













- ♦ All center around one thing
- What happens when learners use your system?
- There's a lot of details beyond that, of course...

Empirical methods What kind of study should I run? What data should I collect? What should I do with that data?

What kind of study should I run?

Many potential types of studies

Types of studies

- Experiments
- Quasi-experiments
- Single-condition study
- Design experiments/design research
- Case studies
- Ethnography
- ✤ Instructor-led role playing

"Fixed Method" studies

- Experiments
- Quasi-experiments
- Single-condition study
- Design experiments/design research
- Case studies
- Ethnography
- ♦ Instructor-led role playing

More conclusive, more objective

 There's a place for all of these methods, but the evidence from fixed method studies is considered more conclusive



- Instructor-led role playing
- The authors claim that this is a really good method
- ♦ I am *not* a fan
- This method is very easy to screw up, and very expensive to conduct
 - Testing method and assessment are both very difficult to standardize
 - Can you react the same way every time?
 - Can you assess the same way in real-time, without taking personality, likability, etc. into account
- Only military labs/contractors with huge budgets use this method













- Resentful Demoralization Participants in the control condition see that they get a less fun system, become angry, and work less hard (a good reason not to give a game at the same time and in the same room as some more boring curriculum)
- Compensatory Rivalry ("John Henry effect") Participants in the control condition decide that they want to beat the participants in the experimental condition

Things to beware

 Hawthorne Effect – the sheer novelty of the serious game makes it seem better to the subjects (over a brief period of time) than it really is













- Measuring factual knowledge is good
- Measuring ability to apply it is better

Even Better tests...

- Measure the knowledge that the student should have gained by using the system, in a concrete fashion
- Measuring factual knowledge is good
- Measuring ability to apply it is better
- Measuring ability to apply it 2 weeks later is even better
 - Is learning maintained?





























































The End

Now get out there and evaluate!