysis

  - determine *power* needs (number of samples needed)

THE RESEARCH PROPOSAL

- **Planning the experiment:**
  - determine $ needs
  - determine source of funding
    - who is interested in your idea?
      - many private foundations, gov, etc
        - e.g., American Heart Association
        - www.apa.org/students/funding.html

THE RESEARCH PROPOSAL

Feasibility:

- scientific rigor vs. practicality
- ethical constraints (IRB/IACUC)
- costs
- time to complete
- necessary cooperation
    - collaborate – but be selective

THE RESEARCH PROPOSAL

- **Proposal**
- Tell an interesting and intriguing story
- Be clear, then try to be even more clear
- Organize, then reorganize

THE RESEARCH PROPOSAL

- **To test effects, you need to set up the experiment:**

  - obtain or prepare *protocol(s)*
  - mentor, published papers, etc
  - tweak to suit your needs

  - collect data

  - analyze data - VISUALLY

  - interpret data
  - do your colleagues agree with your analysis?

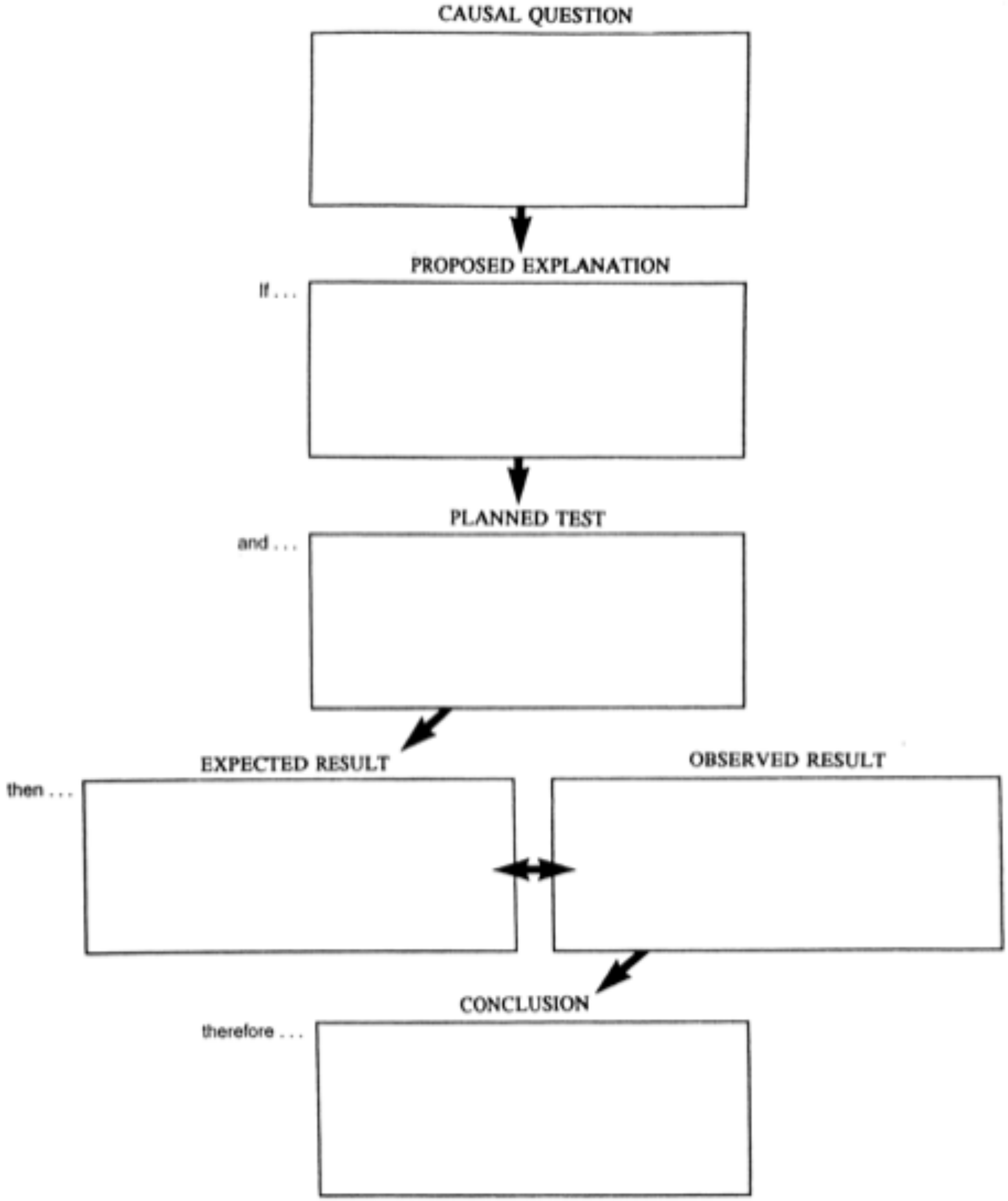
  - repeat (for the rest of your life….)

THE RESEARCH PROPOSAL

- Write it up

- 2 rules:
- keep it simple & clear
- keep it clear & simple

- PUBLISH IT!



CAUSAL QUESTION
PROPOSED EXPLANATION
If . . .
PLANNED TEST
and . . .
EXPECTED RESULT
then . . .
OBSERVED RESULT
CONCLUSION
therefore . . .

CVs and biosketches - things I've noticed:

- Too Many Capitals
- SOAFTU (spell out acronyms first time used)
- Boldface your own **name** in author lists
- Do not include "non-career" oriented work experience
- Try not to repeat. Try not to repeat
- Active voice was not used
- Bullet-point experience, etc., usually with full sentences (including periods)
- Don't use contractions
- Do not include classes taken
- Personal statements should explain why / how you are qualified to participate (including your lab's qualifications)
- Use FULL WORDS (e.g., undergraduate, laboratory)

Psychology 505:
Advanced Research Methods
Rich Hartman

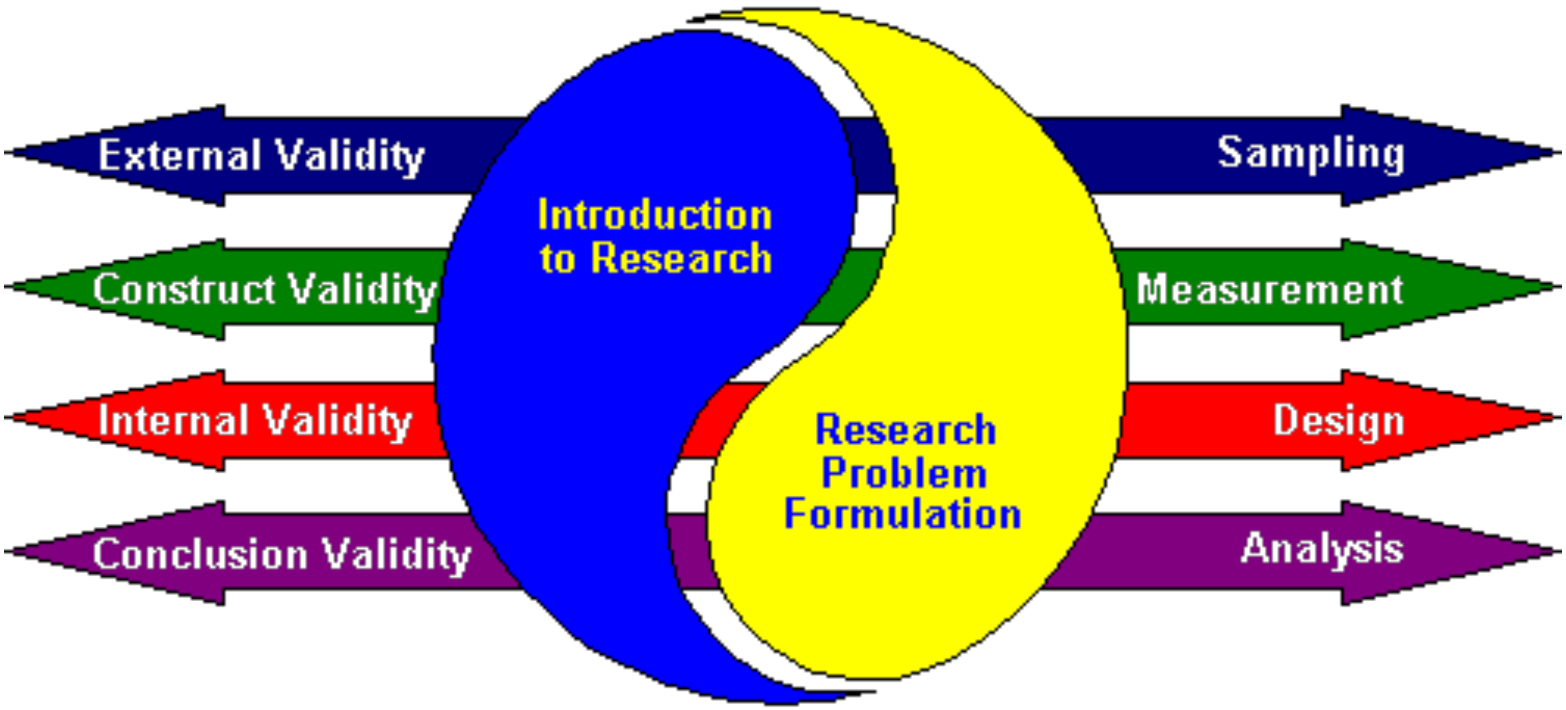


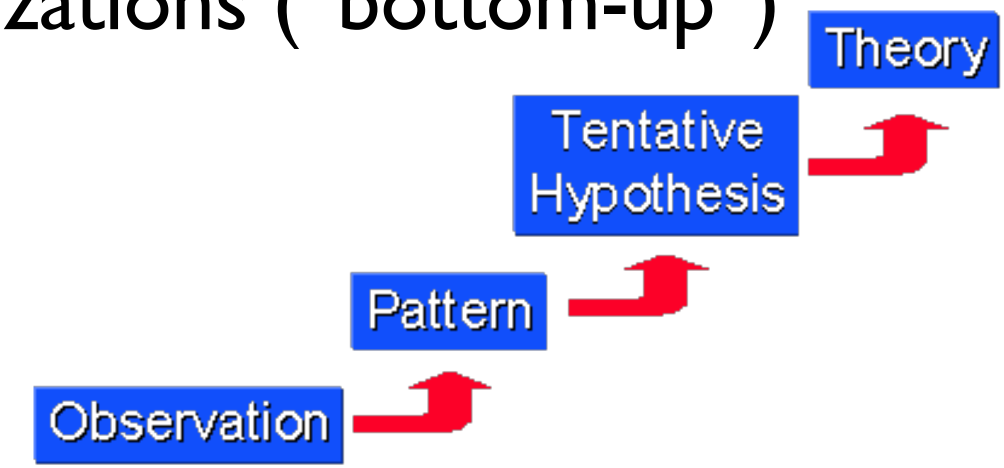
Image taken from
www.socialresearchmethods.net

Some More Advice for a Young Investigator

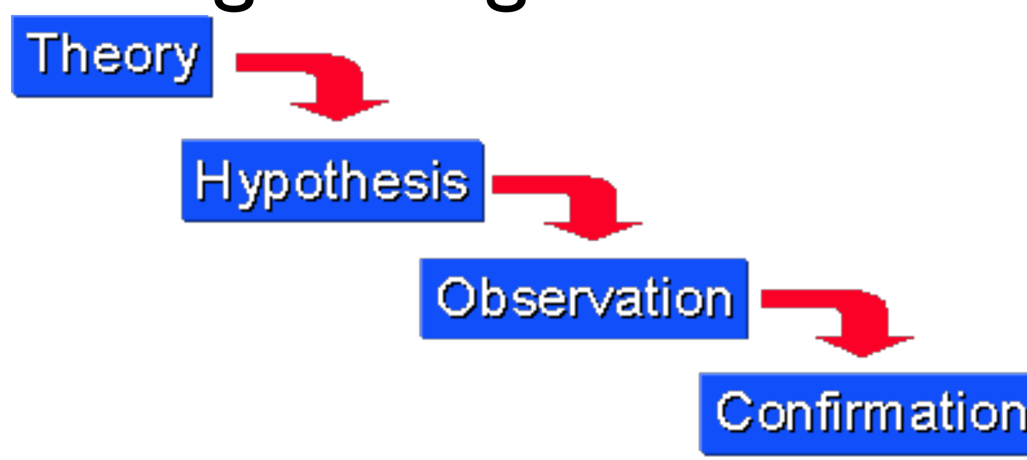
- **Stages of Scientific Research:**
  - **observation** - look at things as if for the very first time. Don’t just examine - observe and reflect. Use the best possible tools and most accurate techniques possible.
  - **experimentation** - allows time scale of observation to be shortened
  - **hypotheses** - “He who refuses to accept the hypothesis as a guide is resigned to accept chance as a master.”

Induction and Deduction

- <u>Inductive</u> reasoning - working from specific observations to broad generalizations ("bottom-up")
  - open-ended / exploratory



- <u>Deductive</u> reasoning - working from general to specific ("top-down")
  - narrow in nature

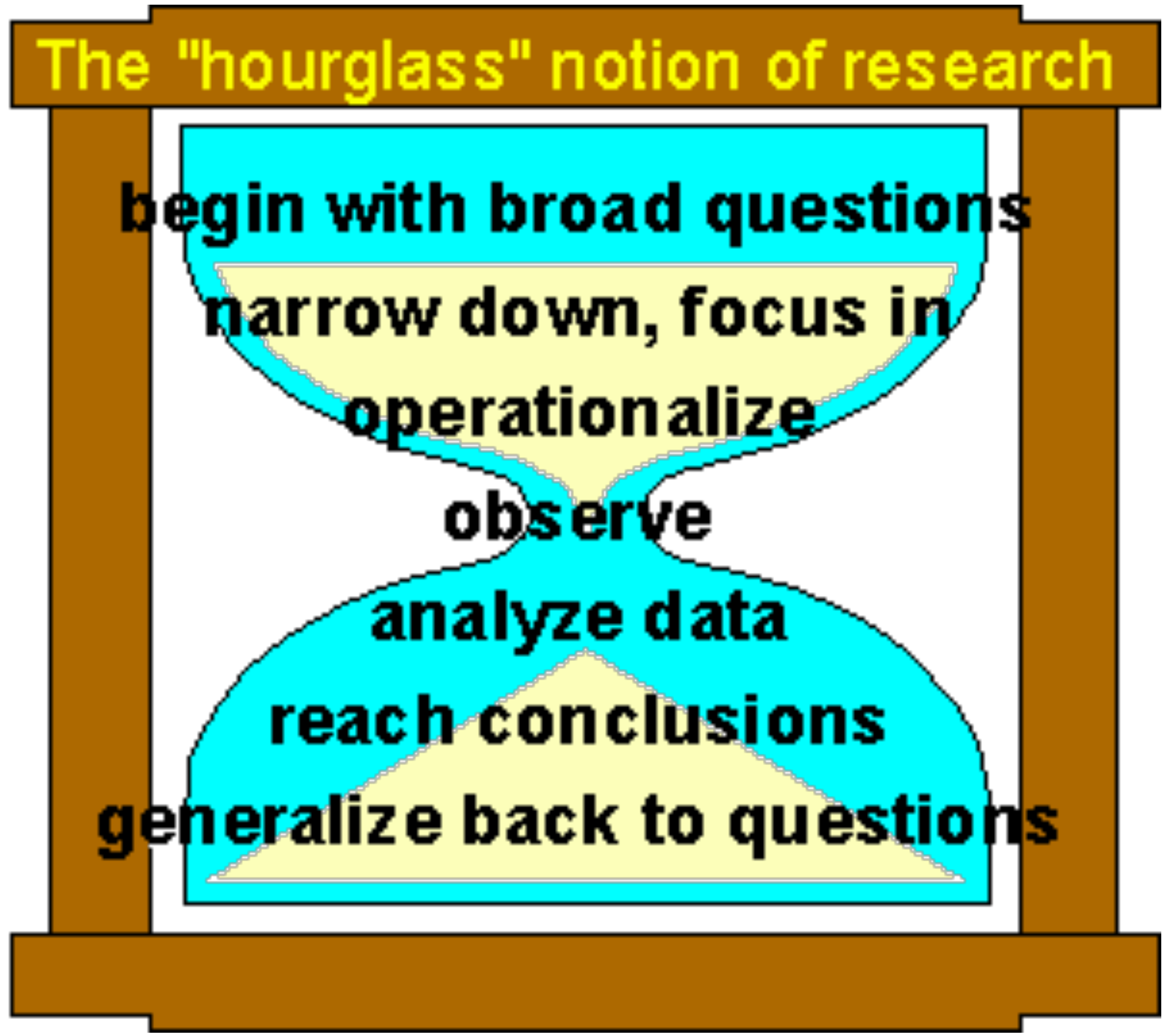


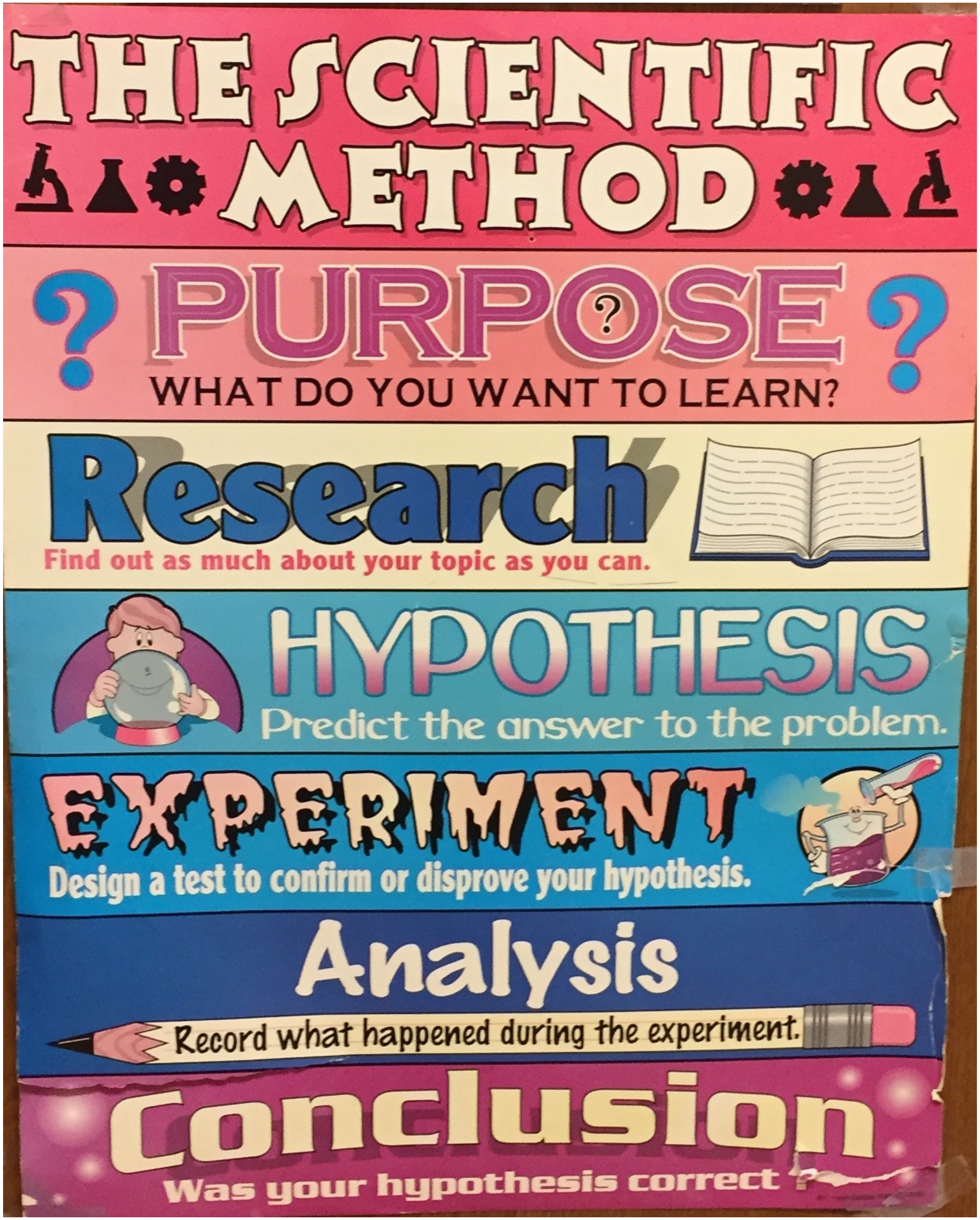
Types of Questions

- 3 basic research questions (*from least to most "demanding"*)
  - <u>descriptive</u> - a study designed to describe what goes on or what exists
    - polls, development of spatial learning in mice, etc
  - <u>relational</u> - a study designed to assess relationships between 2 or more variables
    - analyzing poll data by gender, is hippocampal cell count related to L/M abilities, etc
  - <u>causal</u> - a study designed to determine whether a variable causes or affects another variable
    - analyzing poll data to determine whether an ad campaign affected voting
    - does feeding transgenic mice pomegranate juice prevent the buildup of amyloid plaques in their brains?

Structure of Research

- Major components of a causal study:
  - the problem
  - the question
  - the cause (IV)
  - the units of analysis
  - the effect (DV)
  - design




THE SCIENTIFIC METHOD
PURPOSE
WHAT DO YOU WANT TO LEARN?
Research
Find out as much about your topic as you can.
HYPOTHESIS
Predict the answer to the problem.
EXPERIMENT
Design a test to confirm or disprove your hypothesis.
Analysis
Record what happened during the experiment.
Conclusion
Was your hypothesis correct ?

Research

- Science is <u>empirical</u> - based on observations and measurements of reality
- ultimate goal is to uncover cause-effect relationships (as opposed to “descriptive” or “relational” studies)

Variables

- variable - any "entity" that can take on different <u>values</u>
  - age, country, gender
    - IQ (multiple)
  - attribute - gender, scales, etc
    - must be *exhaustive* and *mutually exclusive*
- dependent vs. independent (for causal studies)
  - independent (IV) - manipulated (by you or "nature")
  - dependent (DV) - what is affected by the IV

- the *unit of analysis* is the major entity that is being studied:
  - individuals, groups, brain cells, etc.

Types of Data

- Qualitative vs. Quantitative
  - *quantitative* - in a numerical form
  - *qualitative* - <u>not</u> in a numerical form
    - descriptive, etc.
  - ALL QUANTITATIVE DATA IS BASED UPON QUALITATIVE JUDGEMENTS, AND ALL QUALITATIVE DATA CAN BE DESCRIBED AND MANIPULATED NUMERICALLY

Time in Research

- cross-sectional vs. longitudinal studies
    - <u>cross-sectional</u> - takes place at 1 point in time
        - slice / cross-section of the variable
    - <u>longitudinal</u> - at least 2 rounds of measurements over time
        - repeated-measures vs. time series

Hypotheses

- hypothesis - a specific statement of *prediction* (causal)
  - concrete terms - what you expect to happen
  - a study may have 1 or more hypotheses
  - *exploratory* studies may not have a hypothesis
- *null hypothesis (H0)* - usually, there is no relationship between the variables
- *alternative hypothesis (H1)* - there is a relationship
- actually, H0 should encompass all other alternatives

- test your hypothes(es) with 1 or 2 SPECIFIC AIMS

Hypotheses

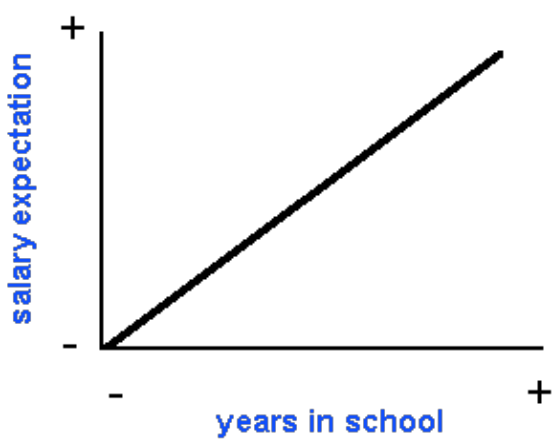
- Hypothetical-Deductive Model:
  - formulate your prediction (directional or not)
  - formulate a 2nd hypothesis that is *mutually exclusive* of the first, and encompasses all possible other outcomes
    - this process exhausts all possible outcomes
  - at the end of a study, you pick one or the other

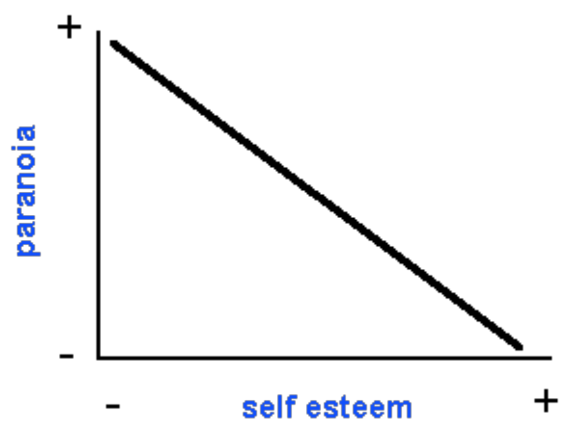
Advice for a Young Investigator

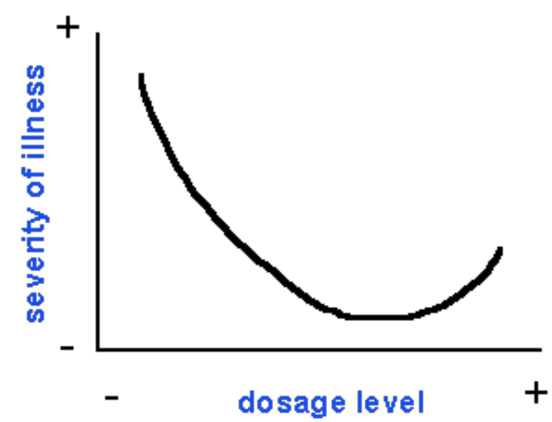
- observe with methods that are clear and precise
- create new conditions with experimentation
- critique and eliminate erroneous hypotheses
- prove the hypothesis with new observations or experiments
- apply the implications of the hypothesis to other spheres of knowledge

Types of Relationships

- relationship - “correspondence” between 2 variables
  - nature vs. pattern:
    - nature - correlational vs. causal
      - “3rd variable” problem (be careful interpreting relationships)
    - patterns - positive / negative / curvilinear

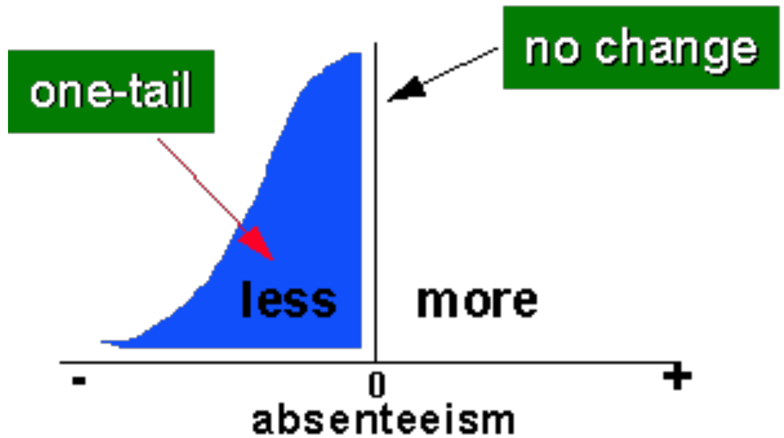




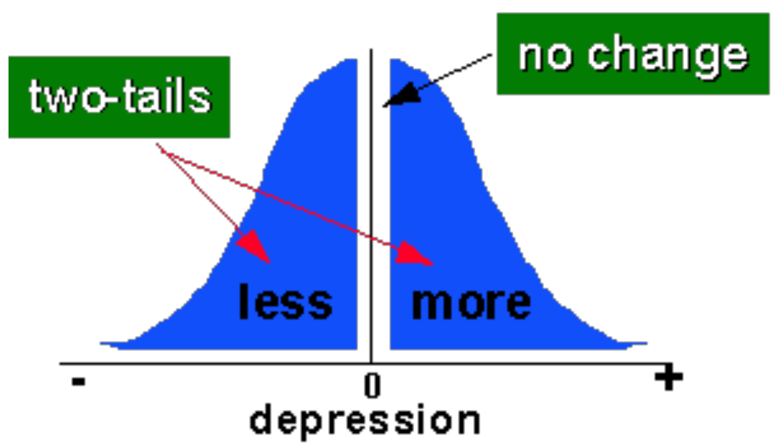


Hypotheses

- one-tailed hypothesis: H1 is *directional*
  - "tail" refers to the tail of the DV distribution



- two-tailed hypothesis: H1 is *non-directional*



Advice for a Young Investigator

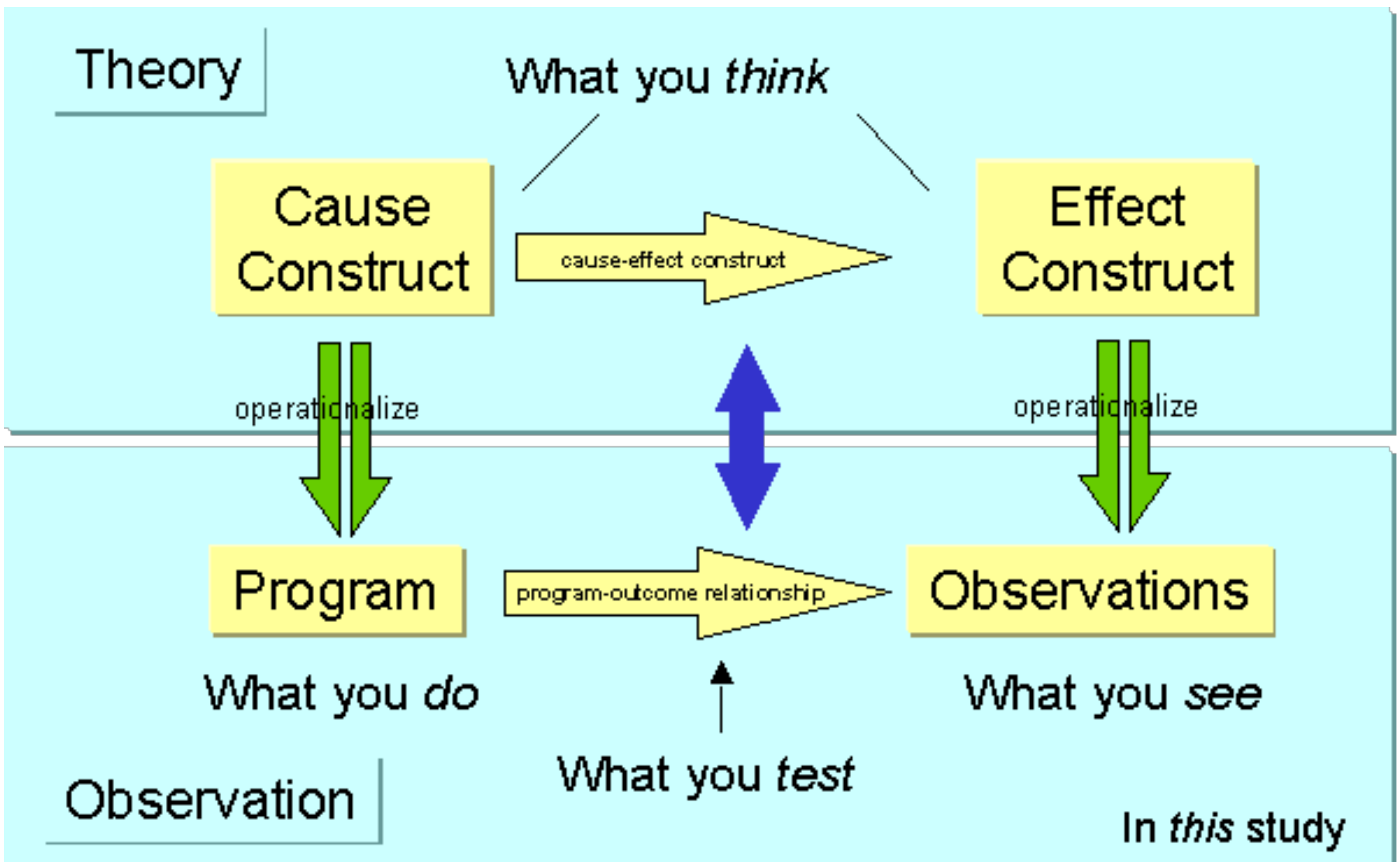
- hypotheses are necessary - without them a phenomenon cannot be explained
- hypotheses that cannot be tested (falsified) leave a problem un-illuminated
- hypotheses should be easily understood, and should lie in the realm of pure mechanism
- hypotheses solve quantitative problems in a qualitative way, avoiding metaphysical considerations
- hypotheses should suggest new research and/or arguments
- how to formulate hypotheses:
  - nature uses the same means for equivalent ends - compare phenomena that you do not understand with those that you do
  - view problems in their simplest forms
  - proof - submit your hypothesis to the ratification of testing. if it does not fit the data, reject it mercilessly.
  - subject yourself to the harshest of self-criticism

2 Important Research Fallacies

- fallacy - an error in reasoning, usually based on mistaken assumptions
  - ecological - making conclusions about individuals (*ideopathic*) based on analyses of group data (*nomothetic*)
  - exception - making conclusions about groups (*nomothetic*) based on analyses of individual data (*ideopathic*)

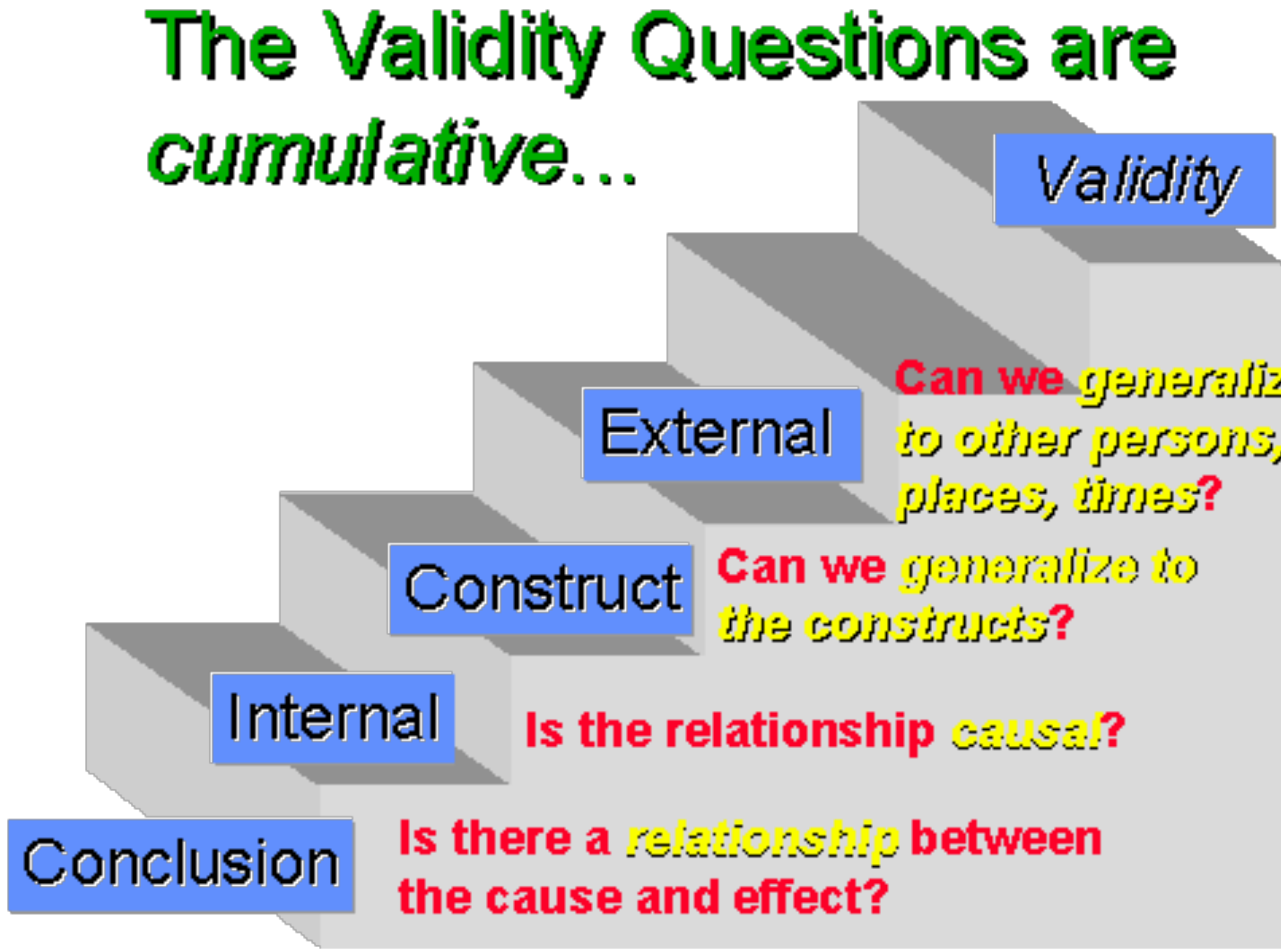
Intro to Validity

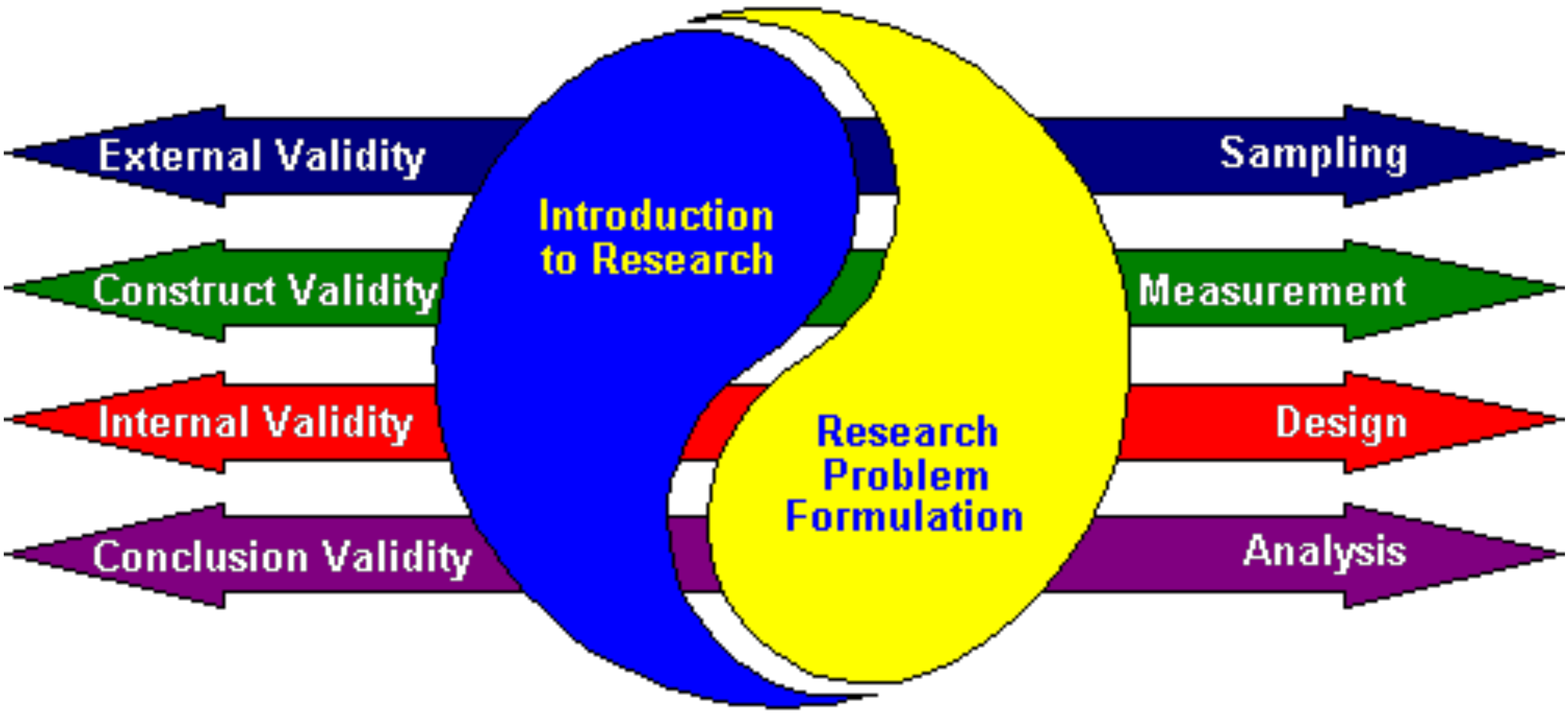
- <u>validity</u> - quality of various conclusions you might reach based on a research project
    - the best available approximation to the truth of a given proposition, inference, or conclusion



Intro to Validity

- 4 types of validity (cumulative):
  - external (theory side)
  - construct (links observation with theory)
  - conclusion & internal (observation side)




External Validity
Construct Validity
Internal Validity
Conclusion Validity
Introduction to Research
Research Problem Formulation
Sampling
Measurement
Design
Analysis

Threats to Validity

- you want to reduce the plausibility of the most likely threats to validity, leaving your conclusion as the most plausible
  - your study will have greater conclusion validity if you can show that these alternative explanations are not credible:
    - insufficient statistical power to detect a relationship even if it exists
      - not enough samples / sample size is too small
    - too much variability in the data

Dealing with External Validity - Sampling

goal is to fairly generalize the results back to the population

sampling - the process of selecting units from a population of interest

Dealing with Construct Validity - Measurement

measurement - the process of observing and recording the observations

fundamental ideas:

- levels of measurement (nominal, ordinal, interval, ratio)
- reliability (true score theory, reliability estimators)

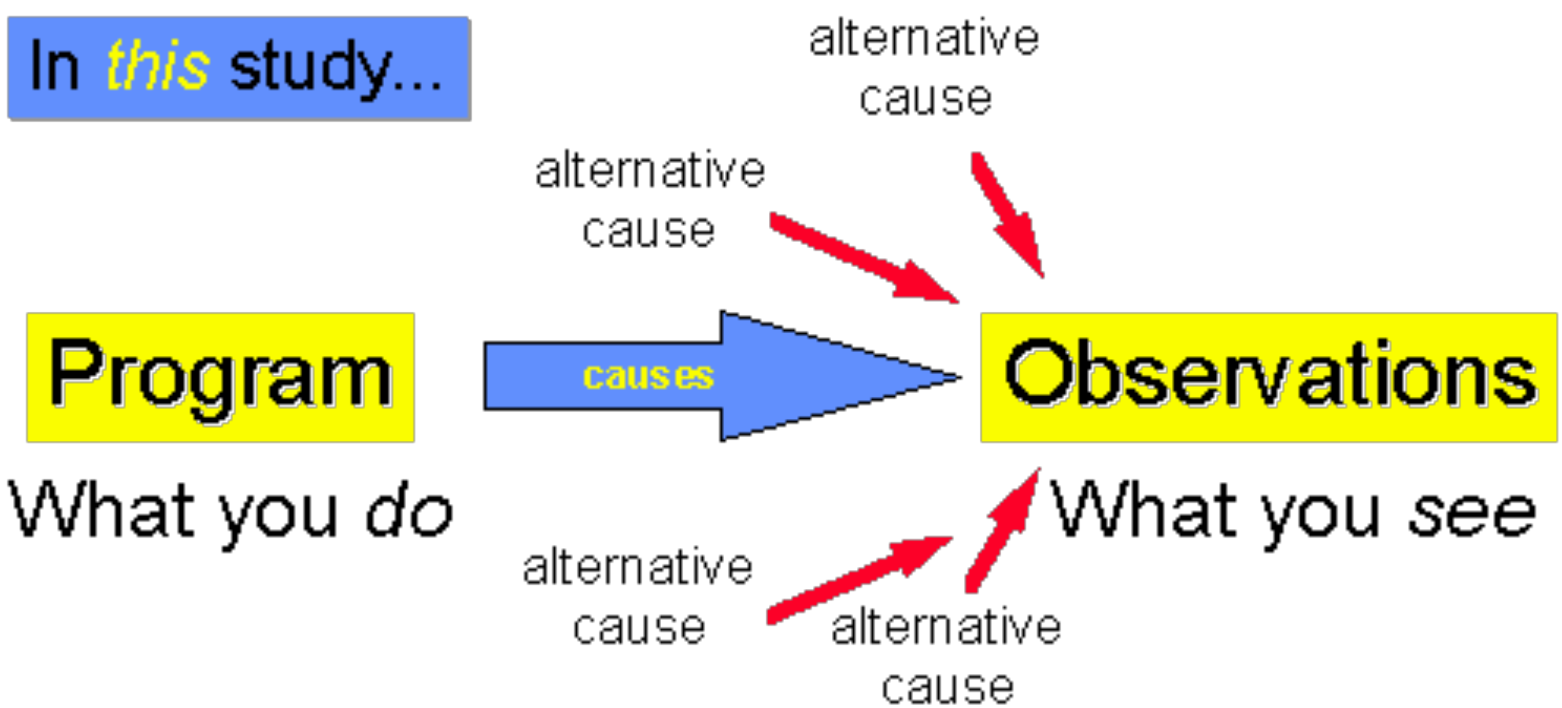
types of measures:

- survey
- scaling
- qualitative
- unobtrusive measures

Dealing with Internal Validity - Experimental Design

Internal validity - are the “cause-effect” inferences in your study valid?

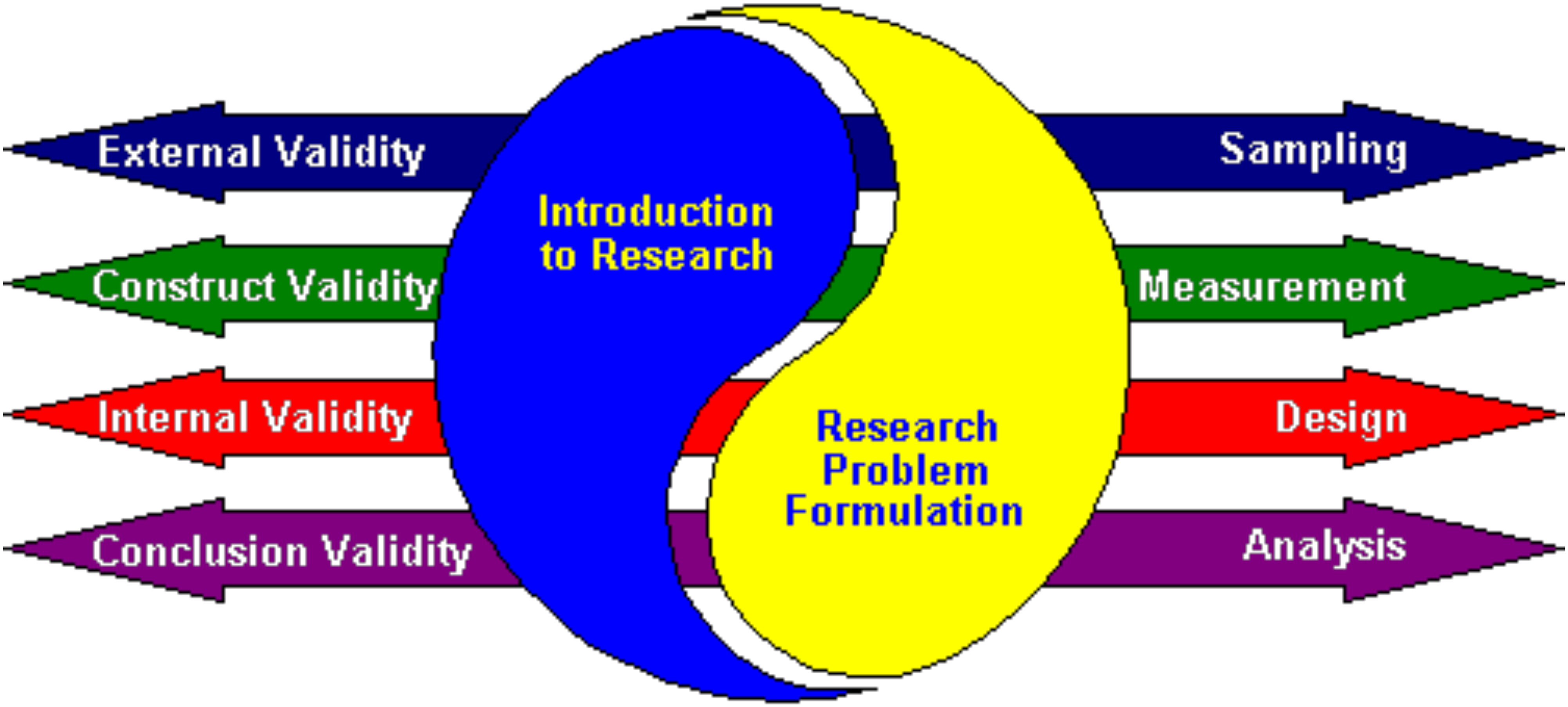
did your treatment / program / independent variable actually have an effect, or are there alternative explanations?



Dealing with Conclusion Validity - Analysis

- **Conclusion Validity - the degree to which conclusions we reach about relationships in our data are reasonable.**
  - **the credibility of your conclusion on the relationship that you found between your IV and DV**
  - **also - *"statistical" conclusion validity***

Dealing with External Validity - Sampling



Dealing with External Validity - Sampling

goal is to fairly generalize the results back to the population

sampling - the process of selecting units from a population of interest

Dealing with External Validity - Sampling

validity refers to the approximate truth of propositions, inferences, or conclusions

external validity is related to generalizing

the degree to which a study's conclusions hold true for other people, places, and times

Dealing with External Validity - Sampling

Sampling:

- identify population you would like to generalize to
- draw a fair sample and do the research
- because the sample is representative, results are generalizable back to the population
  - Problems:
    - only generalizable back to the original population
    - not easy to draw a truly representative sample
    - impossible to sample across all times (e.g., future)

Dealing with External Validity - Sampling

- 3 major threats to external validity:
  - wrong subjects, places, or times

- Improving external validity:
  - do a good job of sampling from your population
    - use random selection (different from random assignment)
    - replicate - use a variety of people, places, or times

Sampling - Terminology

Major question should be: "To whom do you want to generalize?"

- the whole population - probably not just those in your study
  - theoretical population - the population that you would like to generalize to
  - accessible population - those that are actually available

Sampling - Terminology

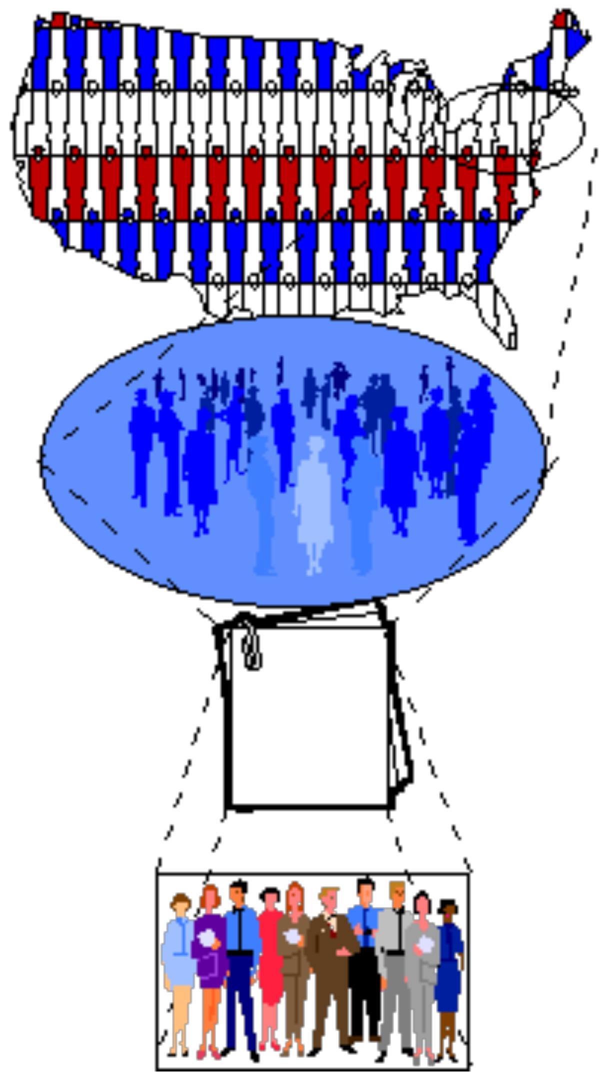
sample - group of people who you select to be in your study

before you collect your sample, you must have a list / procedure for getting your sample - sampling frame

actually, usually end up with a “subsample”

non-respondents, dropouts, etc.

Sampling - Terminology



Who do you want to generalize to?		The Theoretical Population
What population can you get access to?		The Study Population
How can you get access to them?		The Sampling Frame
Who is in your study?		The Sample

as you move though each step, you introduce more systematic error (aka bias)

Sampling - Statistical Terms

the individual "units" in your sample provide one or more responses - a specific measurement value

you use statistics to assess responses from your entire sample

- mean, variability, etc.

if you have a measurement of the entire population, you refer to parameters rather than statistics

the goal of statistics is to get an estimate of the parameter for the population that you sampled from

Sampling - Statistical Terms

Variable

response

1 2 3 4 5

Statistic

sample

Average = 3.75

Parameter

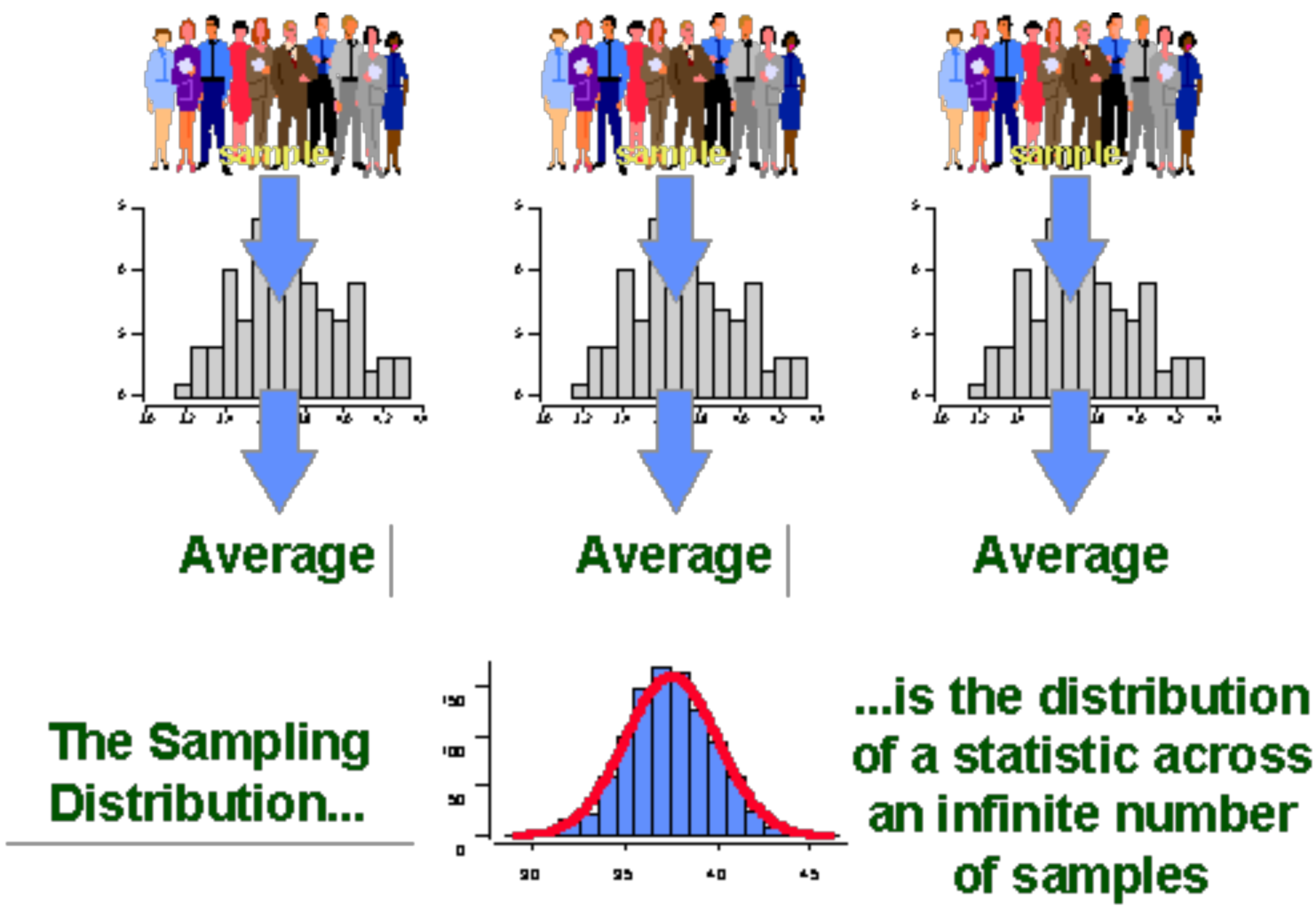
population

Average = 3.72

Sampling - Statistical Terms

getting from a statistic to an estimate of the population parameter:

sampling distribution:



Sampling - Statistical Terms

so...... the mean of the theoretical sampling distribution is the average of the averages of an infinite number of samples - essentially equal to the parameter

standard deviation - the spread of scores (variability) around the average in a single sample

the standard deviation of the theoretical sampling distribution is called the standard error

so.... the standard error is the spread of averages around the average of averages in the sampling distribution

Sampling - Statistical Terms

sampling error - the (lack of) precision of your statistical estimate

- *estimated* based on:
  - standard deviation of your sample
  - sample size (high n usually = low s.d.)
    - greater sample size measures a larger % of the actual population

Probability Sampling

probability sampling - any method that uses some form of random selection method

- assures that different units in your population have equal probabilities of being chosen

basic terms:

- N = number of cases in the sampling frame
- n = number of cases in the sample
- NCn = the number of subsets / combos of n from N
- f = n/N (the sampling fraction)

Nonprobability Sampling

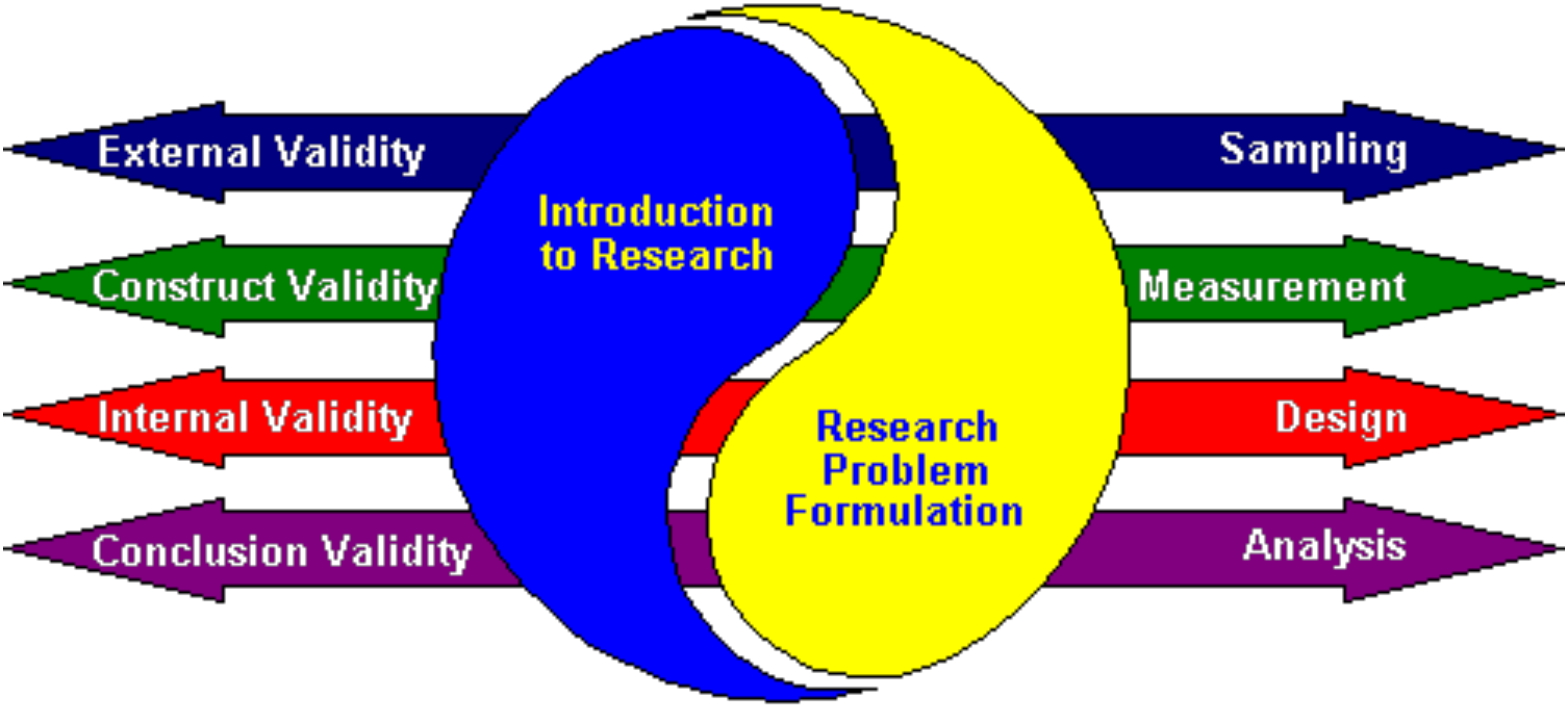
nonprobability sampling does not involve random selection

- used when random methods are not feasible
- statistical inferences not valid

2 broad types:

- accidental, haphazard, convenience sampling:
  - “man on the street” interviews
  - use of college undergrads
    - probably not representative of the population you want to generalize to
- purposive
  - mall market researchers looking for female Libertarians between 15-21
  - snowball sampling - identify some who fits criteria, then ask them to recommend others

Dealing with Construct Validity - Measurement



Dealing with Construct Validity - Measurement

measurement - the process of observing and recording the observations

Dealing with Construct Validity - Measurement

Scaling - involves the construction of an instrument that gives qualitative constructs quantitative units

the assignment of objects to numbers according to a rule

set of procedures for getting numbers that can be meaningfully assigned to objects

Dealing with Construct Validity - Measurement

operationalization - the translation of a concept (construct) into a functioning and operating reality

- formal, complete definition of every measure and manipulation in the study

construct validity - the degree to which inferences can be made from the operationalizations back to the theoretical constructs on which they were based

- generalizing from your measures to the "concept" of your measures
- is the "label" correct? does your measure of "self-esteem" really measure self-esteem?

Dealing with Construct Validity - Measurement

- “translation” validity - the degree to which your construct is accurately translated into the operationalization
  - *face validity* - does the operationalization seem like a good translation “on its face”?
    - relies on subjective judgement
    - weakest way to demonstrate validity
  - *content validity* - checking the operationalization against a good, detailed description of the construct
    - problem - what criteria constitute the content domain?
      - IQ, self-esteem, etc

Dealing with Construct Validity - Measurement

"criterion-related" validity - checking the performance of your operationalization against some criterion

- *predictive* - ability to predict something that it theoretically should be able to predict
- *concurrent* - ability to distinguish between groups that it theoretically should be able to distinguish between
- *convergent* - similarity to other operationalizations that it theoretically should be similar to
- *discriminant* - LACK OF similarity to other operationalizations that it theoretically should be different from

Dealing with Construct Validity - Measurement

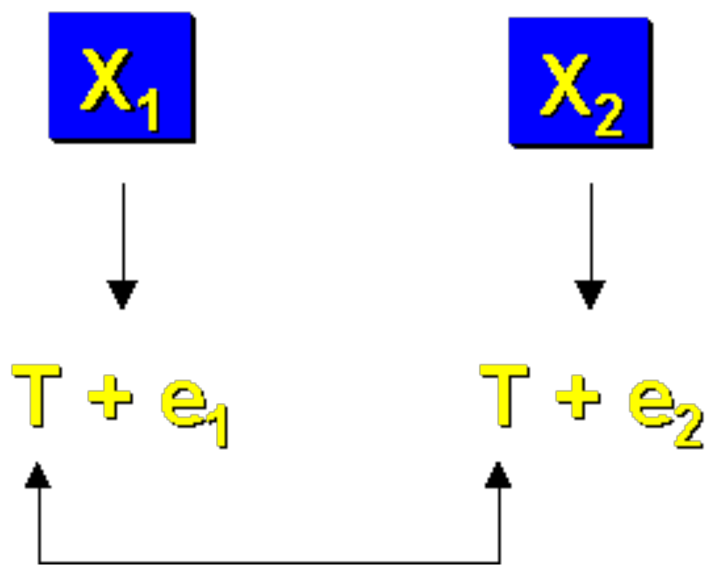
Threats to construct validity:

- *inadequate pre-operational explanation of constructs* - not doing a good job of defining the construct
- *mono-operation bias* - not using enough independent variables (treatments, etc)
- *mono-method bias* - not measuring enough dependent variables
- *unexpected interaction* of different treatments
- *unexpected interaction* of measurements with treatments
- *unintended consequences* of the treatment (side effects, etc)
- *social threats*:
  - hypothesis guessing by participants
  - test anxiety
  - experimenter bias

Dealing with Construct Validity - Measurement

Reliability - a measure is considered reliable if it gives the same result over & over again

measuring something twice should yield a similar answer



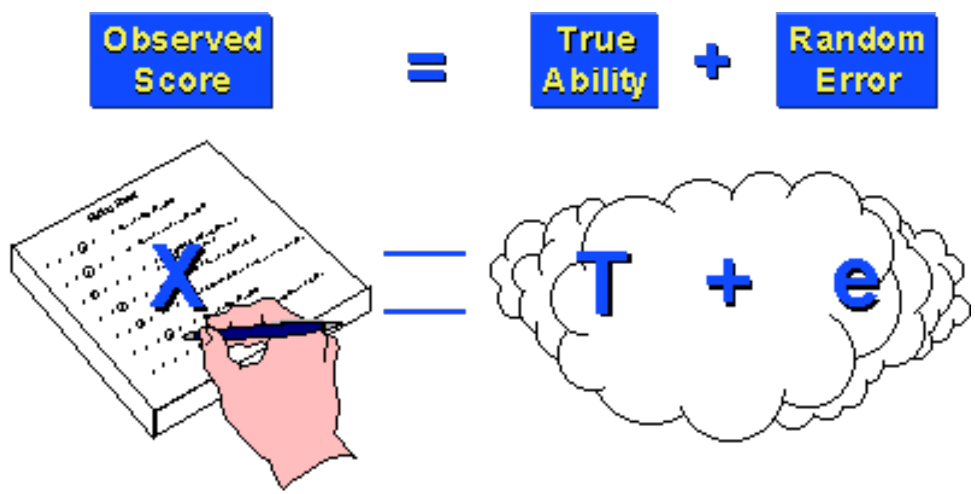
Dealing with Construct Validity - Measurement

True Score theory - every measurment is an additive composite of 2 components:

"true" ability / level

random error

$X = T + e\mathbf{x}$



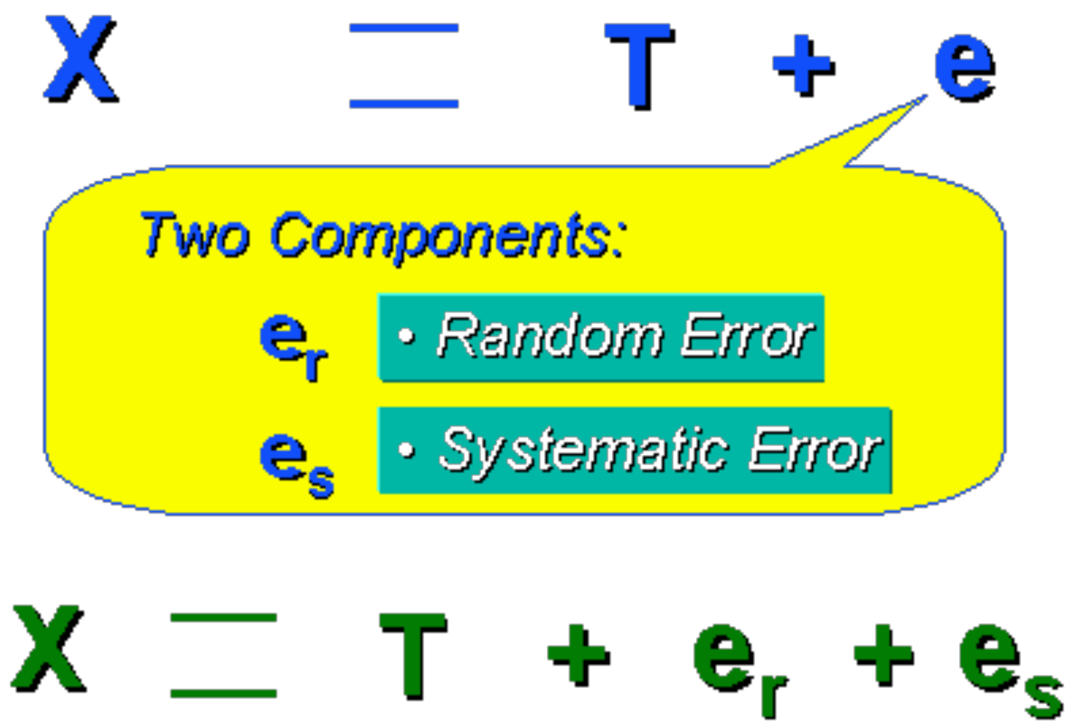
measures with no random error are perfectly reliable

measures with no true score component are not at all reliable

Dealing with Construct Validity - Measurement

Not all error is random

may be partly systematic, across all measurements

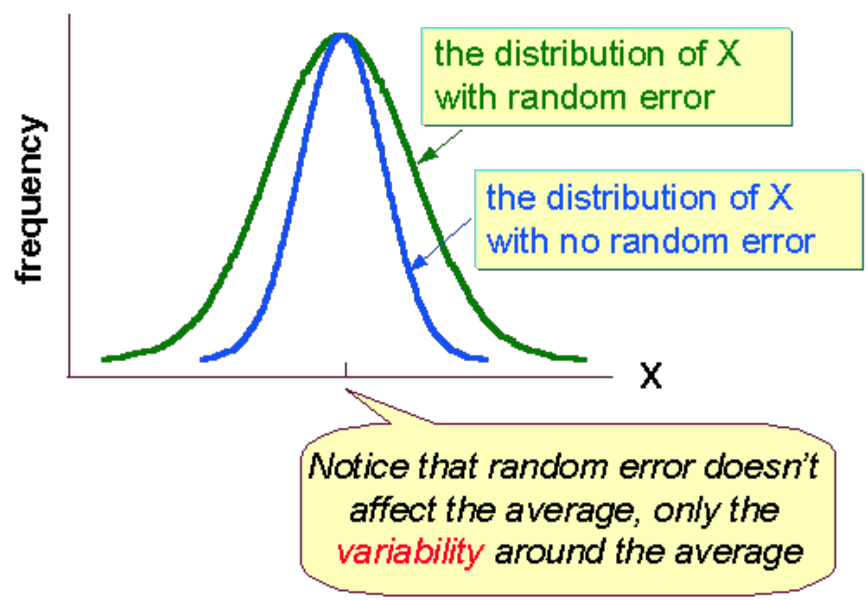


Dealing with Construct Validity - Measurement

Random error:

measurements of the variable are affected randomly (inconsistently) across your sample

adds variability / noise, but does not change the average

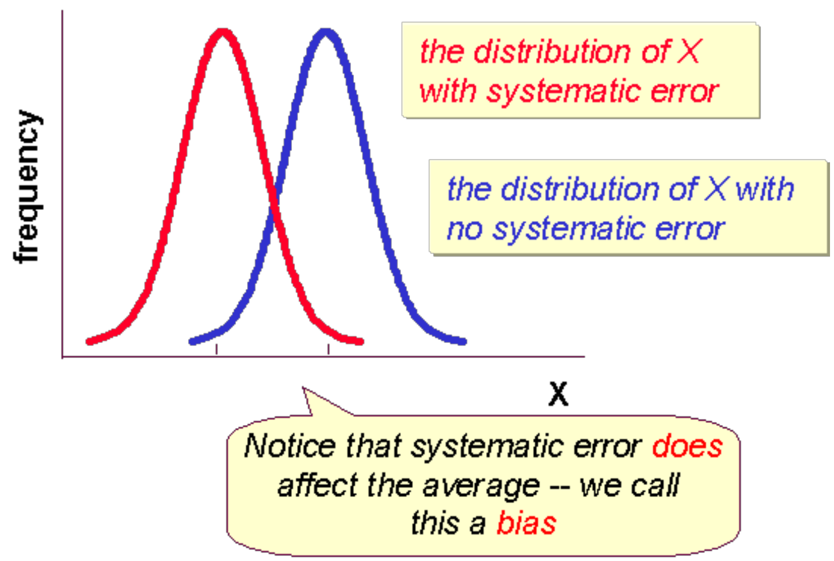


Dealing with Construct Validity - Measurement

Systematic error:

measurements of the variable are affected systematically (consistently) across your sample

adds bias in one or the other direction



Dealing with Construct Validity - Measurement

4 types of reliability estimates:

inter-rater

categorical measurements - % of agreement

continuous measurements - correlation

test-retest

depends on interval between retest

test history effects

parallel forms

using different measures - may control for test history effects

1/2 get Form A at test time 1, the other half get Form B,

switched at test time 2

internal consistency

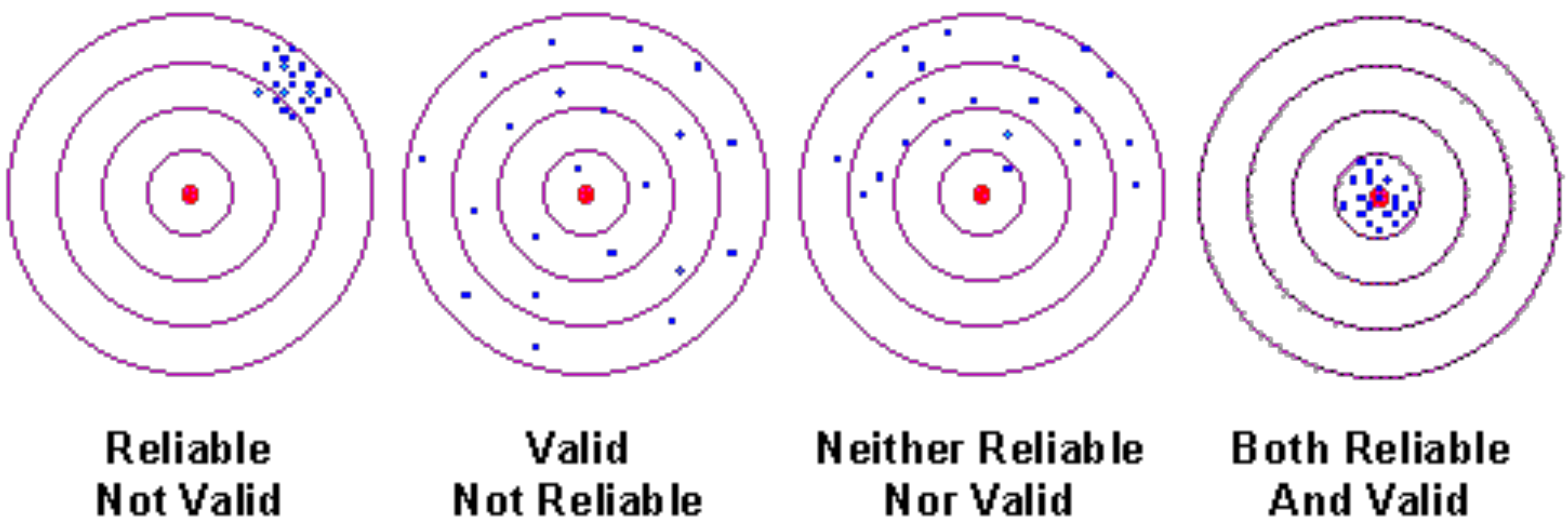
looking at different items for the same construct within a measure

average inter-item correlation, average item total correlation, split-half reliability, Cronbach's Alpha

Dealing with Construct Validity - Measurement

Reliability vs. Validity

- center of the target is the construct that you're measuring
- each "shot" is a measurement

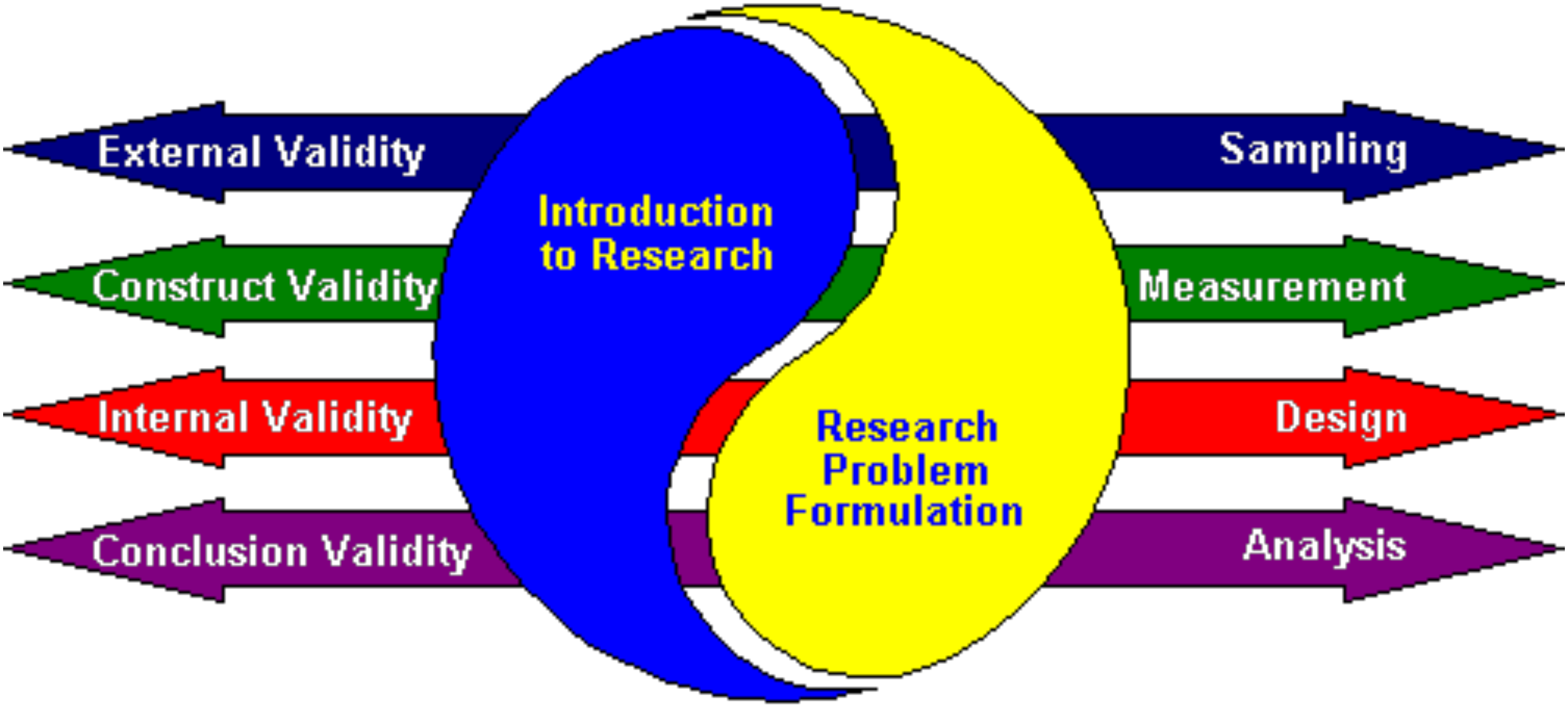


Dealing with Construct Validity - Measurement

Reducing measurement error:

- pilot test your instruments of measurement
- be consistent in your measures and across measurers
- double-check (double enter) data
- use multiple measures of the same construct
  - useful for “triangulating” away from systematic error
- use statistical procedures to adjust for measurement error

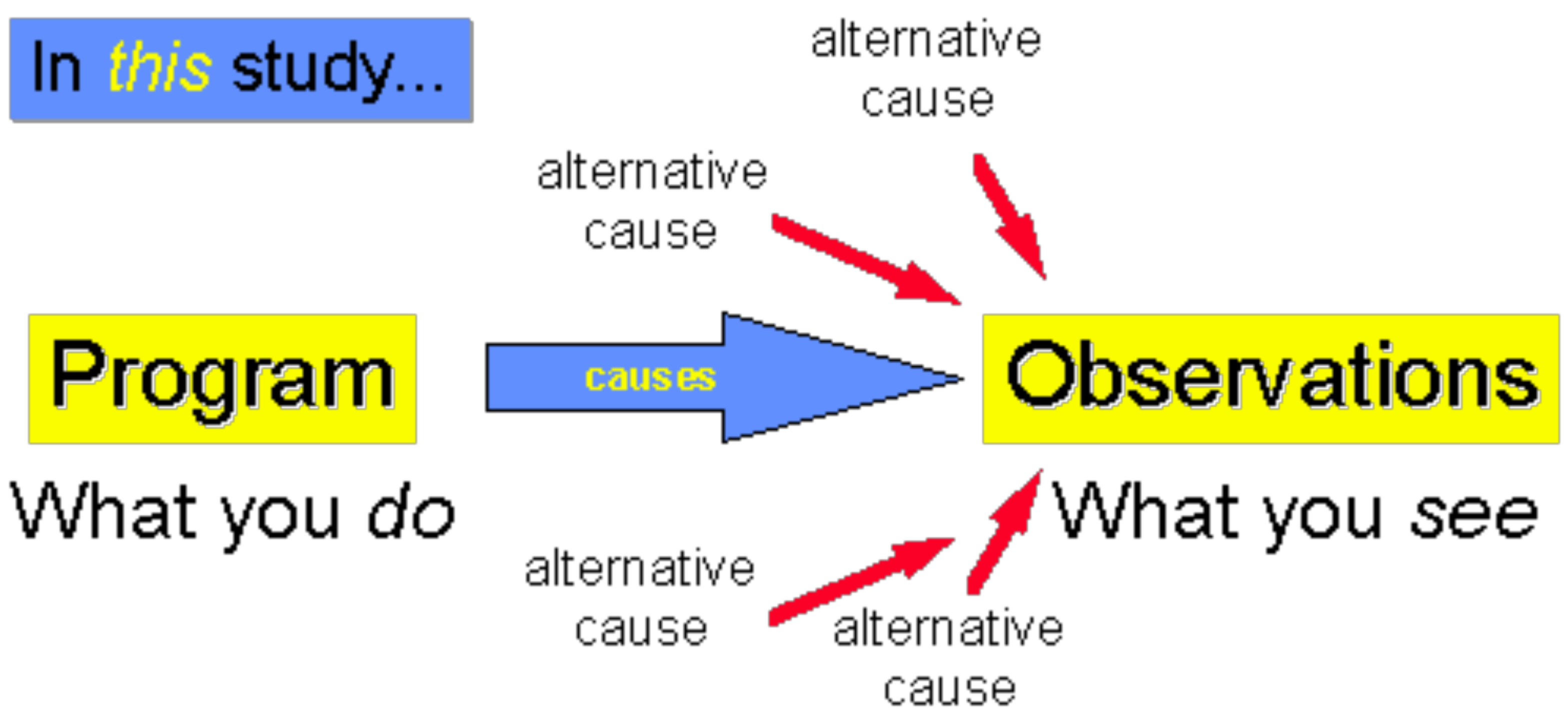
Dealing with Internal Validity - Experimental Design



Dealing with Internal Validity - Experimental Design

Internal validity - are the "cause-effect" inferences in your study valid?

did your treatment / program / independent variable actually have an effect, or are there alternative explanations?



Dealing with Internal Validity - Experimental Design

- Good designs are:
  - hypothesis-driven
  - redundant / flexible - “over-engineered”
  - efficient - balance between redundant and over-engineered
  - feasible

Dealing with Internal Validity - Experimental Design

- **design strategies:**
  - depict the simple hypothesized causal relationship
  - over-expand the design across:
    - time
    - treatment
    - measurements
    - groups
  - scale back design using a cost-benefit analysis

Dealing with Internal Validity - Experimental Design

- most research is dedicated to determining whether some treatment causes some result
- 3 conditions must be met to infer <u>cause & effect</u>:
  - <u>covariation</u> - changes in the "cause" should be associated with changes in the "effect"

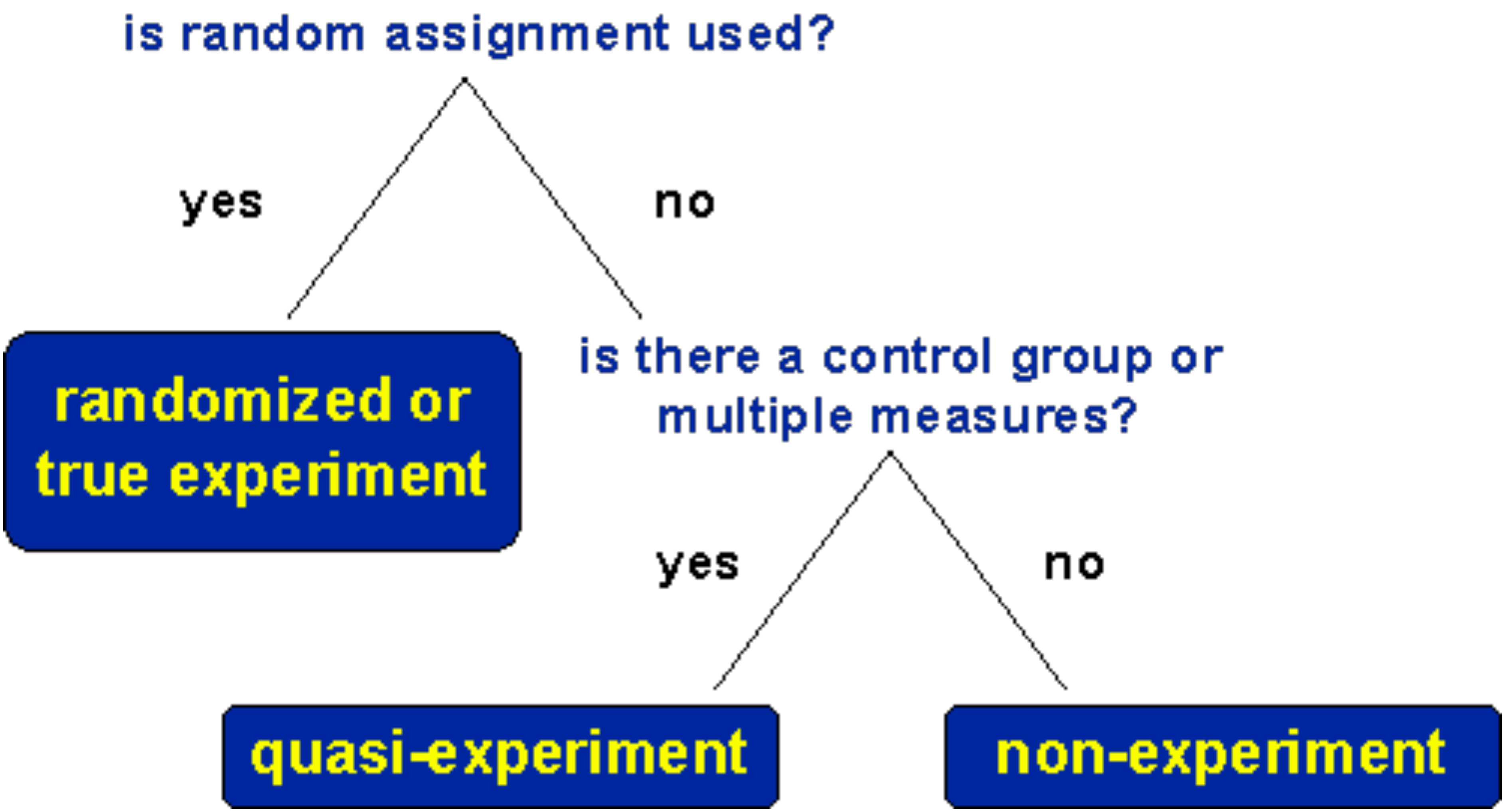
    with your treatment, you see an effect

    without your treatment, you see no effect

    does *more* treatment cause *more* effect?
  - <u>temporal precedence</u> - the "cause" must occur prior to the "effect"
  - <u>no other *alternative explanations*</u>
    - most difficult condition to meet
      - <u>*alternative explanations*</u> are threats to internal validity
      - a relationship does not imply a causal relationship
      - "3rd" or "missing" variable?
      - proper experimental design is key!!!

Dealing with Internal Validity - Experimental Design

3 major designs - in order of strongest to weakest in terms of being able to demonstrate internal validity / casual assessment:



Dealing with Internal Validity - Experimental Design

Quasi-Experimental Design

nonequivalent group design

looks pretty much like a "true experimental" design, but no random assignment

inferior to experimental designs with respect to internal validity

but, may be more easily implemented

Dealing with Internal Validity - Experimental Design

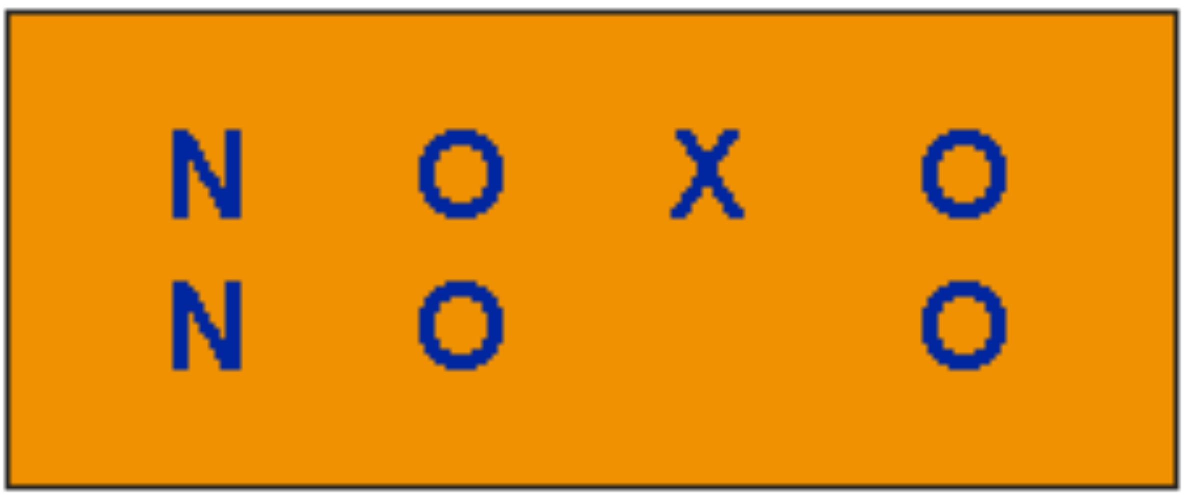
Quasi-Experimental Design

Nonequivalent Groups Design

most frequently used design in social research

looks like a pretest-posttest experiment:

- no random assignment



Dealing with Internal Validity - Experimental Design

"Experimental" Design

- most rigorous / gold standard design
  - strongest with respect to internal validity
    - cause & effect:
      - if X, then Y
      - if *not* X, then *not* Y
- generally uses 2+ "equivalent" groups (control, treatment, etc.)
- random assignment creates "probabilistically equivalent" groups
- achieving greater *internal validity* often reduces *external validity*
  - can set up artificial situations that are not generalizable

Dealing with Internal Validity - Experimental Design

"Experimental" Design
2-Group Designs

Simplest experimental design:

2-group posttest-only randomized design

due to random assignment, the groups are "probably equivalent" before treatment

are the groups different after treatment?

R	X	O
R		O

Dealing with Internal Validity - Experimental Design

an experiment or quasi-experiment is internally valid if it shows a cause-effect relationship between the IV & DV

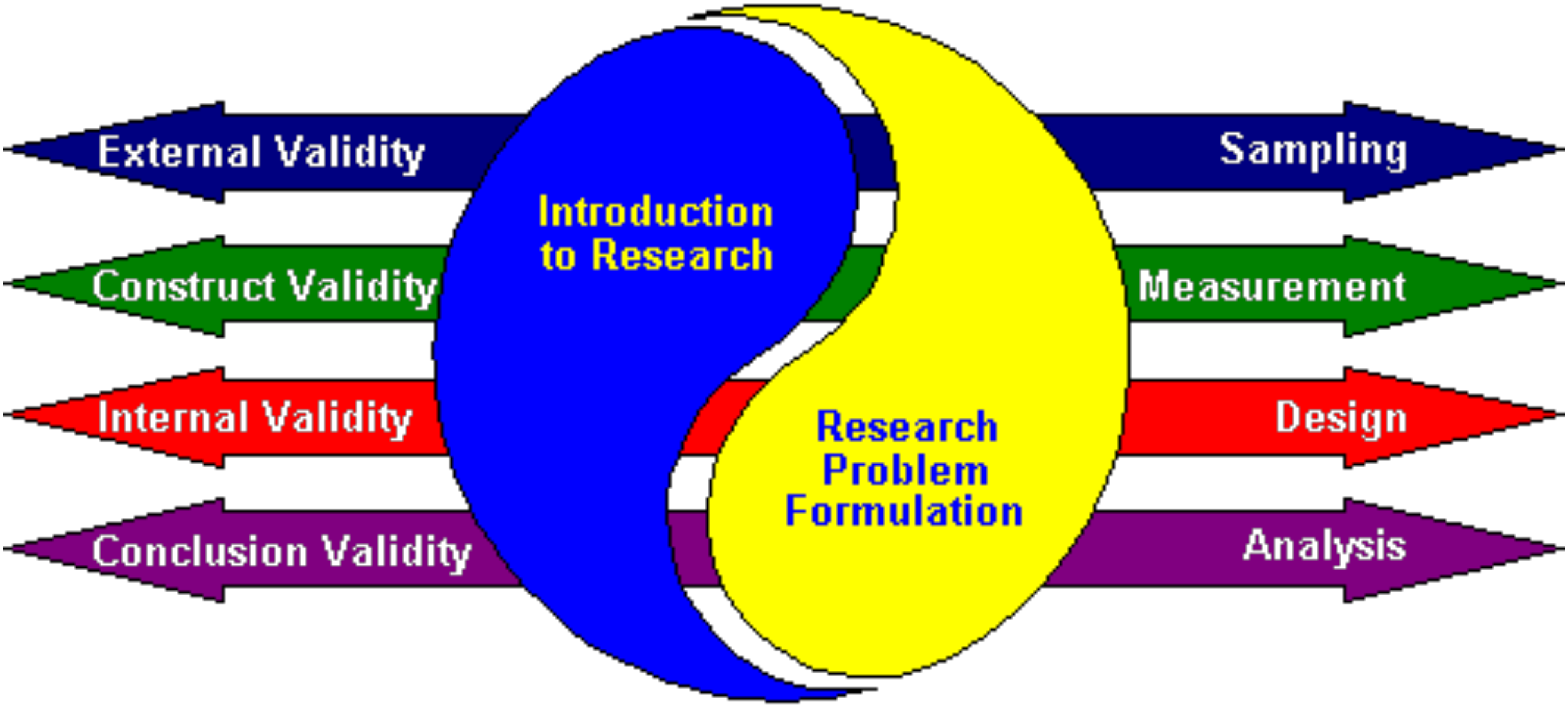
Threats to internal validity:

- selection
- maturation
- selection-maturation interaction
- history
- repeated testing
- instrumentation
- regression to the mean
- experimental mortality
- experimenter bias

Dealing with Internal Validity - Experimental Design

- ways to minimize threats to validity / alternate explanations
  - argument - weakest approach (but *before* is better than after)
  - measurement / observation / analysis - demonstrate that a threat occurs minimally or not all, or co-vary out alternate explanations
  - design - add control groups, extra treatment groups, waves of repeated measurements, etc.

Dealing with Conclusion Validity - Analysis



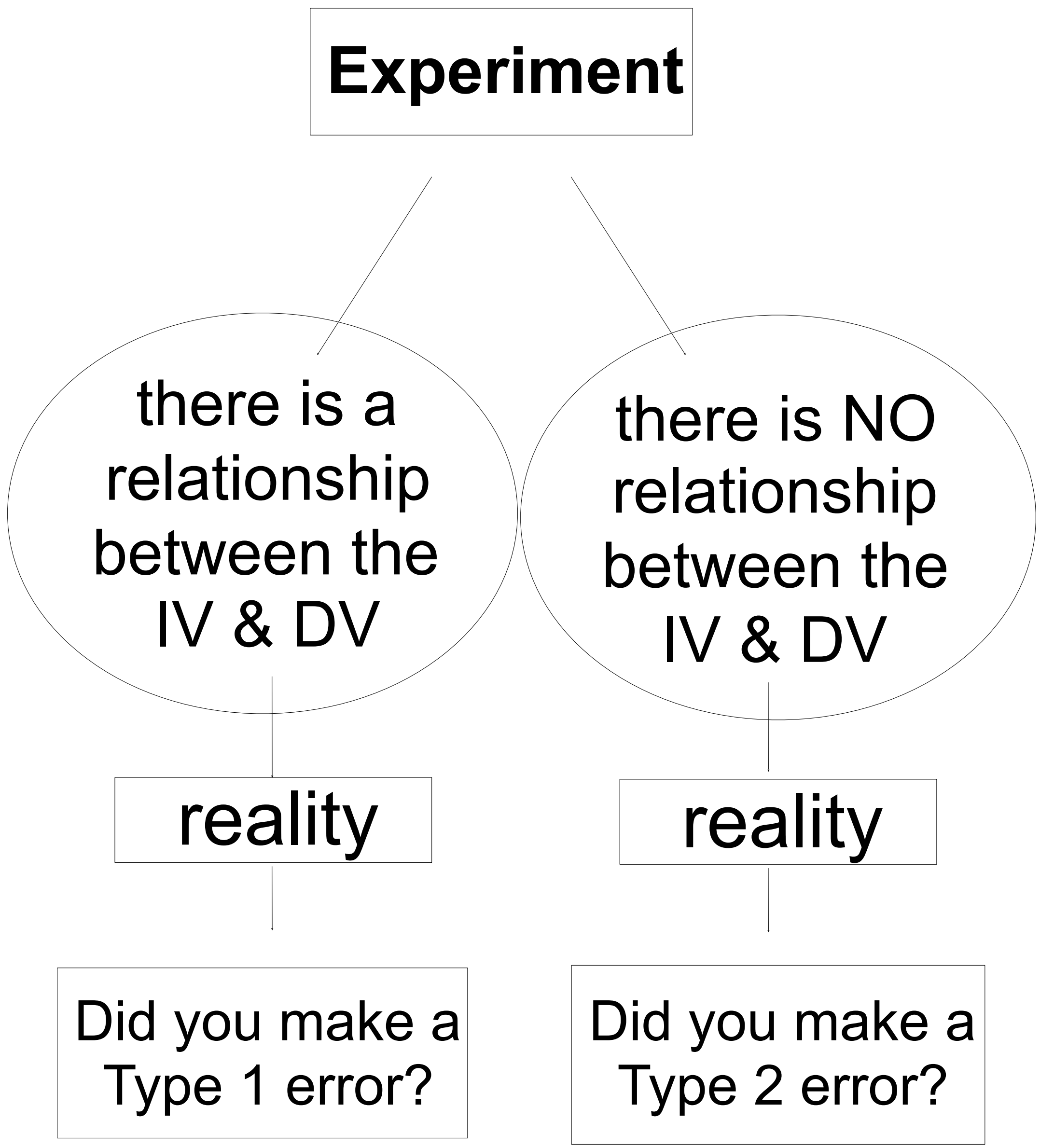
Dealing with Conclusion Validity - Analysis

- probabilistic - absolute certainty is almost unobtainable (that why we use statistics)

- **Conclusion Validity - the degree to which conclusions we reach about relationships in our data are reasonable.**
  - **the credibility of your conclusion on the relationship that you found between your IV and DV**

Dealing with Conclusion Validity - Analysis

“Statistical” Conclusion Validity



Dealing with Conclusion Validity - Analysis

- big threat to validity: insufficient statistical power
  - sample size too small
  - too much error in measurements

Dealing with Conclusion Validity - Analysis

Power: Alpha

- we <u>test</u> our hypothesis to determine whether our experimental condition (H_1) is different from null (H_o).
  - $(H_1) \neq (H_o)$

- Alpha (α) is the probability that we are <u>rejecting the null hypothesis</u> when the null hypothesis <u>true</u>
  - TYPE 1 (α) ERROR - "false positive"
    - generally (.05) 1 in 20

Dealing with Conclusion Validity - Analysis

– TYPE 1 (α) ERROR - “false positive”

Dealing with Conclusion Validity - Analysis

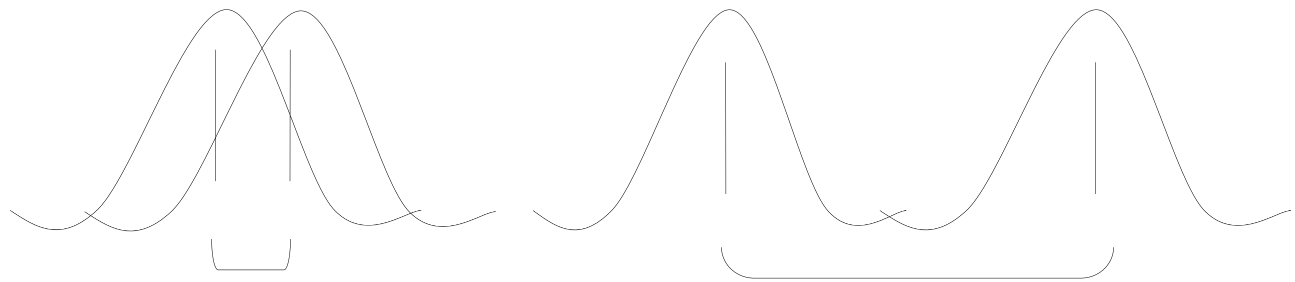
Power: Beta

- If we decide that our hypothesis is <u>not true</u>, we are <u>accepting</u> the null hypothesis
    - $(H_1) = (H_o)$
- rejecting H_1 when H_1 <u>is true</u> is a TYPE 2 (β) ERROR
    - Beta (β) - generally .2 (1 out of 5 times)
        - “false negative”

Dealing with Conclusion Validity - Analysis

– TYPE 2 (α) ERROR - “false negative”

Dealing with Conclusion Validity - Analysis



Small effect size Large effect size

Power: Effect Size

- **Effect size is the <u>difference</u> between the mean of your distribution and the null distribution**
- **The "strength" of our statistical test to detect <u>real</u> differences in our treatment from the null treatment is POWER**

– **1- β (80% chance of not making a type 2 error)**

Dealing with Conclusion Validity - Analysis

Threats to Conclusion Validity

- Low Power:
  - calculate sample size based on power estimates from either pilot studies or the literature before you start collecting subjects
    - Use larger sample sizes
    - try to get equal n per cell
    - reduce random setting irrelevances (NOISE)
      - **MAINTAIN *CONSISTENCY***

Dealing with Conclusion Validity - Analysis

Threats to Conclusion Validity

- Violated Assumptions of statistical tests
  - severe skewness?
  - non-normal data?
    - non-linear transformation
    - nonparametric tests

Dealing with Conclusion Validity - Analysis

Threats to Conclusion Validity

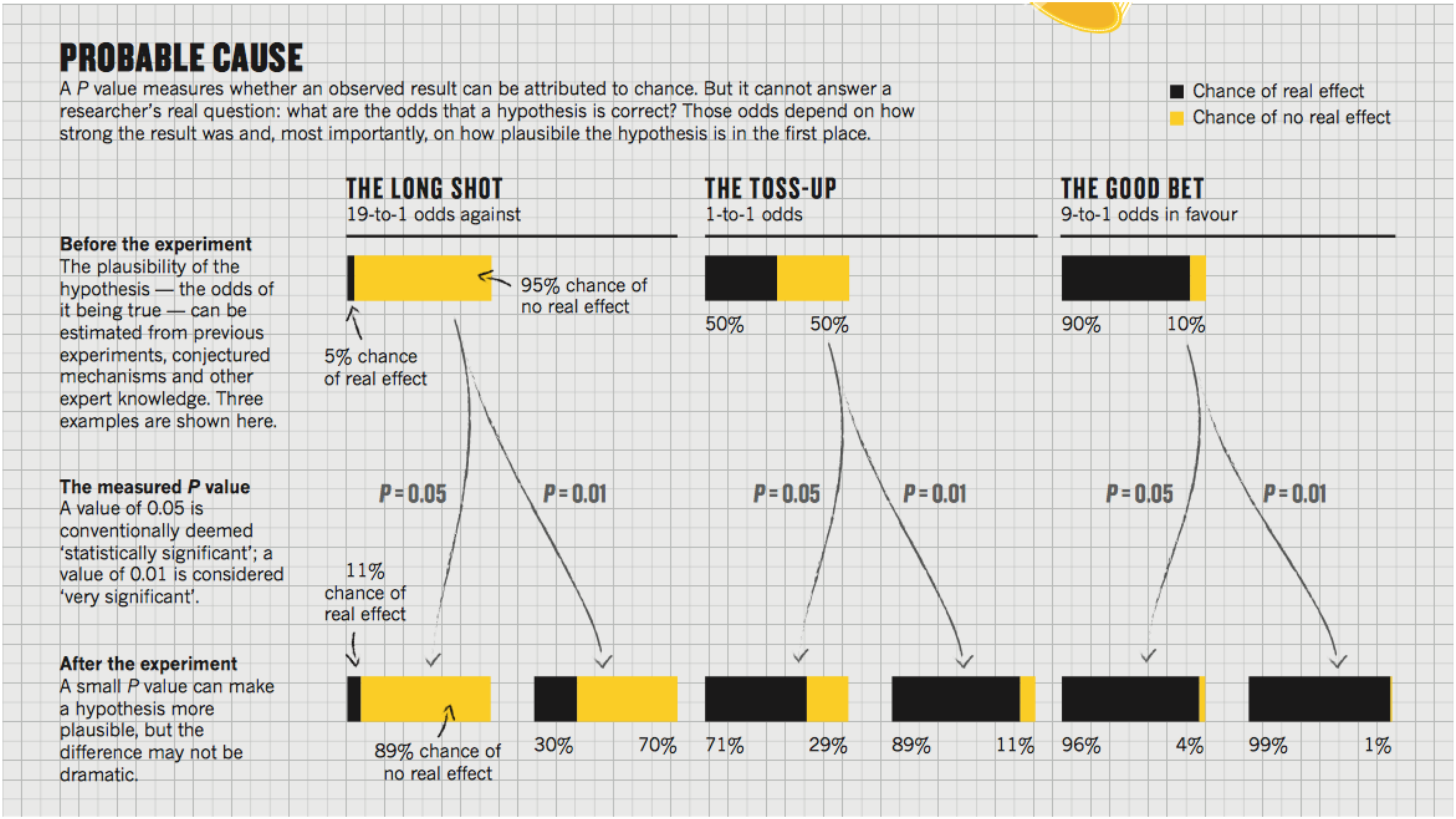
- Data fishing
  - the more tests you do to look at, the more you are likely to encounter “significant” results
    - So correct for your post-hoc alpha inflation by using appropriate statistical corrections
      - Bonferroni, Tukey’s, LSD
        - e.g., use the number of post-hoc tests/.05 to recalculate the significance level for the Bonferroni test statistic.
  - if your “significant results” don’t make sense or you can’t interpret them easily, maybe they are spurious

Dealing with Conclusion Validity - Analysis

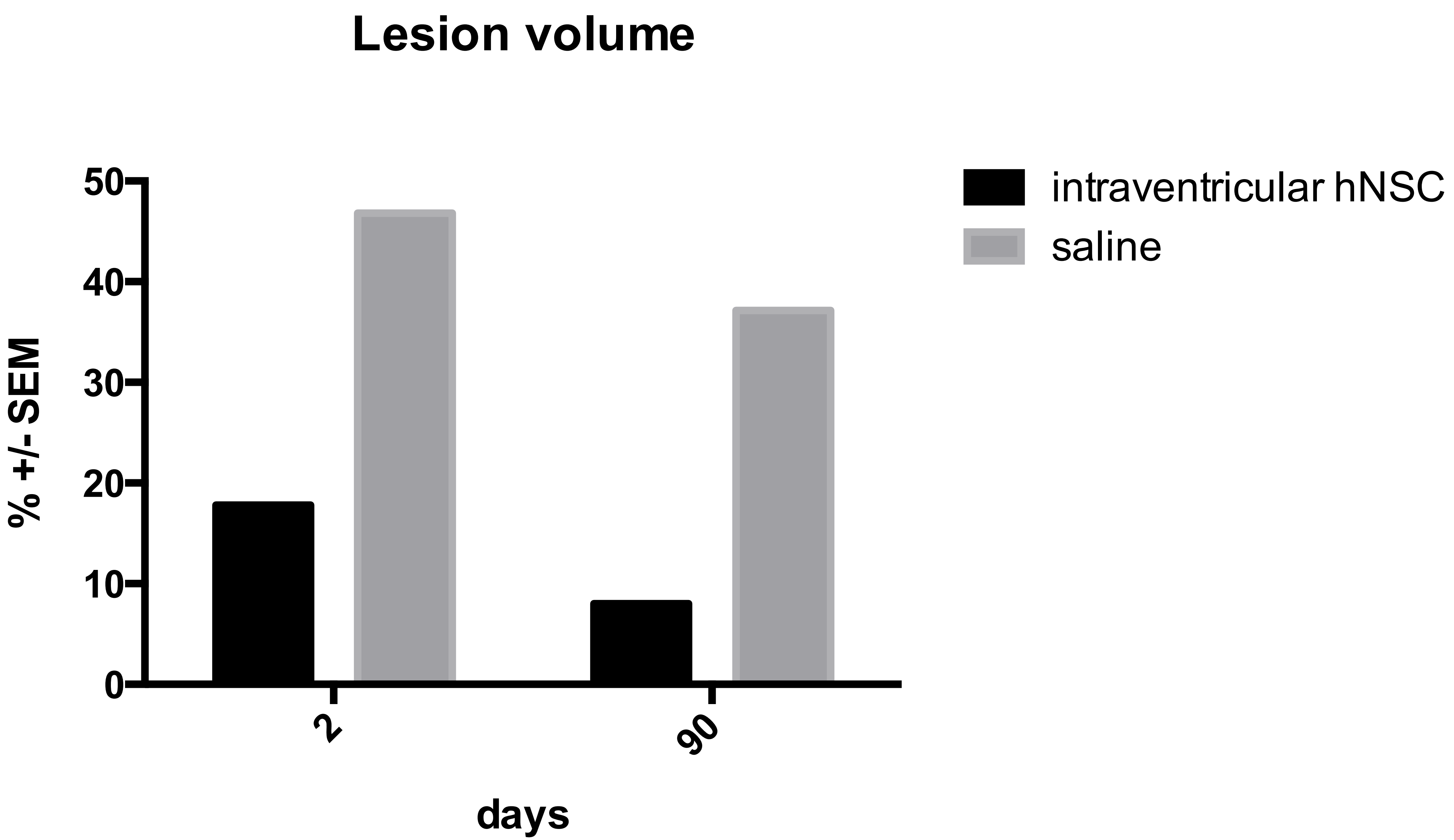
Threats to Conclusion Validity

- Unreliable measures
  - use measures that have been used before
  - unless your scale is very straightforward (e.g., “how many days per week do you have at least 1 alcoholic drink?”) don’t attempt to reinvent the wheel

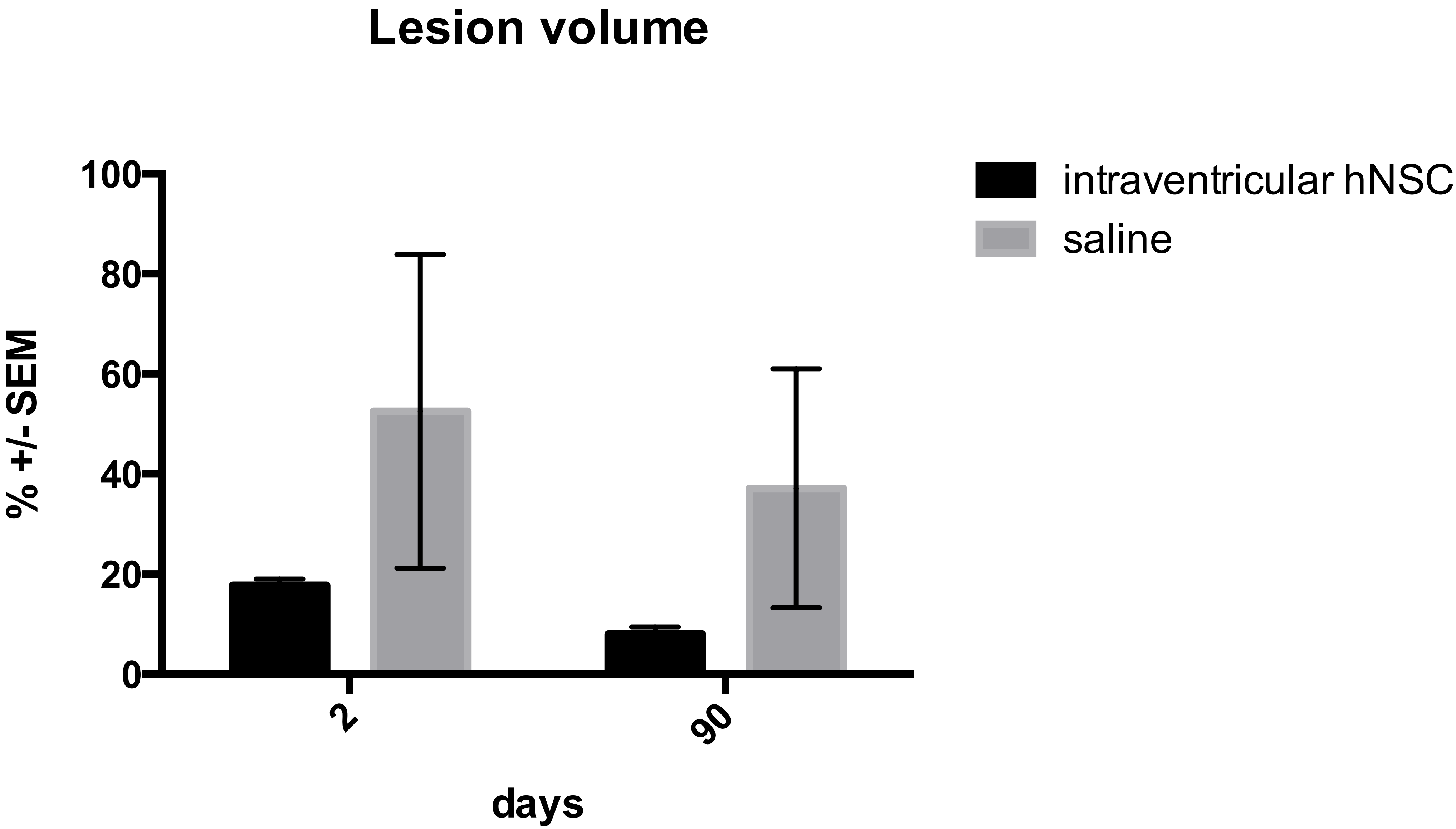
- Unreliable treatment
  - use the literature / pilot experiments as your guide


PROBABLE CAUSE
A P value measures whether an observed result can be attributed to chance. But it cannot answer a researcher's real question: what are the odds that a hypothesis is correct? Those odds depend on how strong the result was and, most importantly, on how plausibile the hypothesis is in the first place.
Chance of real effect
Chance of no real effect
THE LONG SHOT
19-to-1 odds against
THE TOSS-UP
1-to-1 odds
THE GOOD BET
9-to-1 odds in favour
Before the experiment
The plausibility of the hypothesis — the odds of it being true — can be estimated from previous experiments, conjectured mechanisms and other expert knowledge. Three examples are shown here.
95% chance of no real effect
5% chance of real effect
50%
50%
90%
10%
The measured P value
A value of 0.05 is conventionally deemed 'statistically significant'; a value of 0.01 is considered 'very significant'.
P = 0.05
P = 0.01
P = 0.05
P = 0.01
P = 0.05
P = 0.01
11% chance of real effect
After the experiment
A small P value can make a hypothesis more plausible, but the difference may not be dramatic.
89% chance of no real effect
30%
70%
71%
29%
89%
11%
96%
4%
99%
1%

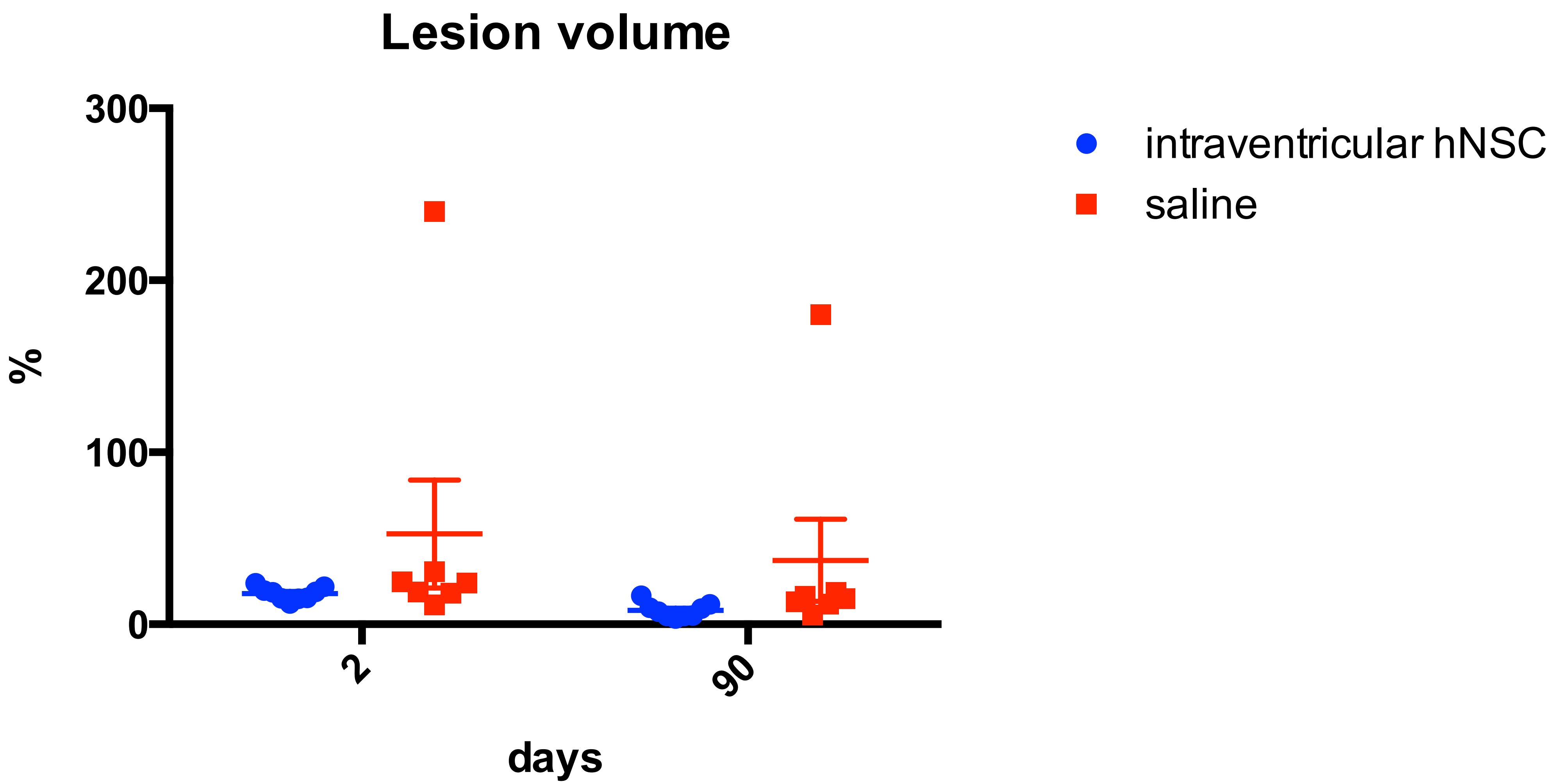
Dealing with Conclusion Validity - Analysis



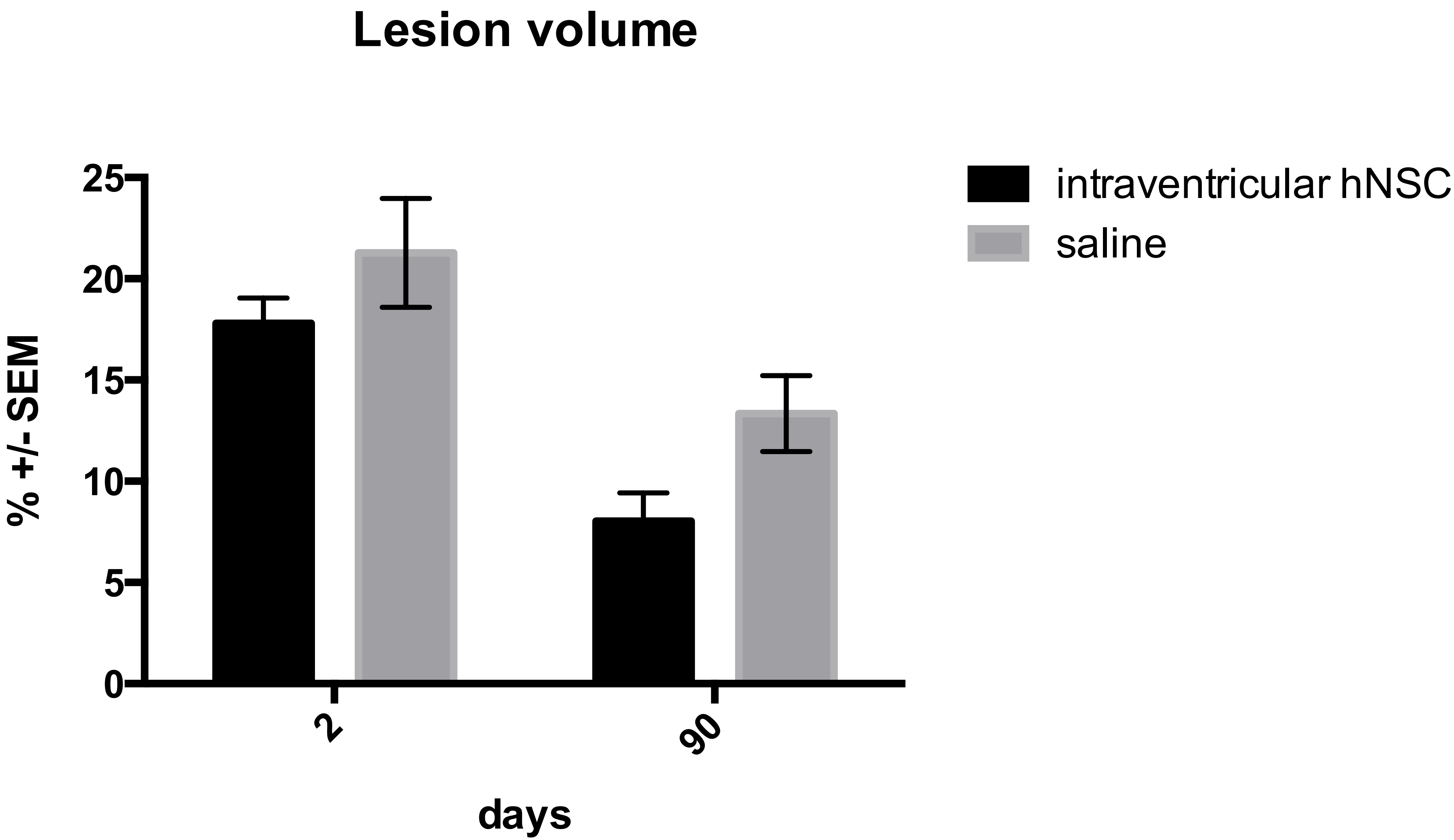
Dealing with Conclusion Validity - Analysis



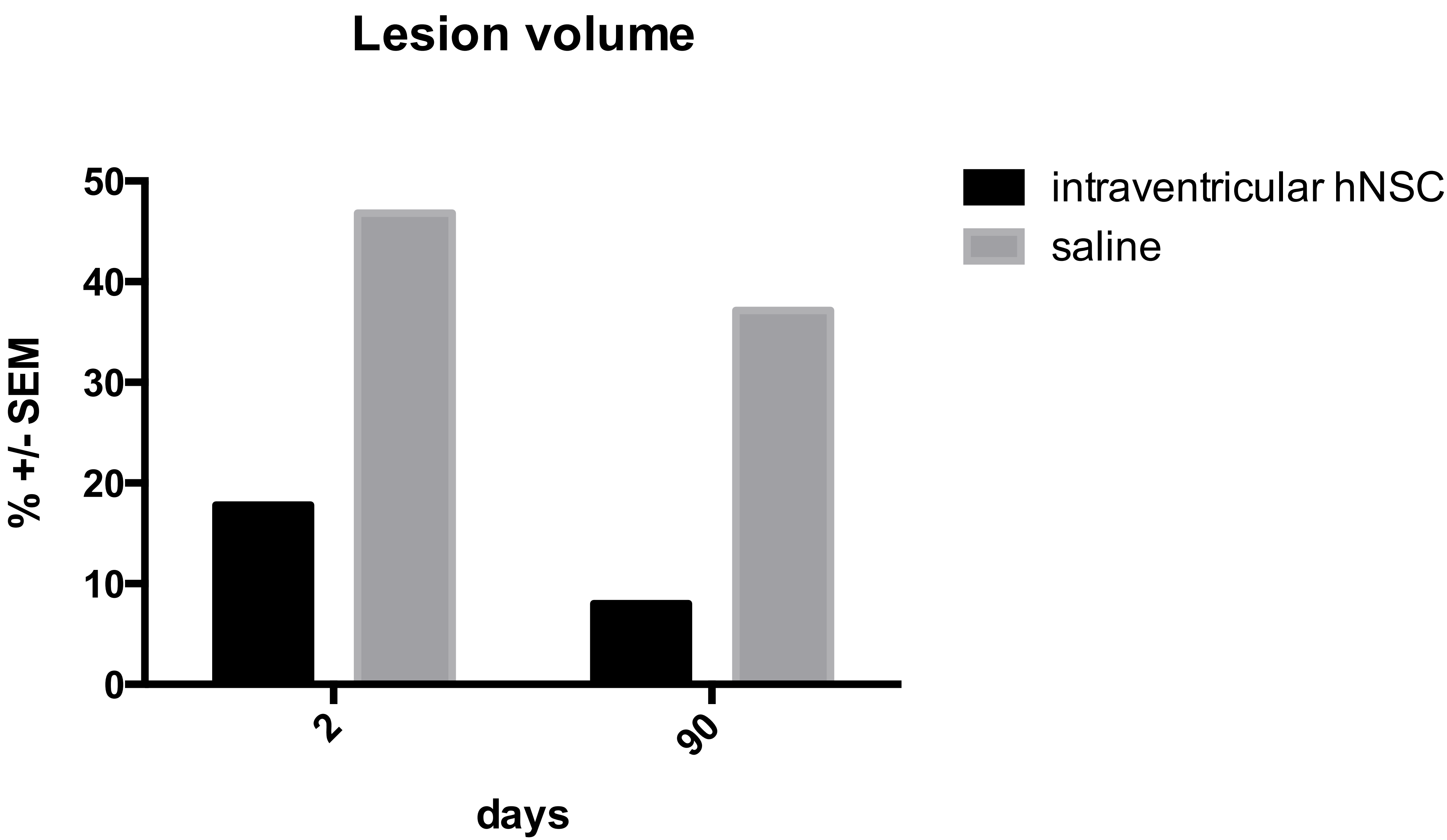
Dealing with Conclusion Validity - Analysis



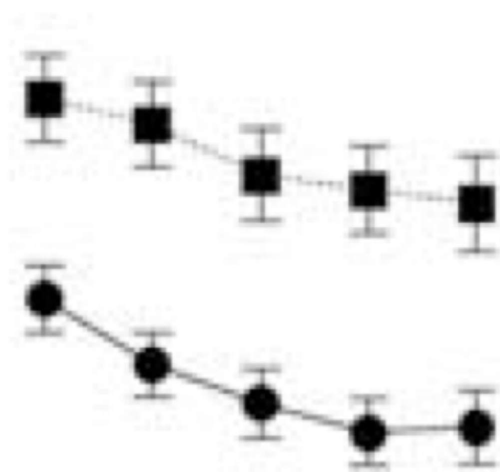
Dealing with Conclusion Validity - Analysis



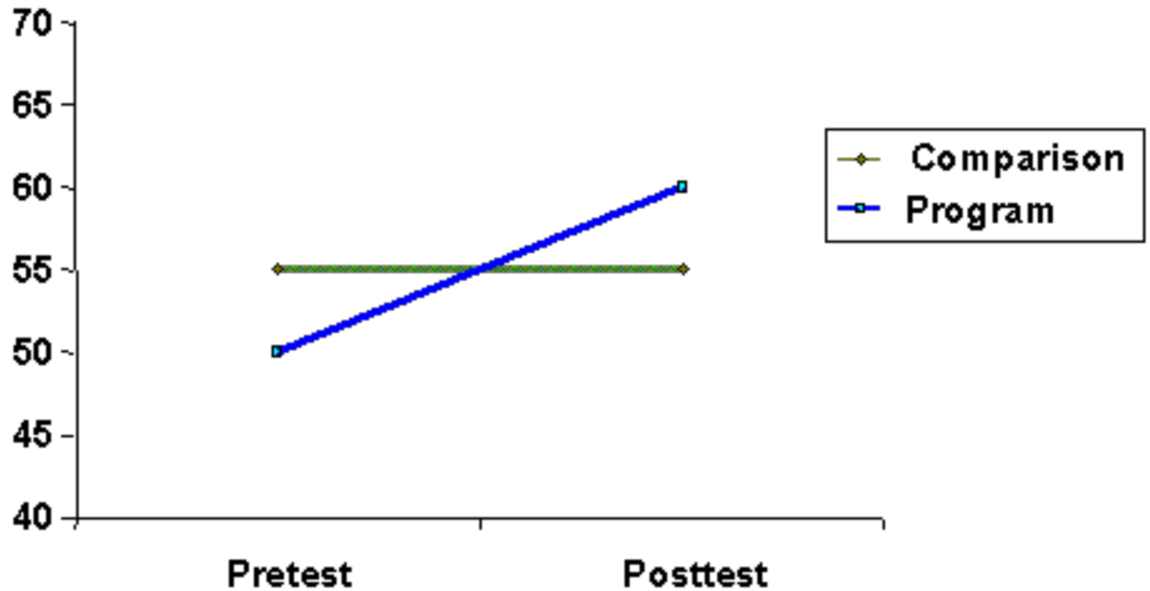
Dealing with Conclusion Validity - Analysis



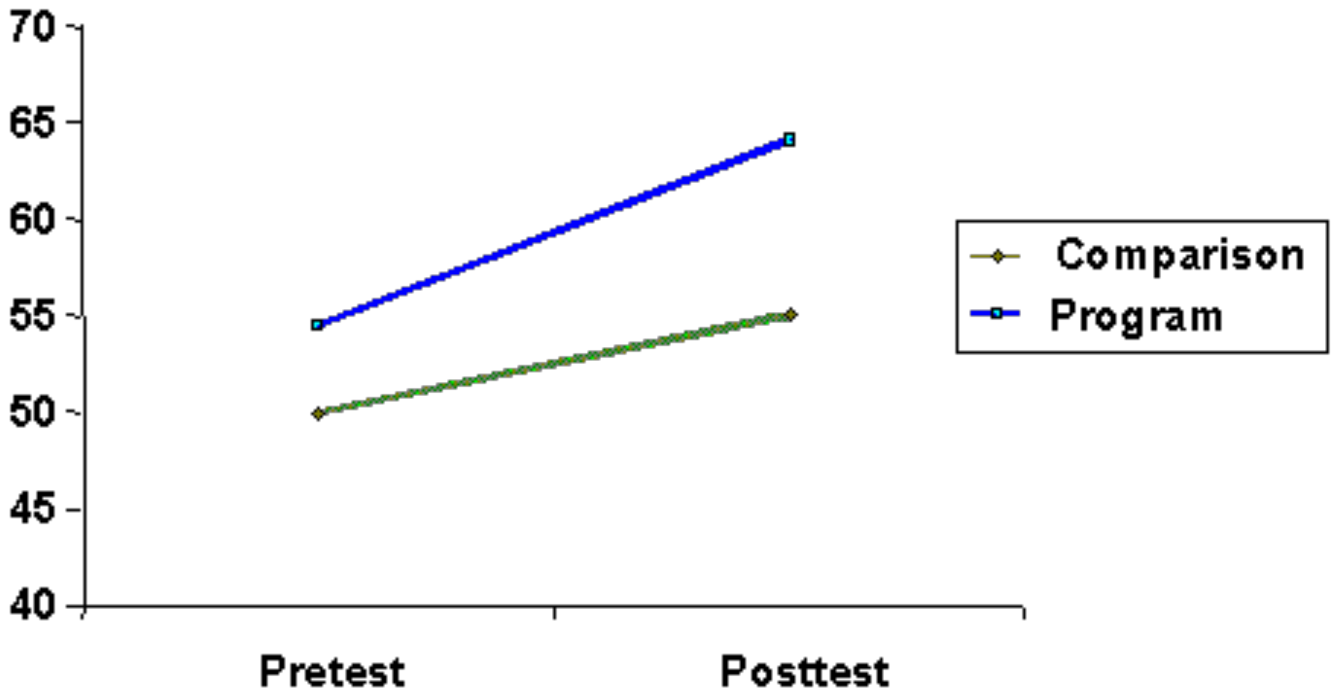
Main Effect or Interaction?



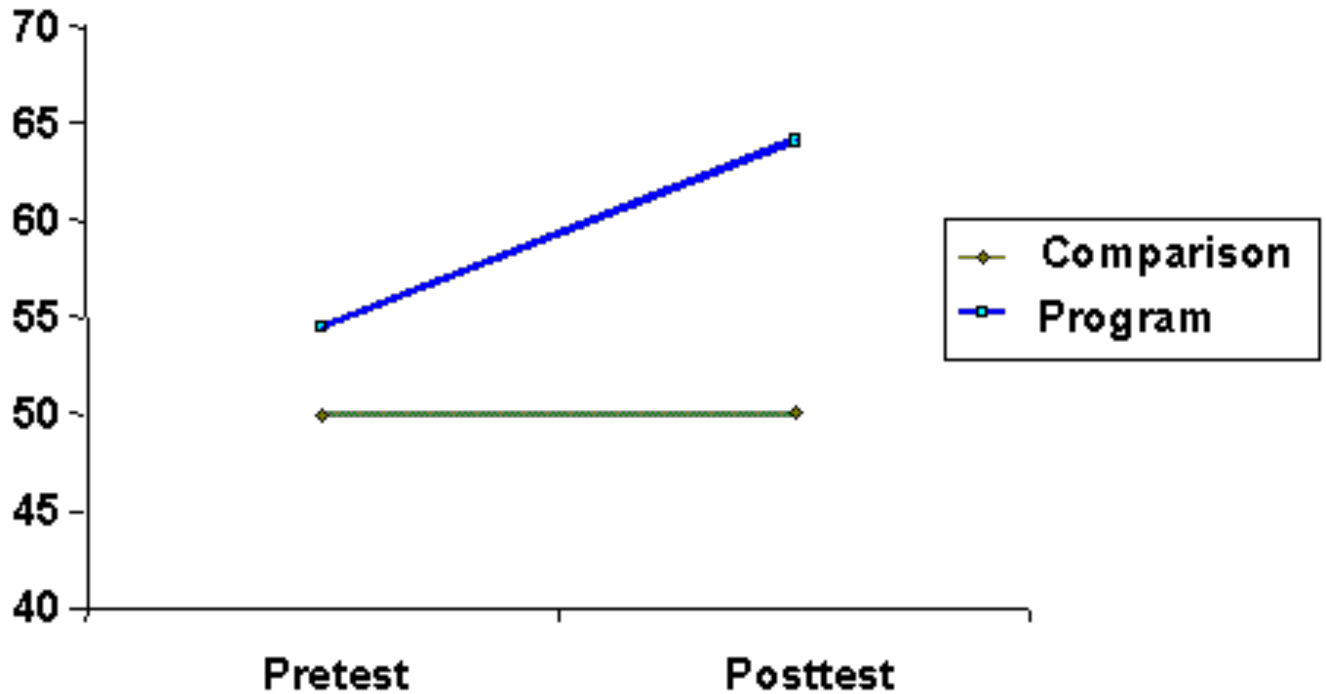
Main Effect or Interaction?



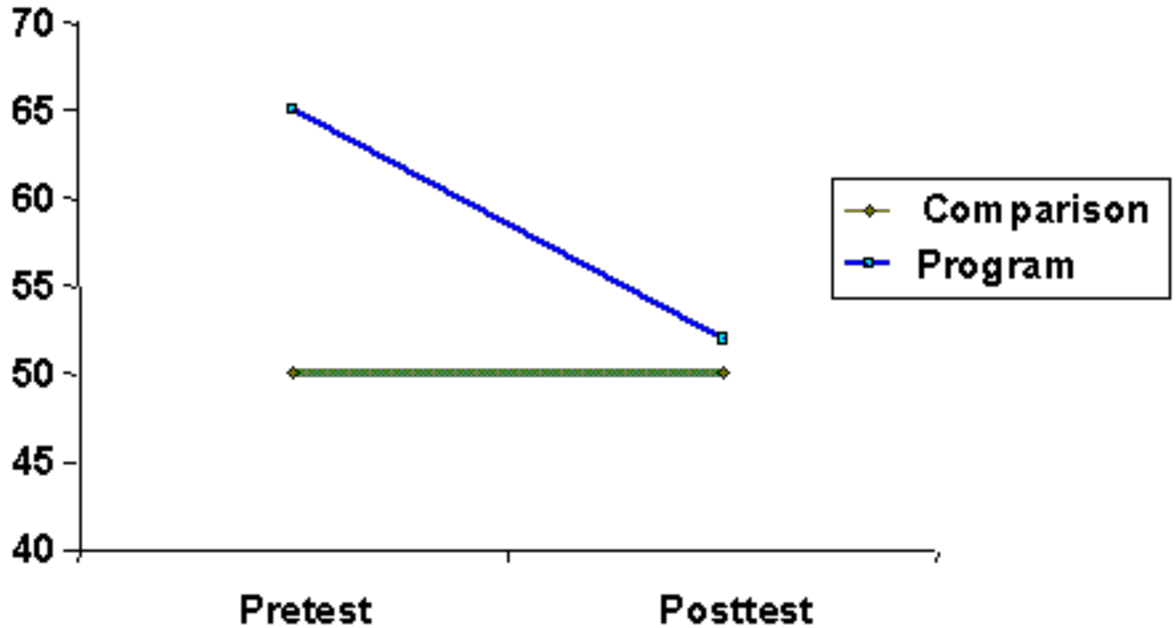
Main Effect or Interaction?



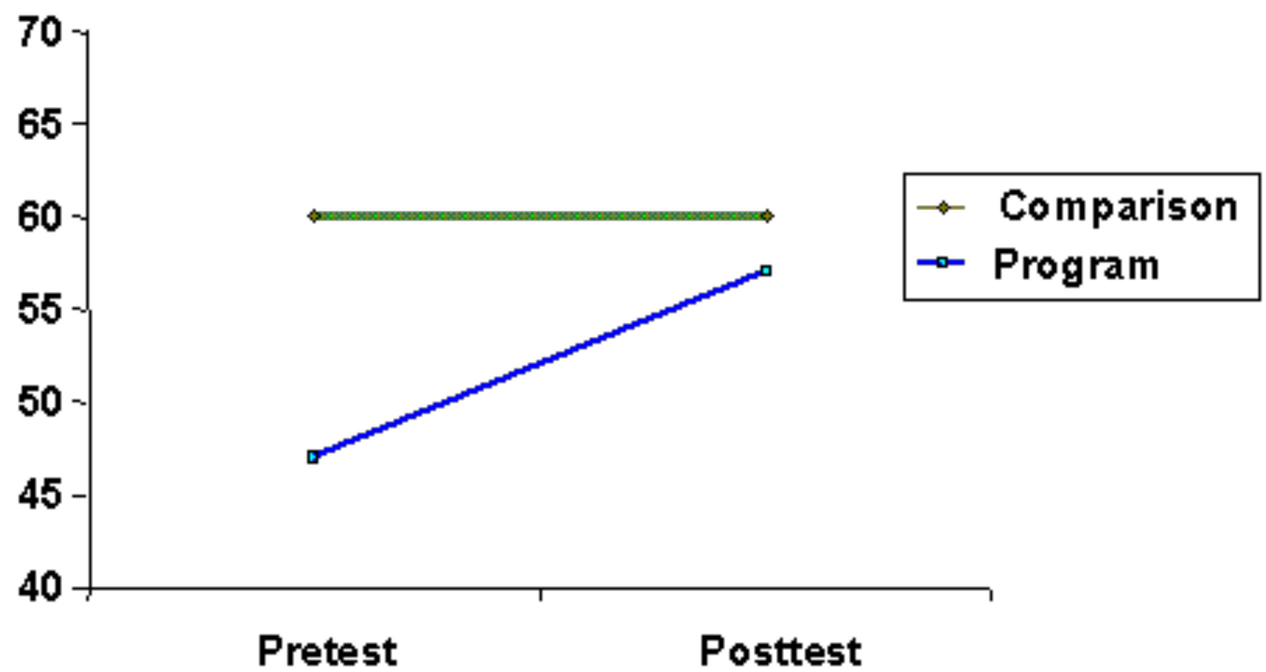
Main Effect or Interaction?



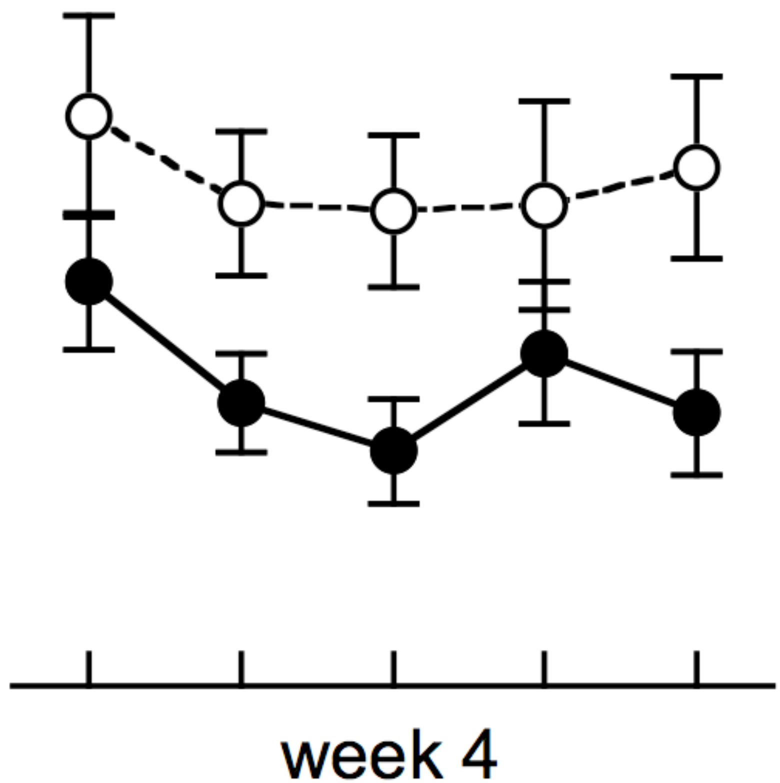
Main Effect or Interaction?



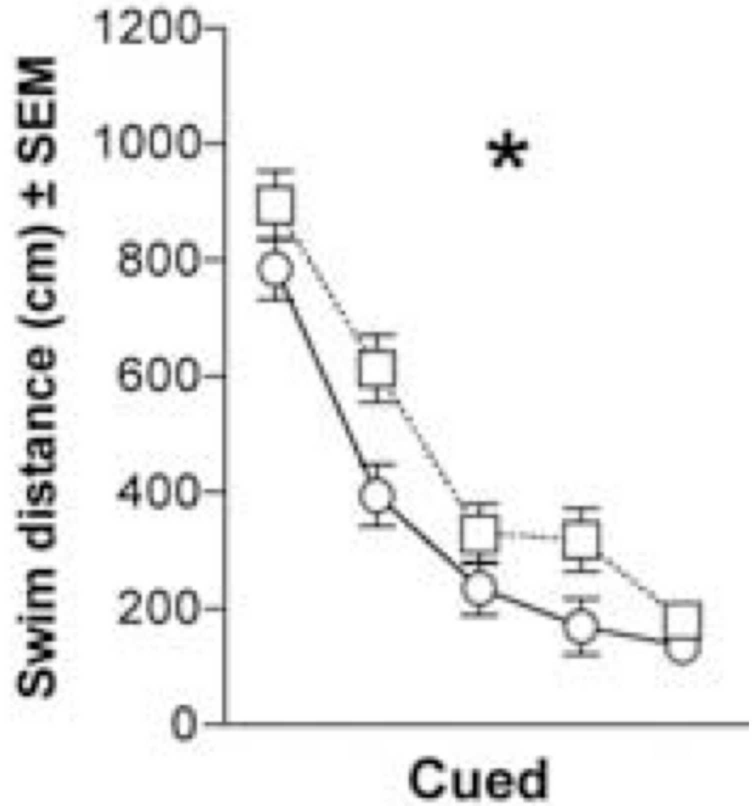
Main Effect or Interaction?



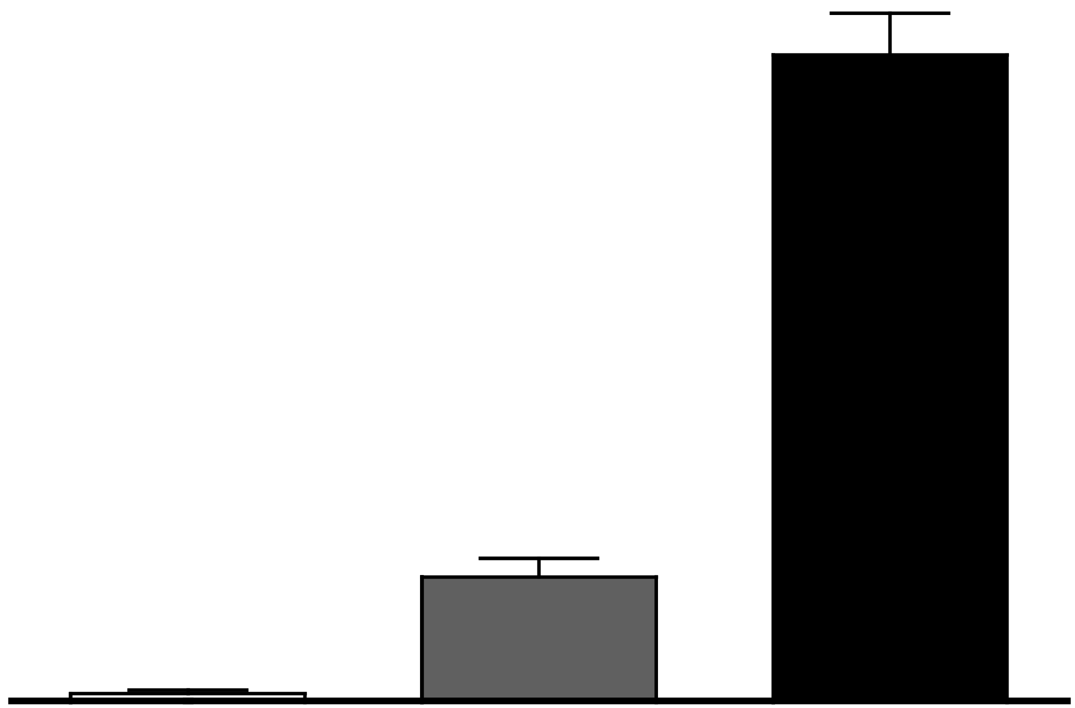
Main Effect or Interaction?



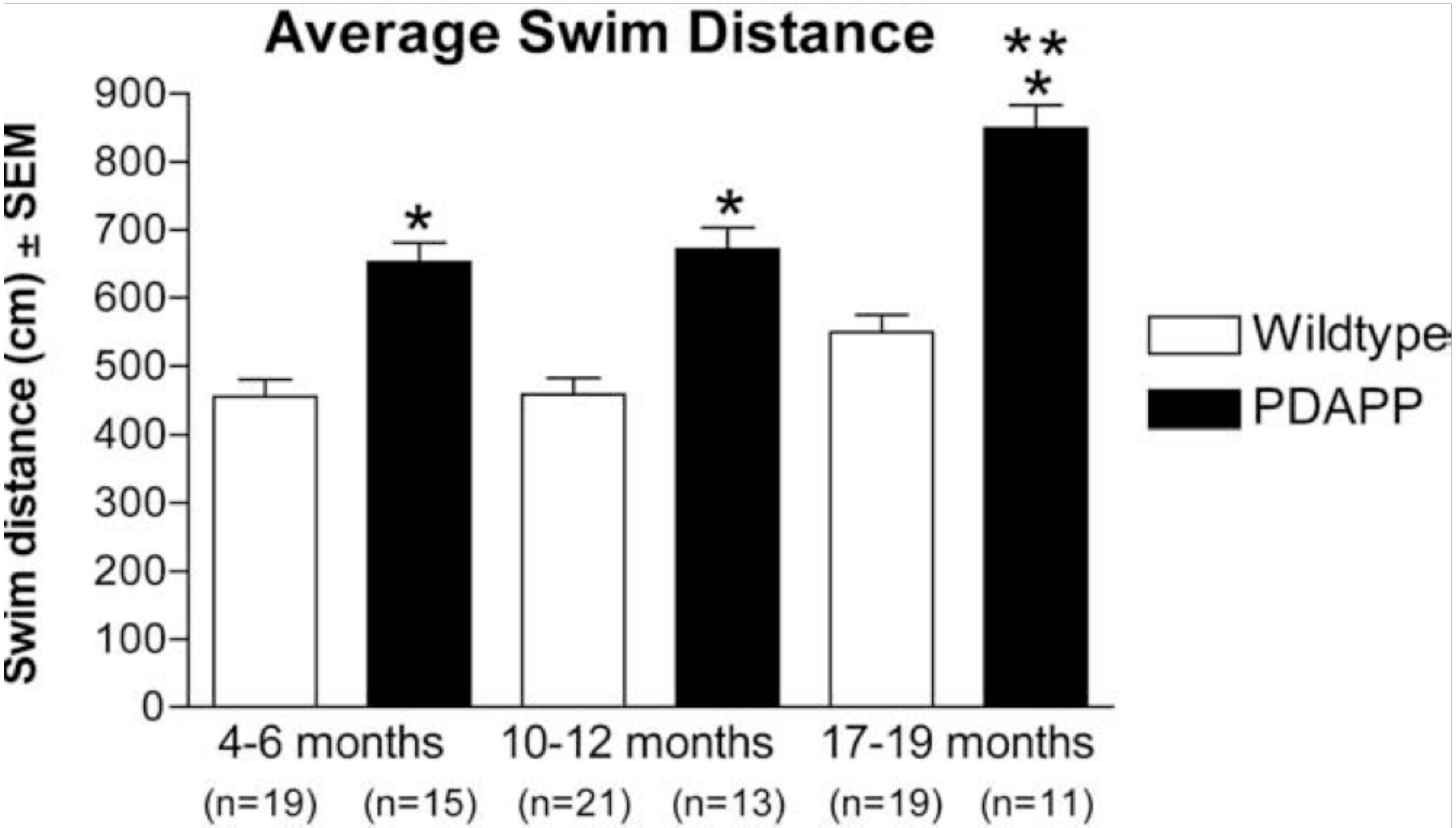
Main Effect or Interaction?



Main Effect or Interaction?



Main Effect or Interaction?



Ethics in Research

2 big ethical no-no's:

- Nuremberg War Crimes trial
  - use of captive humans in medical experiments (40's)
- Tuskegee Syphilis Study
  - withheld treatment for syphilis (50's-60's)

both led to the consensus that humans cannot be used as "guinea pigs"

Ethics in Research

By the 1990's, society witnessed a reversal:

- cancer / AIDS patients demanded to be given experimental treatments
  - protesting a system that had been designed to "protect" them

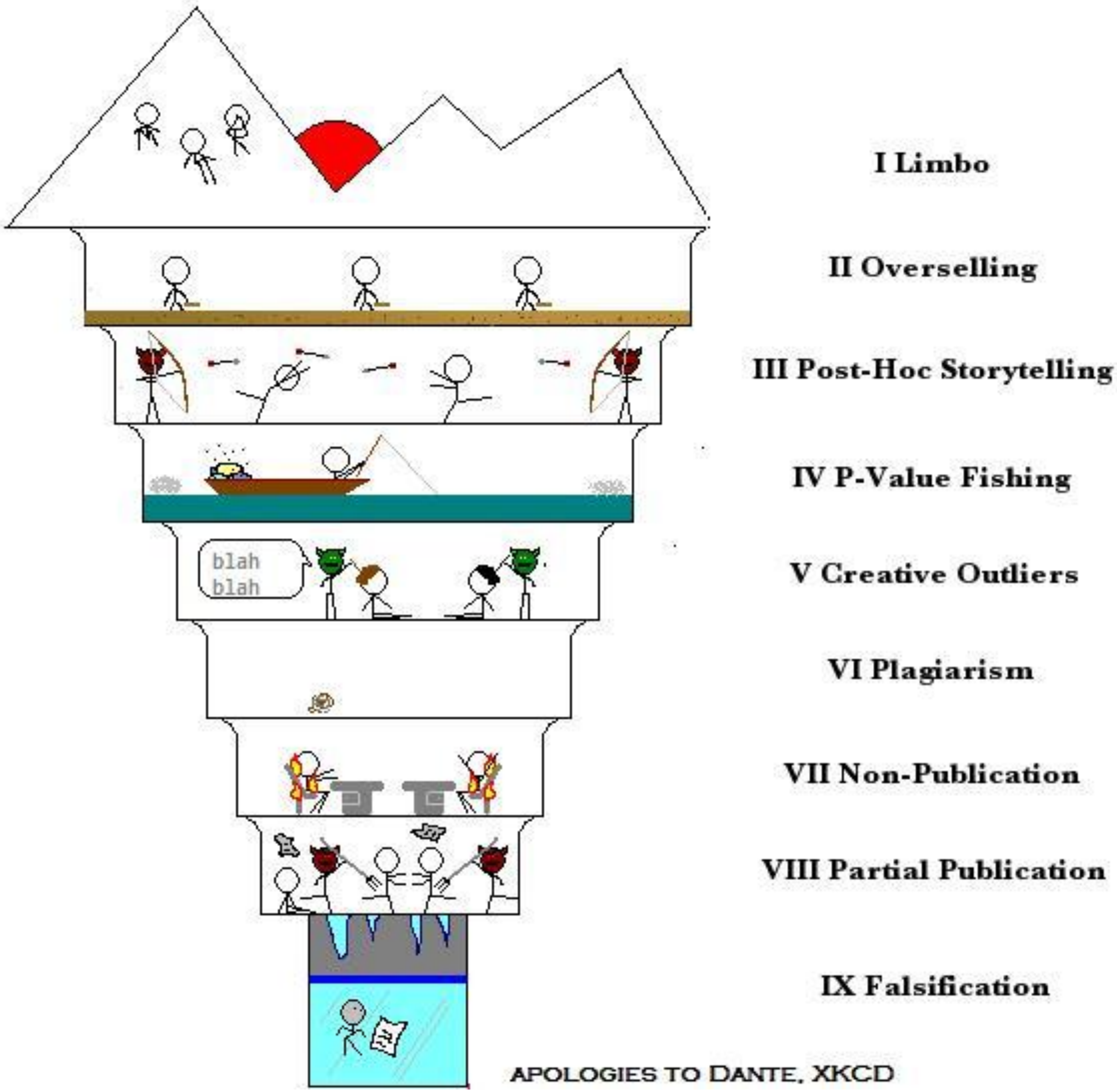
2 extreme views:

- protect against human experimentation at all costs
- allow anyone who is willing to be experimented on

Ethics in Research

The modern ethical system

- enforced by the Institutional Review Board (IRB)
  - voluntary participation - not coerced
  - informed consent - fully informed about procedures before-hand
  - no risk of harm (physical or psychological)
  - privacy:
    - confidentiality - no identifying info given out
    - anonymity - even the researchers can't ID
      - difficult for longitudinal studies
  - right to service - "no-treatment" groups


I Limbo
II Overselling
III Post-Hoc Storytelling
IV P-Value Fishing
blah blah
V Creative Outliers
VI Plagiarism
VII Non-Publication
VIII Partial Publication
IX Falsification
APOLOGIES TO DANTE, XKCD

The Write Up / Presentation

- audience understanding is the goal
- use informative, catchy titles to hook the audience
- do a spelling AND grammar check AND E.of.S
- define any abbreviation or acronym when 1st used
- preemptively address any and all questions that others may have
- be transparent!
- have others read it
- results - the data are the data (don't massage)

Advice for a Young Investigator

- **On Writing Scientific Papers:**
  - have something to say
  - say it
  - stop once it is said
  - give the article a suitable title

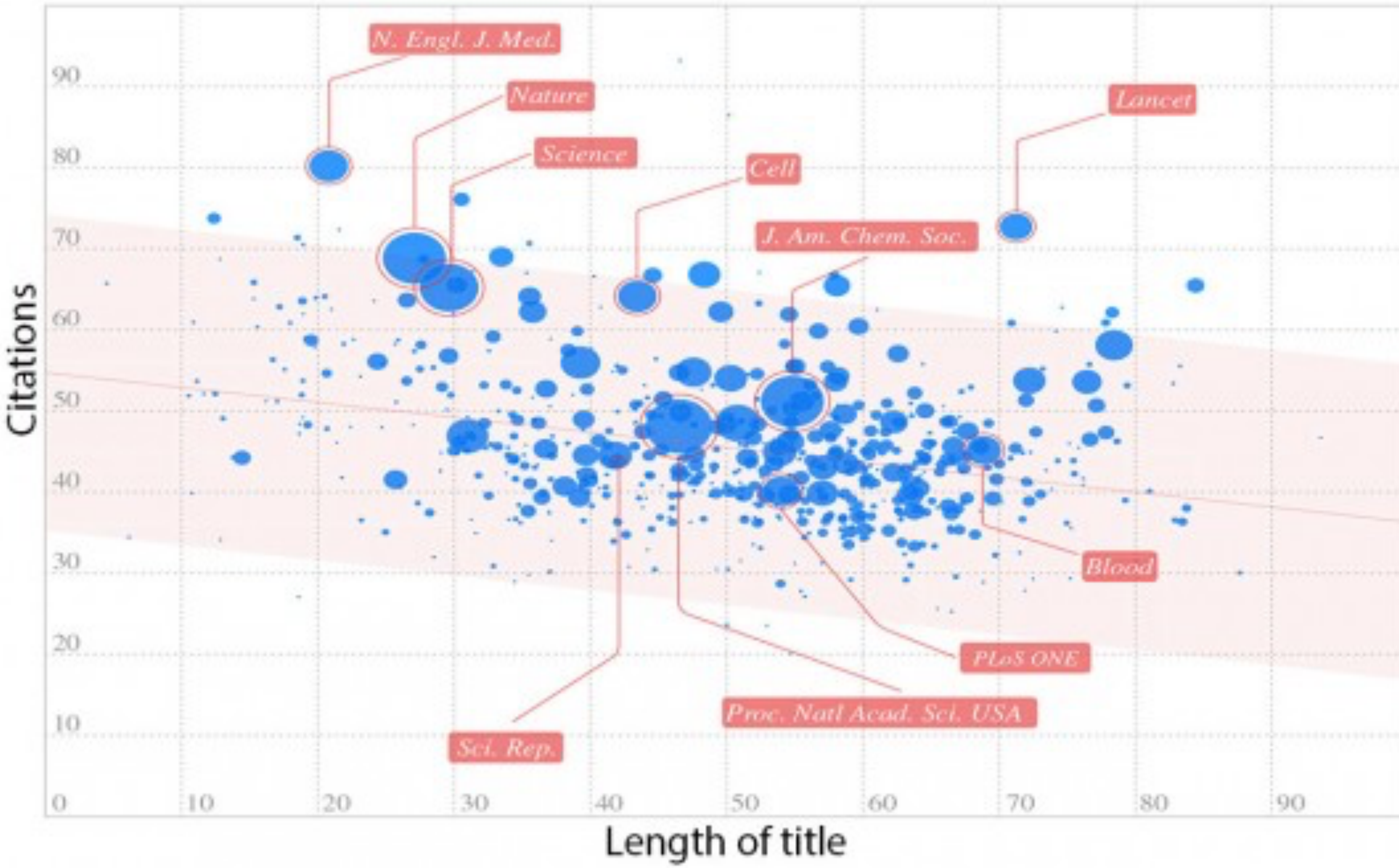
Advice for a Young Investigator

- references - it is customary to trace the history of a problem
- overlooking references will be repaid in kind
- the right to make mistakes is reserved for the famous
- be respectful and unassuming with the mistakes of our scientific forefathers
- deal with errors of fact, observation, or interpretation
- when unjustly attacked, defend yourself, but nobly
- fully outline your methods - **be clear**
- explain observations in a clear, concise, and systematic way
- summarize briefly the way in which your data contribute to science
- call reader’s attention to problems that remain to be solved

Advice for a Young Investigator

- no matter how exact and minute a verbal description may be, it is always less clear than a good, simple illustration
- style - avoid "rhetorical flatulence"
  - <u>whatever is good, if brief, is twice as good</u>
- one should speak as in a will - fewer words mean less litigation

- **<u>FINALLY</u> - an unpublished experiment is an experiment that never happened**



Papers with shorter titles get more citations

The Abstract

- only present the most pertinent information
- it should be clear and concise
- 1 or 2 most important findings
- use direct, active-voice sentences
- avoid unnecessary details
- avoid abbreviations
- no references

The Abstract

- Introduction: 2-3 sentences in present tense
  - background information
  - purpose / goal
  - identify hypotheses

- Method: 1-2 sentences in past tense
  - explain procedures used to test your hypothesis

The Abstract

- Results: 3-4 sentences in past tense
  - state the 1 or 2 most important results
- Discussion: 1-2 sentences in present tense
  - “these results suggest / demonstrate / etc”
- Conclusions: 1-2 sentences in present tense
  - avoid hype / speculation

The Talk

- Respect audience's intelligence, but AIM LOW
- intelligent non-scientists / bright high-schoolers
- no jargon
- Be smooth / rehearse <u>out loud</u> many times
- listen to what you are saying
- memorize the order of IDEAS, but not specific words
- video / audio recording - work on vocal inflections ("questioning" tone, etc) and, you know, "fillers"

The Talk

- memorize EXACTLY the 1st and last things you will say
- begin with a global/general statement that everyone will agree with
- tell a story with a logical flow (beginning / middle / end) using 1st person
- End with a general conclusion or significant statement (do NOT just end with your last slide and say "That's it! questions?")
- an acknowledgment slide
- then thanks for the invite

The Talk

- Delivery:
  - know your slides
  - segue into the next slide - prepare audience for what's coming next
  - use metaphors, ask rhetorical questions, propose dichotomies
  - say the same thing several times in different ways (use synonyms)
- Tempo: ~1 slide every 3 minutes
  - rushing to finish (or going over) is <u>disastrous</u>
  - interact with audience / change pace if needed

The Talk

- Dealing with questions
  - answer what was asked
  - ask for clarification if necessary
  - pause for a drink if necessary
  - don’t be afraid to say “I don’t know....”
- Slides:
  - use simple colors / consistent themes
  - use visual props effectively
  - show several simple figures rather than 1 complex figure
  - make figures as big as possible / remove extraneous info