



Empirical Methods

Bogdan Vasilescu

08-30-2018

(many slides courtesy of Jim Herbsleb)

Today's Agenda

- Introductions
 - Who are you?
 - What is your research?
 - What would make this course valuable to you?
- Why empirical methods?
- Research designs
- One way of thinking about how methods are related
- Course overview

Today's Agenda

- Introductions
 - Who are you?
 - What is your research?
 - What would make this course valuable to you?
- Why empirical methods?
- Research designs
- One way of thinking about how methods are related
- Course overview

A Variety of Reasons for Doing a Study

- Observation
 - E.g., a better understanding of how software engineers/designers/... work
- Identification of problems with state-of-the-art
- Evaluating a new tool/technique
 - E.g., evidence that approach A is better than B
- ...

How to validate your claims?

A Variety of Types of Questions

- What's going on here?
 - What does an example look like?
 - How are several instances same/different?
 - What things/events co-occur?
 - Do X and Y have a causal relation?
 - Does the effect of X on Y vary depending on the value of Z?
 - ...
- How to approach an answer?**

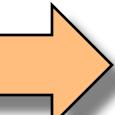
Aside: Logic of Causality

- In a sense, the logic of hypothesis testing is flawed
 - If X, then Y (if theory true, hypothesis must be true)
 - Observe that Y
 - Therefore, X
- Example
 - If my technology is effective, people using it will perform better
 - People using it perform better
 - Therefore, my technology is effective
- Fallacy: affirming the consequent
- Problem of confounding – multiple possible causes

Empirical Methods

- Used in many forms and phases of research
 - Understand problem
 - Current practice
 - Demonstrate utility of solution
- Selection of methods depends on
 - State of knowledge
 - Question researcher is asking
 - Nature of contribution
- **Each method has its own standards and techniques for rigor**

Today's Agenda

- Introductions
 - Who are you?
 - What is your research?
 - What would make this course valuable to you?
 - Why empirical methods?
 - Research designs
 - One way of thinking about how methods are related
 - Course overview
- 

Research Designs

- Step 1 to **design** empirical research: adopt a general (and guiding) **framework**
- Three main frameworks:
 - **Qualitative**
 - **Quantitative**
 - **Mixed Methods**
- Qual. vs Quant.:
 - Not dichotomies
 - Rather, different ends on a **continuum**

Qual. vs Quant.

- Often:
 - words (qual) vs numbers (quant)
 - open-ended questions (qual interview questions) vs closed-ended questions (quant hypotheses)
- More completely, differences in:
 - **philosophical** assumptions (e.g., postpositivist, constructivist)
 - strategies of **inquiry** (e.g., experimental, ethnographic, mixed methods design)
 - research **methods** (data collection and analysis)

Philosophical Worldviews

- Postpositivism
- Constructivism
- Advocacy/Participatory
- Pragmatism

Positivist View

- Traditional form of research, also referred to as the *scientific method*; or *empirical science*; or *postpositivism*
 - More often quant. research than qual.
- Never absolute truth of knowledge
 - Can't be “positive” about our claims of knowledge when studying human behavior and actions
 - Do not prove a hypothesis; instead, fail to reject the hypothesis

Positivist View (2)

- Deterministic philosophy: causes probably determine effects or outcomes
 - Research seeks to identify the causes that influence outcomes (e.g., experiments)
- Reductionistic in nature:
 - small, discrete set of ideas to test, e.g., variables that comprise hypotheses and research questions

Positivist View (3)

- Knowledge develops through careful observation and measurement of the objective reality “out there” in the world.
- Laws or theories govern the world
 - need to be tested, verified, refined so that we can understand the world
- Accepted approach: begin with a theory → collect data that either supports or refutes the theory → make necessary revisions → perform additional tests

Social Constructivist View

- Typically qual research
- Individuals develop subjective meanings of their experiences, directed toward objects or things
 - Meanings are varied and multiple
 - Meanings are not imprinted but rather formed through interaction with others (hence social constructivism)
- Researchers look for complexity of views, not narrowing meanings into a few categories
 - The goal is to rely as much as possible on the participants' views of the situation being studied.

Social Constructivist View (2)

- Researchers want to make sense of (or interpret) the meanings others have about the world.
 - generate or inductively develop a theory rather than start with one

Advocacy/Participatory View

- Research inquiry needs to be intertwined with politics
 - the research contains an action agenda for reform that may change the lives of the participants
 - focuses on the needs of groups and individuals that may be marginalized or disenfranchised
 - addresses important social issues of the day, such as empowerment, inequality, oppression
- It is practical and collaborative, i.e., completed *with* others rather than *on* or *to* others

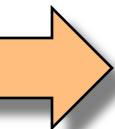
Pragmatic View

- Arises out of actions, situations, and **consequences** rather than antecedent conditions (as in postpositivism)
- Instead of focusing on methods, researchers emphasize the research problem and use **all approaches available** to understand it
 - Typical for mixed methods studies

Which do you subscribe to?

Positivist (or “Post-positivist”) <ul style="list-style-type: none">Knowledge is objective“Causes determine effects/outcomes”Reductionist: study complex things by breaking down to simpler onesPrefer quantitative approachesVerifying (or Falsifying) theories	Constructivist/Interpretivist <ul style="list-style-type: none">Knowledge is socially constructedTruth is relative to contextTheoretical terms are open to interpretationPrefer qualitative approachesGenerating “local” theories
Advocate / Critical Theorist <ul style="list-style-type: none">Research is a political actKnowledge is created to empower groups/individualsChoose what to research based on who it will helpPrefer participatory approachesSeeking change in society	Pragmatist <ul style="list-style-type: none">Research is problem-centered“All forms of inquiry are biased”Truth is what works at the timePrefer multiple methods / multiple perspectivesSeeking practical solutions to problems

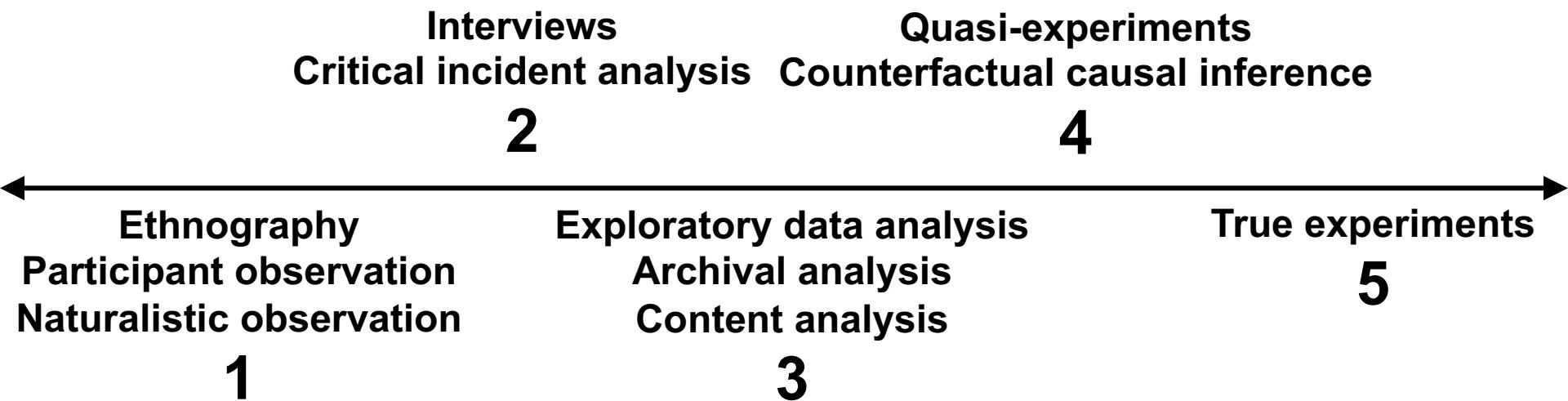
Today's Agenda

- Introductions
 - Who are you?
 - What is your research?
 - What would make this course valuable to you?
 - Why empirical methods?
 - Research designs
 - One way of thinking about how methods are related
 - Course overview
- 

One Way of Thinking About How Methods Are Related

- Roughly speaking, methods form a dimension
- Often most effective to use in combination or in sequence (rich to precise)
- Poles
 - Very little knowledge: “What’s going on?”
 - Extensive knowledge: “Precisely how is this different?”
- Generally, from qualitative to quantitative
 - From observation to experimentation
 - From richness to precision
 - From reliance on human interpretation to reliance on decision rules

Points on Empirical Methods Dimension



Example: Ethnography Rich precise

- What is it?
 - Immersion in the environment, group
 - Attempt to see the world through their eyes
- What questions can it answer?
 - How do the participants think about their work?
 - What are the problems?
- What makes it rigorous?
 - Constantly testing interpretations
 - Triangulation – multiple sources of data
- Example study?
 - Grinter: software architects

Example: Ethnography Rich precise

- What contributions can it support?
 - Problem as seen by persons of interest
 - Work in context
- What are its limitations?
 - May get trapped by participants' perceptions
 - Small samples, no causality
- What resources are needed?
 - Time and labor intensive
 - Access to right people
 - Willingness and ability to join group

Example: Interviews Rich precise

- What is it?
 - Structured interaction
 - Questions, answers, followup
- What questions can it answer?
 - Perceptions, opinions, processed observations
 - How things are done, exceptions, problems
- What makes it rigorous?
 - Preparation with well thought-out topics
 - Cross-validation in questions, interviewees, checking interpretations
- Example study?
 - Herbsleb & Grinter: software integration

Example: Interviews Rich precise

- What contributions can it support?
 - Nature of problem, as perceived, current process
 - Examples, exceptions, incidents
- What are its limitations?
 - Information is processed, filtered by interviewees
 - May be inappropriately biased by questions
- What resources are needed?
 - Willing interviewees, correctly positioned
 - Ability to sample all relevant perspectives
 - Preparation, follow-up

Example: Quasi-experiment Rich precise

- What is it?
 - Naturally-occurring differences
 - Examination of effects of variables in situ
- What questions can it answer?
 - What are the effects of introducing X?
 - What is the difference between X and Y?
- What makes it rigorous?
 - Good quasi-control groups
 - Access to data for control variables
- Example study?
 - Wagstrom et al: effects of commercial participation on volunteers in open source

Example: Quasi-experiment Rich precise

- What contributions can it support?
 - Value of tool, method, process, training
 - Influence of context factors
- What are its limitations?
 - Never sure cause-effect relation is established
 - Relying on luck, that situation occurs naturally
- What resources does it take to do it well?
 - Sophisticated statistics (e.g., multiple regression)
 - Contextual knowledge of experimental situation

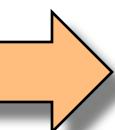
Example: True Experiment Rich precise

- What is it?
 - Comparison that is engineered
 - Random assignment of values of independent vars
- What questions can it answer?
 - Cause and effect
 - Size of effect, interaction of factors
- What makes it rigorous?
 - Well designed experimental and control conditions
 - Attention to reliability, validity
- Example study?
 - Dekel et al: pushing directive information

Example: True Experiment Rich precise

- What contributions can it support?
 - Value of tool, method, process, training
 - Influence of context factors
- What are its limitations?
 - Must be able to isolate critical variables
 - Limited by ability to create situations that manipulate people
- What resources does it take to do it well?
 - Access to appropriate subjects
 - Statistics, measurement instruments

Today's Agenda

- Introductions
 - Who are you?
 - What is your research?
 - What would make this course valuable to you?
 - Why empirical methods?
 - Research designs
 - One way of thinking about how methods are related
 - Course overview
- 

All Methods are Flawed

- This course teaches strategies to overcome weaknesses

Goals of This Course

- Better consumers of empirical research results
- See the variety of kinds of methods available, clear idea about when to use them
 - Direct you towards specialized courses
- Help you focus in your own research
 - Learn to plan, conduct, report on, and present empirical studies
 - Understand the key steps of a research project:
 - **formulating research questions**
 - theory building
 - data analysis (using both qual. and quant. methods)
 - building evidence
 - assessing validity

Activities

- Read method descriptions / how-tos
- Present and critique papers that use these methods
- Exercises where you get some experience with some of the methods
- Project on your own research topic
 - Review literature
 - Design and plan a study
 - Formulate research questions
 - Write as proposal

Format

- Seminar
 - 1 three-hour seminar per week
- Typical outline of one seminar:
 - Homework debrief
 - Lecture/discussion of a new method
 - In-class practice
 - Short (conference-style) student presentations of example papers
- Assessment:
 - 40% assignments (one each week, due before class)
 - 50% research project
 - 10% initial project description (proposal)
 - 10% final presentation
 - 30% final report
 - 10% participation and in-class presentations

Coordinates

- Web
 - <https://github.com/bvasiles/empirical-methods> (schedule)
- Canvas
 - <https://canvas.cmu.edu/courses/7434> (readings, assignments)
- Office hours (open door policy)

Schedule (draft)

Date	Topic
08/30	Introduction
09/06	Literature Review and Theory
09/13	Interviews
09/20	Grounded Theory
09/27	Surveys
10/04	Introduction to Measurement
10/11	Your Research Project Proposal
10/18	Experiments
10/25	Quasi-experiments
11/01	Time Series Analysis
11/08	Mixed-methods
11/15	Text Mining
11/22	No Class - Thanksgiving
11/29	Social Network Analysis
12/06	Final Presentations

Topics

- Overarching concerns
 - Epistemic base
 - Human subjects
- Research Methods
 - Case studies
 - Grounded theory
 - Experiments
 - Quasi-experiments
 - Mixed methods
- Data Collection
 - Interviews
 - Surveys
 - Archival data
- Data Analysis
 - Coding
 - Statistics (regression, time series)
 - Social Network Analysis
- Drawing conclusions, presenting results

References

- Dekel, U. & Herbsleb, J.D. (2009). Improving API documentation usability with knowledge pushing. Proceedings of the International Conference on Software Engineering, Vancouver, Canada, pp. 320-330.
- Grinter, R. E. (1999) "Systems Architecture: Product Designing and Social Engineering" in Proceedings of ACM Conference on Work Activities Coordination and Collaboration (WACC '99). San Francisco, California: February 20-22. 11-18.
- Herbsleb, J. D. & Grinter, R. E. (1999). Splitting the Organization and Integrating the Code: Conway's Law Revisited. In proceedings, International Conference on Software Engineering (ICSE), pages 85-95, Los Angeles, CA, May 16-22.
- Olson, G.M., Olson J.S., Carter, M.R., Storrosten, M. (1992). Small Group Design Meetings: An Analysis of Collaboration Human-Computer Interaction, 7(4) p.347-374
- Wagstrom, P., Herbsleb, J.D., Kraut R.E., & Mockus, A. (2010). The Impact of Commercial Organizations on Volunteer Participation in an Online Community. Under review.