IAT 804
Foundations of Research Design for Human-Centred Design of Interactive Technologies

Instructor: Dr. Alissa N. Antle

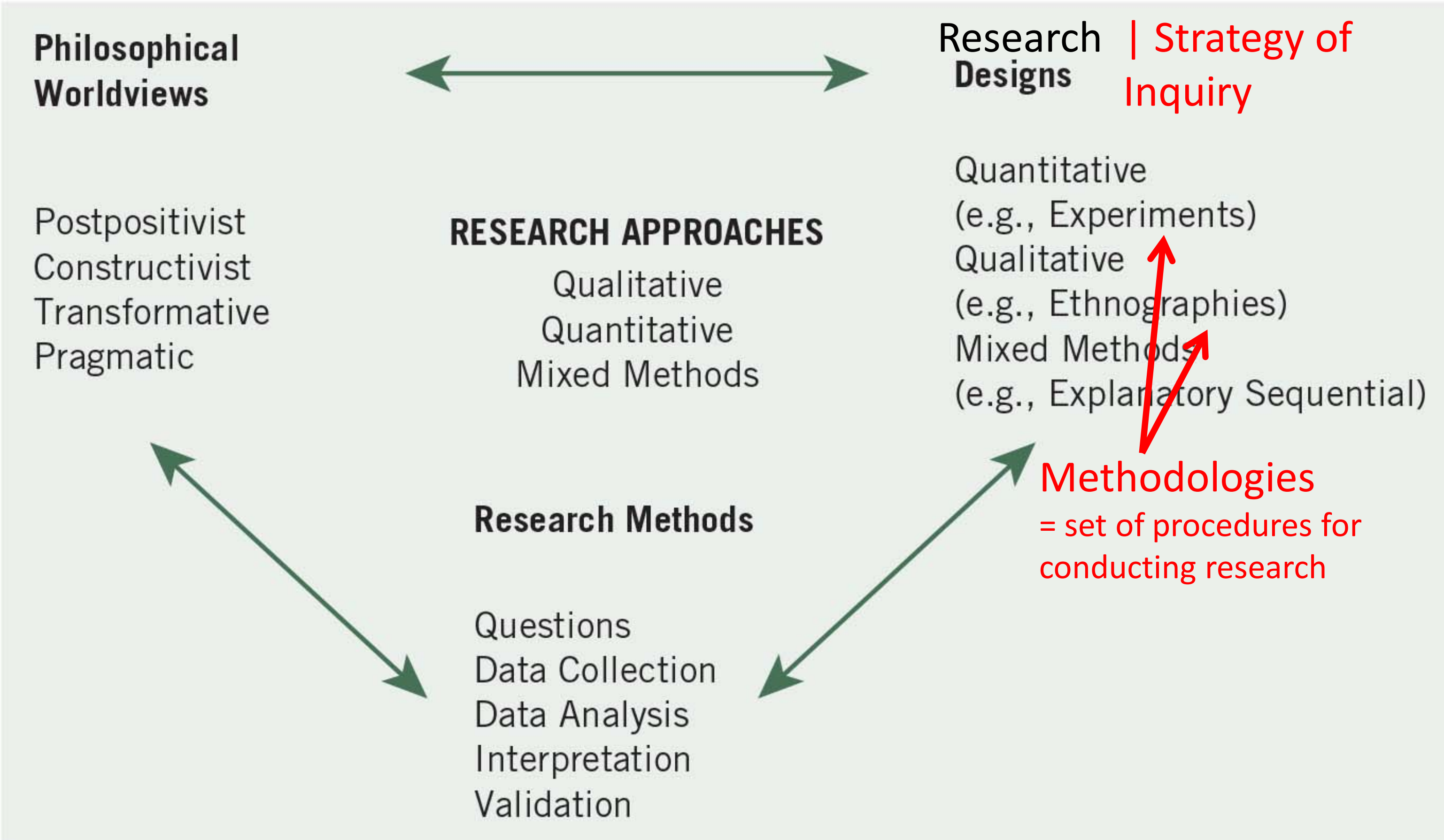
TA: Alex Kitson

Lecture 5 Fall 2019

Today

- Hand in Assignment 2: Short Paper
- Clarify strategies of Inquiry vs Methodology
- Assignment 3A: An Experiment
- Experiments
- Mouse Experiment Class Activity

Figure 1.1 A Framework for Research—The Interconnection of Worldviews, Design, and Research Methods



Assignment 3A Experiment

Work to date

- P: People
- A: Activity
- T: Interactive Tech (System)
- C: Context (Environment)s
- Research Question
    - Quantitative question

    e.g. What is the impact of X on Y?

Methodology Quantitative Experiment

- Scientific method/empirical observation
- In HCI: causes and effects explain PATC relationships
- Research question
- Hypotheses
- Construct
  - Operational definition
- Variables

Two main types of variables

- Independent variable
  - **Causes** an effect
  - Researcher can "manipulate" it (change it in the experiment i.e., has control over it)
  - Can be about any of P,A,T,C
    - E.g. People -- Group 1: Male, Group 2: Female
    - E.g. Interface – Design 1, Design 2, Design 3
- Dependent variable
  - Is **affected** by cause – varies as a result
  - Researcher observes/measures through data collection
  - Often about "interaction" or "performance"

Theory + Positioning (in) → Generate Hypothesis

- Theory IN → create hypothesis
- Theory → identifies variables AND relationships between them
- Must be causes and effects
- Hypothesis is a statement of these relations

Theory + Deductive Reasoning → Hypothesis Testing

- Deduction … if premises true, then conclusion is true

  General premise from theory: All crows are black

  Specific premise from observation (data): I see a crow

  Inference → Conclusion: The crow I see is black.

  Hypothesis: If I observe a crow, it will be black.

  Hypothesis testing: Does data (observation of crows) "fit" prediction of hypothesis (observed crows should all be black)

Science Experiments + Hypothesis Testing

- Deduction → create and test a hypothesis
- Test hypothesis: observe | prediction holds?
- Scientific experiment assumptions
  - Hypothesis -> causes and effects
  - Hypothesis must be falsifiable
  - Must be able to control/manipulate cause
  - Must be able to measure effect(s) w/ quant data

Summary so far …

1. Generate general hypothesis deductively
2. Identify variables
3. Determine prediction

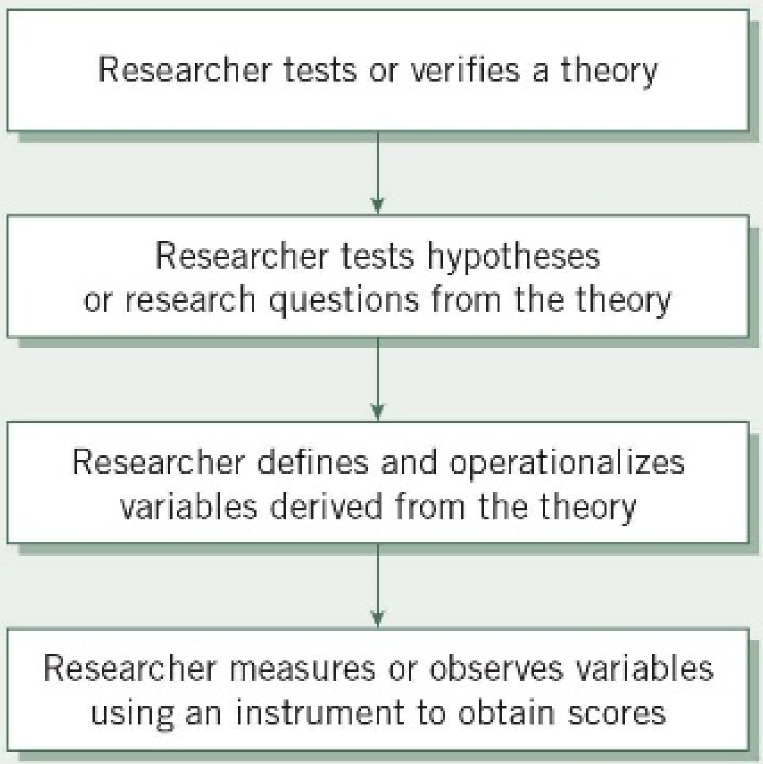
   What is the effect of changing one of the variables (Independent variable) on the other variable you expect to change (Dependent variable)?

4. Create specific hypothesis statement

   e.g. If (independent variable) (is changed by you), then the (dependent variable) will (describe the effect)

Summary …

Figure 3.4 The Deductive Approach Typically Used in Quantitative Research



Hypothesis must be Falsifiable

- Collect data to test hypothesis
- Call this evidence
- What can evidence tell us?
  - Recall: All crows are black – no
  - No crows are white – yes!

- Null hypothesis in experiment
- Cause does NOT impact effect
- Changing Independent variable (manipulated) does NOT impact dependent variable (data/measure)

Example: Generating a Hypothesis

- Theory: There are sex differences (male vs female) in many spatial abilities (e.g. spatial navigation).
- Motivation: Spatial skills predict success in science and engineering. We want to prepare young girls to succeed in science and engineering.
- Fact: Performance in first person shooter (FPS) video games is related to spatial navigation skills.

Example – variation 1: between groups design

- Deduction: Males have stronger spatial navigation skills. FPS require spatial skills. Males are better at FSP.
- RQ: Do males perform better than females in FPS games?
- Independent variable: group: male/female
- Dependent variable: Performance in FPS
- Hypothesis: If males and females play the same number of levels of a FPS, males will performance better.
- Null: There is no difference between males and females performance playing FPS games.
- Look for evidence – was there a M/F difference?

Back to Research Problem

- Motivation: Spatial skills predict success in science and engineering. We want to prepare young girls to succeed in science and engineering.
- How does this compare between groups (males vs females) design addresses research motivation?
- Try again

Example – variation 2 – within group (pre/post or repeated measures) design

- Deduction: Playing FPS improves spatial navigation skills. Women can play FPS. Women who play FPS will improve spatial navigation skills.
- RQ: Does training in a female-friendly FPS improve females' spatial navigation skills?
- Independent variable: Timing of spatial test
- Dependent variable: Test of spatial navigation
- Hypothesis: If females play 30 hours of a female-friendly FPS game over a month, then their spatial navigation skill after playing will be higher than before playing.
- Null: There is no difference in spatial test score before and after playing 30 hours of FPS.
- Look for evidence: Change in test score before/after playing.

Back to Research Problem

- Motivation: Spatial skills predict success in science and engineering. We want to prepare young girls to succeed in science and engineering.
- How does this repeated measures within group design addresses research motivation?
- What might have impacted results?

Strength of knowledge claim?

Confounding and Control Variables

Confounding variables

- Systematically interfere w/predicted cause and effect relationship
- Examples:
  - Expectation/motivation
  - Carry over (e.g. learning)
  - Fatigue
  - Maturation
  - Contamination

Controlling variables

Note about directionality

Non-directional hypothesis

- Gender **impacts** spatial skill

Directional

- Males are **better** than females at spatial navigation
- Females will **improve** in spatial navigation after playing FPS

Review Experiment Design

1. Generate hypothesis from theory
2. IVs – what are they? How to manipulate?
3. Construct/DVs – how to measure? Data type?
4. Confounding/Control variables

What else?

Next steps (part of a write up)

5. Materials/system – valid research instrument?

(A. Dix chapter on ways to evaluate system)

6. Participants – key characteristics that are relevant to hypothesis? How many? Where from? Ethical issues?

Split participants into groups?

Within groups

Between groups

Within vs Between group design

Independent variable	
Level 1	Level 2
P1	P1
P2	P2
…	…

Independent variable	
Level 1	Level 2
P1	P11
P2	P12
…	…

P1, P2, P3 etc are participants

More next steps

7. Procedure …

TASKS! (task order/carry over effects)

When is data collected?

Is there something about task should be consider as another independent variable?

More … 8. DATA

- Data collection for each DV
- Data Types
  - Nominal
  - Ordinal
  - Interval
  - Ratio

The big scary part

9. Quant Data Analysis

1. Descriptive statistics: Describe a single data set

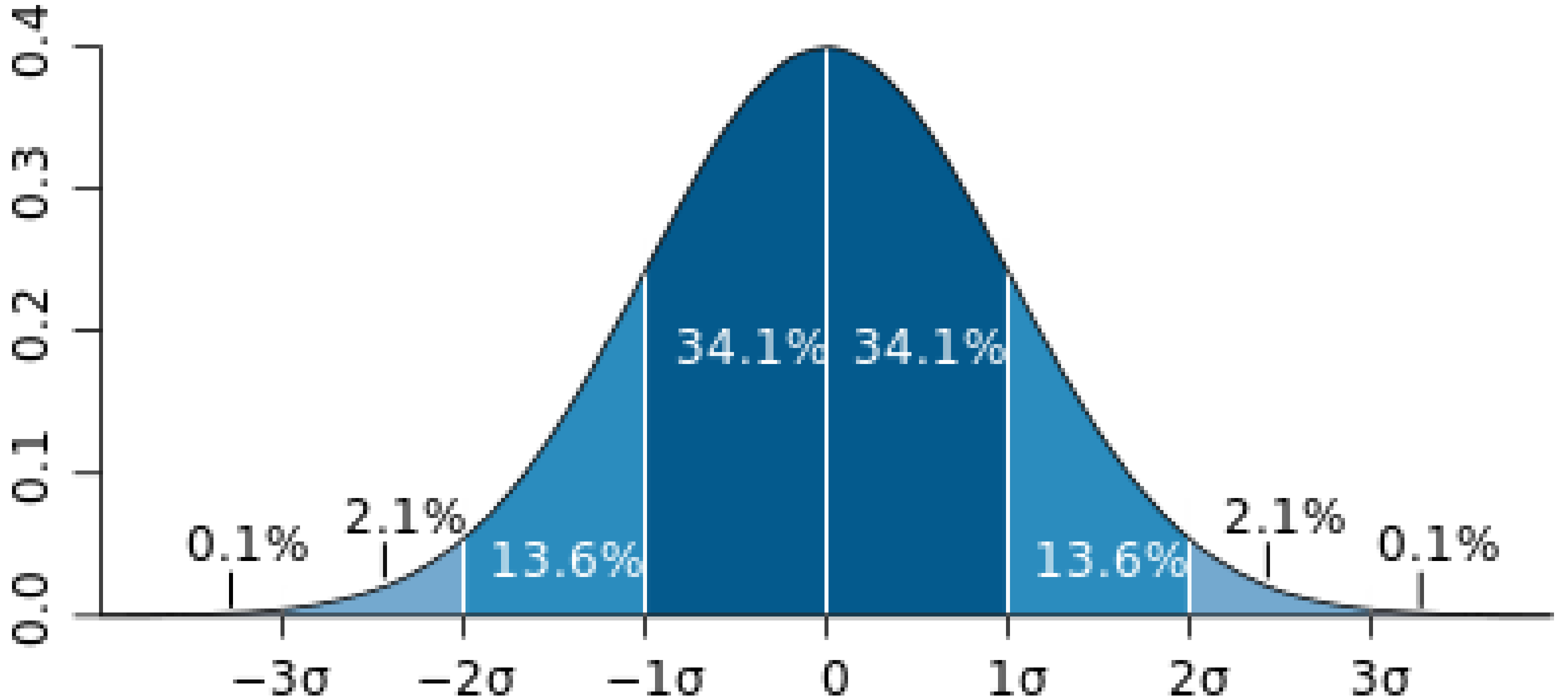
Mean – What is the average value?

Standard deviation – How much variation is there in data?

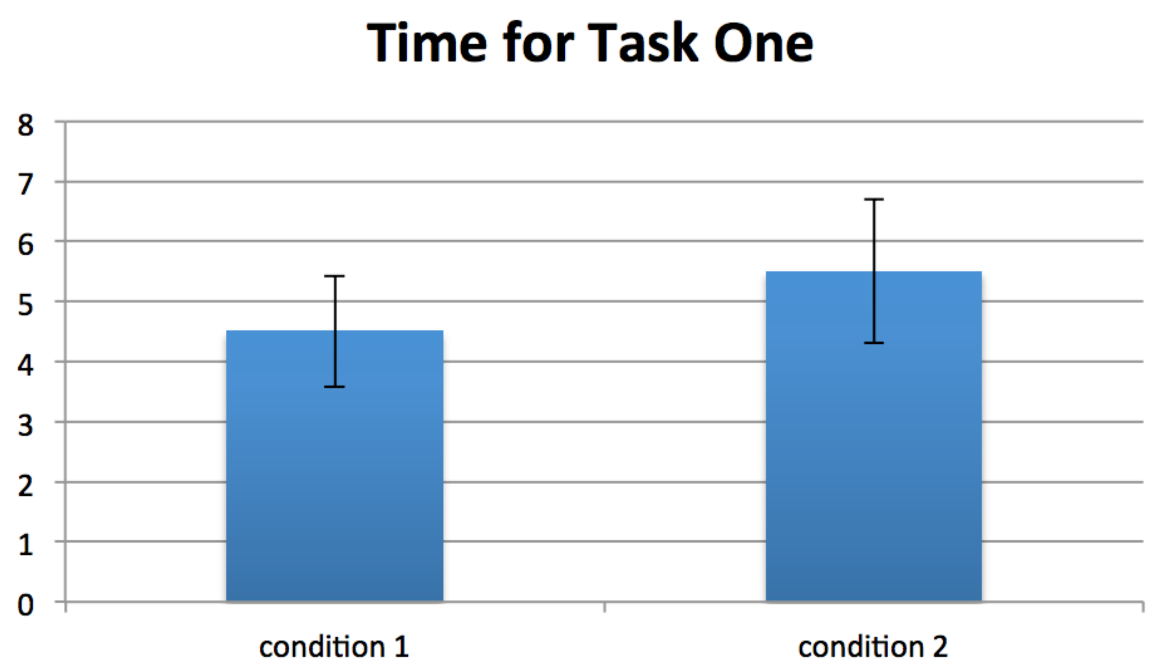
Why mean is not enough

Test score	Group 1	Group 2
	7	10
	7	10
	7	10
	7	8
	7	6
	7	8
	7	3
	7	4
	7	4
Average	**7**	**7**
Std Dev	**0**	**2.8**

Normal distribution

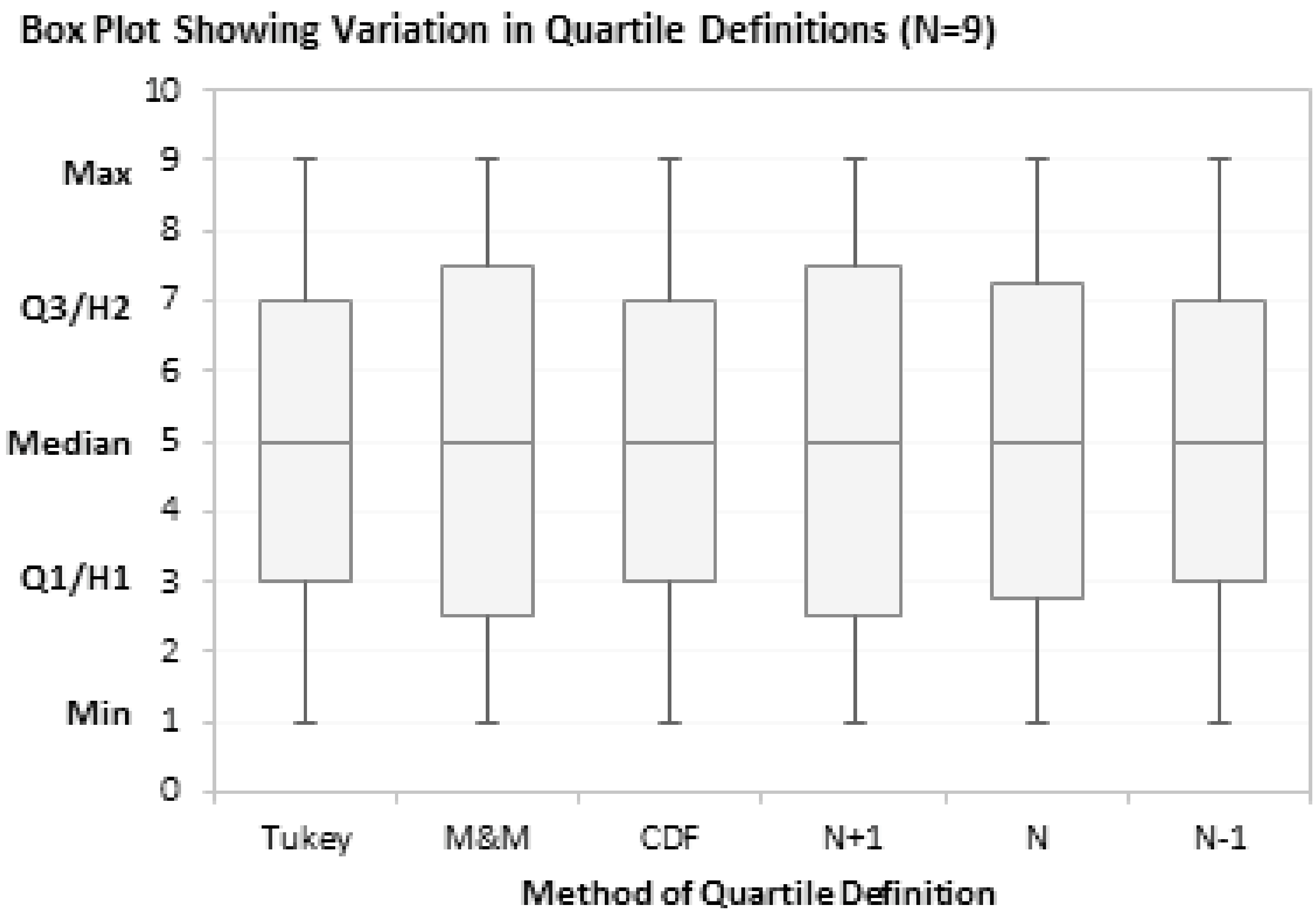


Bar Plot w/ Error Bars



mean = 4.5 +/- 0.9 mean = 5.5 +/- 1.2

Box Plot



Box Plot Showing Variation in Quartile Definitions (N=9)
10
Max
9
8
Q3/H2
7
6
Median
5
4
Q1/H1
3
2
Min
1
0
Tukey
M&M
CDF
N+1
N
N-1
Method of Quartile Definition

Hypothesis testing needs more!

- Hypothesis had levels of IV (called groups or conditions)
- Is there a statistical difference between conditions/groups?

- Null hypothesis – predicts no difference
- If you find a difference
    - You’ve found the one white crow! That is you’ve found evidence to support your original hypothesis.
- If you find no difference
    - Original hypothesis cannot be true (via deduction).

So we need Inferential Stats!

Inferential Statistics …

- Is there a statistically significant difference between IV levels/groups' data sets for DV(s)?
- OR is there evidence to support/refute hypothesis
- Based on likelihood of difference occurring by chance
- Tests consider: mean, variance, # participants, and how confident you want to be there is a difference.

Which test?

- How to pick test
- Data type(s)
- Assumptions
- Choose confidence level (the p value!)

$\alpha < .05$ → means that there is a 5% chance the difference you observed occurred by chance.

Also means -- 95% sure observed difference did not occur chance …

Type of IV → type of test

- One IV (two levels) & one DV → t-test
  - Within (repeated measures) or between (independent)?
  - Single tail (directional) or two-tailed (non-directional)
- One IV (3+ levels) → One way ANOVA
  - Within (RM) or between (indep)
- Two IVs → Two way ANOVA
  - “Factors” = number of IVs
  - Within (RM) or between (indep) or mixed

Randomization & Errors

- The idea of Randomization
- Systematic errors
- Random errors

Reliability

- A good experiment design can be repeated and will get same results.
- What might impact reliability?

Validity

- Threats to internal validity – things that interfere with IV causes effect we measure with DV
  - E.g. error in how we measure constructs (unreliable)
  - What might other threats be?

- Threats to external validity – how it generalizes beyond specifics of experimental study
  - What might impact generalizability?

Advanced: Other Things

- Assumptions
  - Normality
  - Homogeneity (between) or sphericity (within)
  - Else non-parametric tests
- Type I error
- Type II error
- Effect size

What does it all mean?

- Back to RQ! And/or Hypothesis

- Statistical evidence about the observations you made in the world – were they what you predicted based on hypothesis generated from theory?

- If not, why not? What does that mean for science/design/art knowledge?
- If so, what does that mean?
- What is the contribution?