Introduction to Research Methods in Education

Research Methods

- Why is it important to understand research methods for interdisciplinary researchers?
- Types of research
- How do you measure learning experimentally?
  - Within subjects design
  - Between subjects design
- What are you measuring?
- Important Statistical terms
- [other methods]

Research methods as boundary object in interdisciplinary research

- Study of cross-disciplinary research collaboration (Mercier, Penuel, Remold, Villalba, Kuhl & Romo, under review)

- The Bilingual Baby Project
  - 4-year longitudinal study
  - Neuroscience, cognitive science, sociology
  - How does growing up in a bilingual environment influence cognitive development & school readiness?
  - Biggest issues was sample selection (methods)

… [a child] was brought back for me to test, and I was putting [an ERP] cap on and then his mom said, well, he's used to this from going to the neurologist … and then she told me that he has epilepsy. And I thought to myself, well, okay, I'm wasting the next two-hours because I'm not going to be able to use the data….. A quantitative researcher would never include a child that—you wouldn't even waste the time and resources.

It's difficult because you do form close relationships with the families…. So it was then much harder for our group to eliminate them from the pool, which we didn't. We have followed up, we've gone and interviewed her, and it would be very interesting to see what kind of school readiness skills that child has and what kind of problems that mother has encountered. I know that they wanted children who had no other disabilities in order to focus on their language acquisition, but not all children are of this type

Two Types of Research

- Quantitative
  - Experimental
  - Surveys (usually)
- Qualitative
  - Biography, phenomenology, grounded theory, ethnography & case study
- Mixed methods
  - You can't account for context with numbers
 - The plural of anecdote is not data

What do you want your data & results to look like?

- Do you want to show learning, engagement, the process of the activity?

- Do you want to show that your tool works, or that it is better than an alternative?

- Do you want to describe, code, run statistics, present a case study?

- How will you design the study to get the type of results you want to present?

How do you measure learning experimentally?

Ecology lesson

- Aim: to teach 5 year olds about complex systems
- Ten 1-hour long sessions over 5 weeks
- "Embodied curriculum"
  - Technology; dancing; drawing
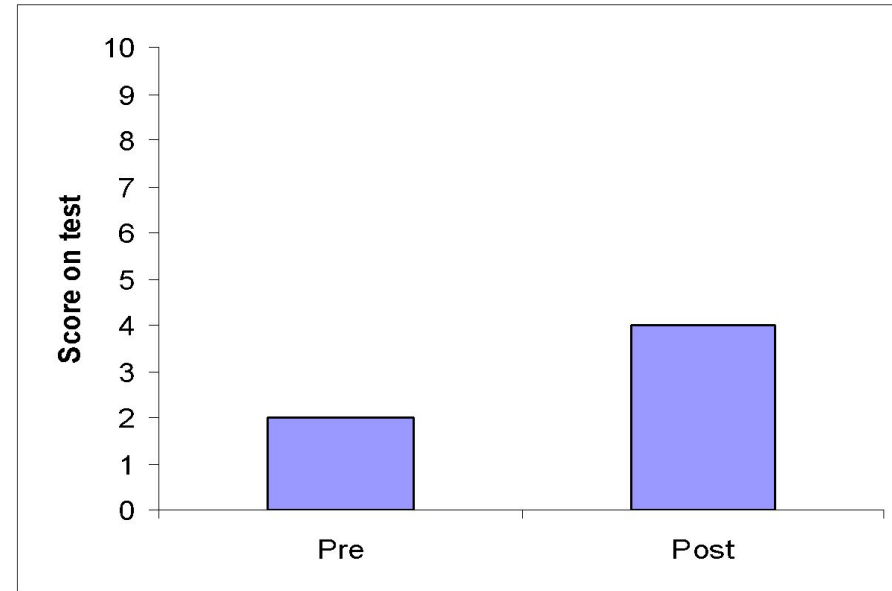
CLAIM: Embodied curriculum is a good way to teach complex systems

No way to know whether it was the curriculum, or just being taught that led to learning

How do you measure learning experimentally?
(and know that it's because of what you did…)

Pre/post test design (within subjects)

- Sequestered problem solving (SPS)
- Preparation for future learning (PFL)
- Free-write/free recall
- Delayed post-test
- Multiple baseline/single case design

How do you measure learning experimentally?
(and know that it's because of what you did…)

Control design (between subjects)

- Only some of the participants receive intervention & compare post-test scores between control & experimental groups

- Compare 'experimental group' with previous groups who did not receive intervention

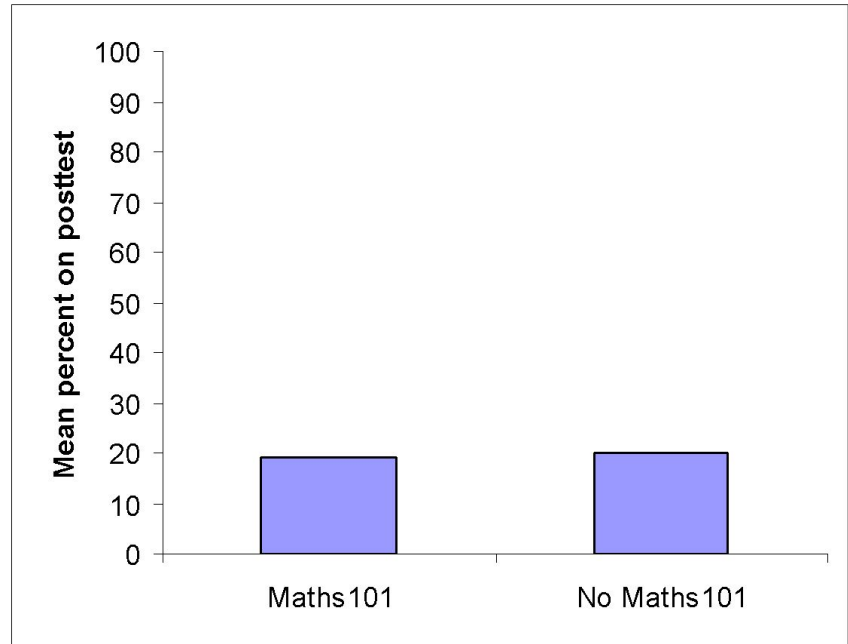- Randomized control trials (the medical model)

2x2 design

Godden & Baddeley (1975)

		Recall Condition	
		On Land	Under water
Learning condition	On Land	20	20
	Under Water	20	20

But what are you measuring?

Prerequisites for Maths Course

- Is maths101 necessary to pass stats202?
- Half of students had taken maths101
- All students take stats202
- Contrast outcomes

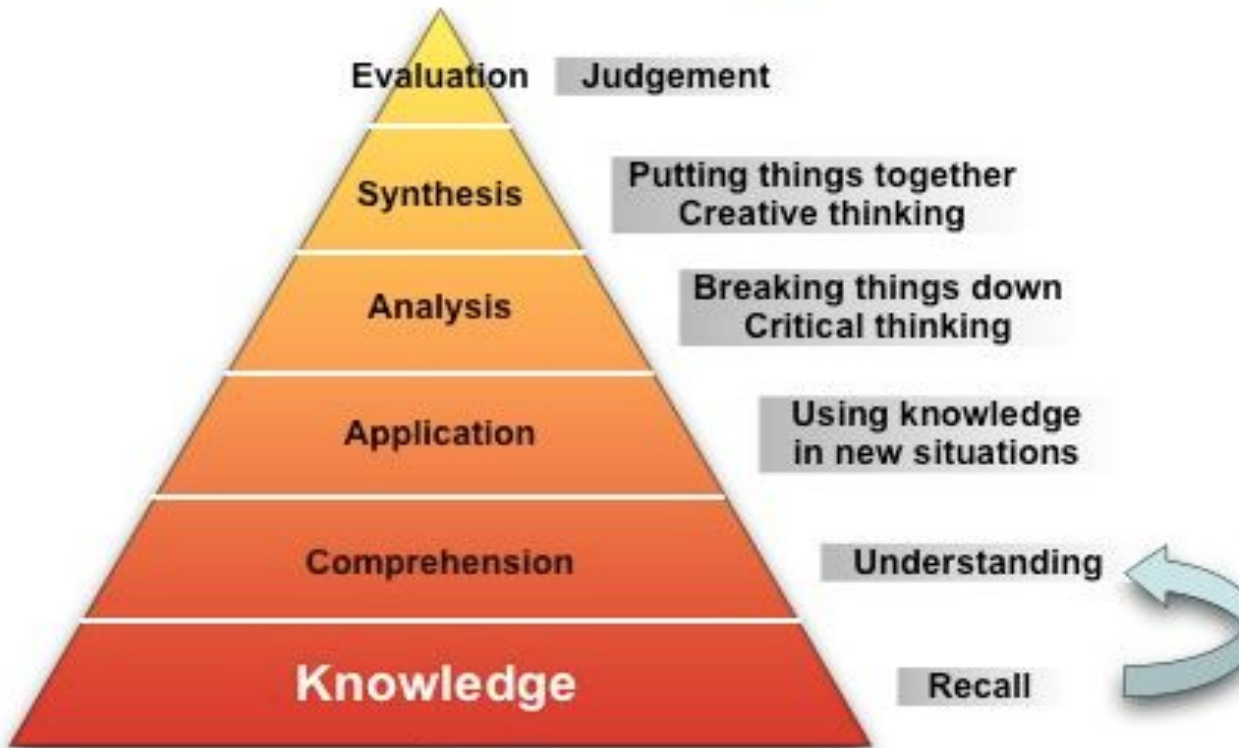CLAIM: No need to take maths101 before taking stats202

Floor effect: either the post-test didn't measure the content or very little was learned from stats202.

But what are you measuring…

- Is it valid?
  - Internal validity (is the effect caused by the IV)
  - External validity (would it replication beyond the sample)
- Is it reliable?
  - test-retest reliability
  - Split-test reliability
- [piloting]

What sort of learning will you measure?

Bloom's Taxonomy for Thinking

Level	Description
Evaluation	Judgement
Synthesis	Putting things together / Creative thinking
Analysis	Breaking things down / Critical thinking
Application	Using knowledge in new situations
Comprehension	Understanding
Knowledge	Recall

Important statistical ideas

- Independent variables
  - The thing you manipulate/control for
  - Main effects & Interaction effects between IV
- Dependent variables
  - The outcome measure
  - Floor effects & ceiling effects
- Statistical significance
  - Usually <.05 for social science
  - Indicates whether the effect is genuine or due to chance

Important statistical ideas

- Level of measurement
  - Nominal
  - Ordinal
  - Interval (& ranking)
- Population & sample
  - Normal distribution
- Descriptive statistics
  - Means, standard deviations, standard errors,
- Parametric and non-parametric statistics

Assumptions for parametric statistics

- Level of measurement must be at least interval
- Sample is drawn from a normally distributed population
- Homogeneity of Variance
  - Variance of two samples is not significantly different
- Independence of scores

Questions?

Key things to look for:

- Are the *differences* between conditions significant:
  - T-tests, ANOVA, Chi Square
- Is there a *relationship* between variables?
  - Correlations (note: you can't tell causation from this)
  - Pay attention to r values (between 1 and -1).
- Which of the IVs cause the DV?
  - Regression Analysis (note: need very large sample size; controversial technique)

Survey design

- **In the past 24 hours, did you watch more than an hour of television programming, or not?** Yes/No

- **In the past 24 hours, did you read a daily newspaper, or not?** Yes/No

- **On a scale of 1 to 7, please rate**

 How satisfied were you with what you learned and the usability of the software?

(1) Agree strongly…….(7) disagree strongly

Survey design (things to remember)

- Is there only one question in each item?
- Pilot with a number of people – do they read the question the way it was intended?
- Are all your scales in the same direction (if not, reverse them before analysis)
- Do the answers match the questions?
- How will you make sense of the answers
    - What sort of analysis can you do on rating, frequency, open-ended items?
- Are particular answers 'socially desirable?'