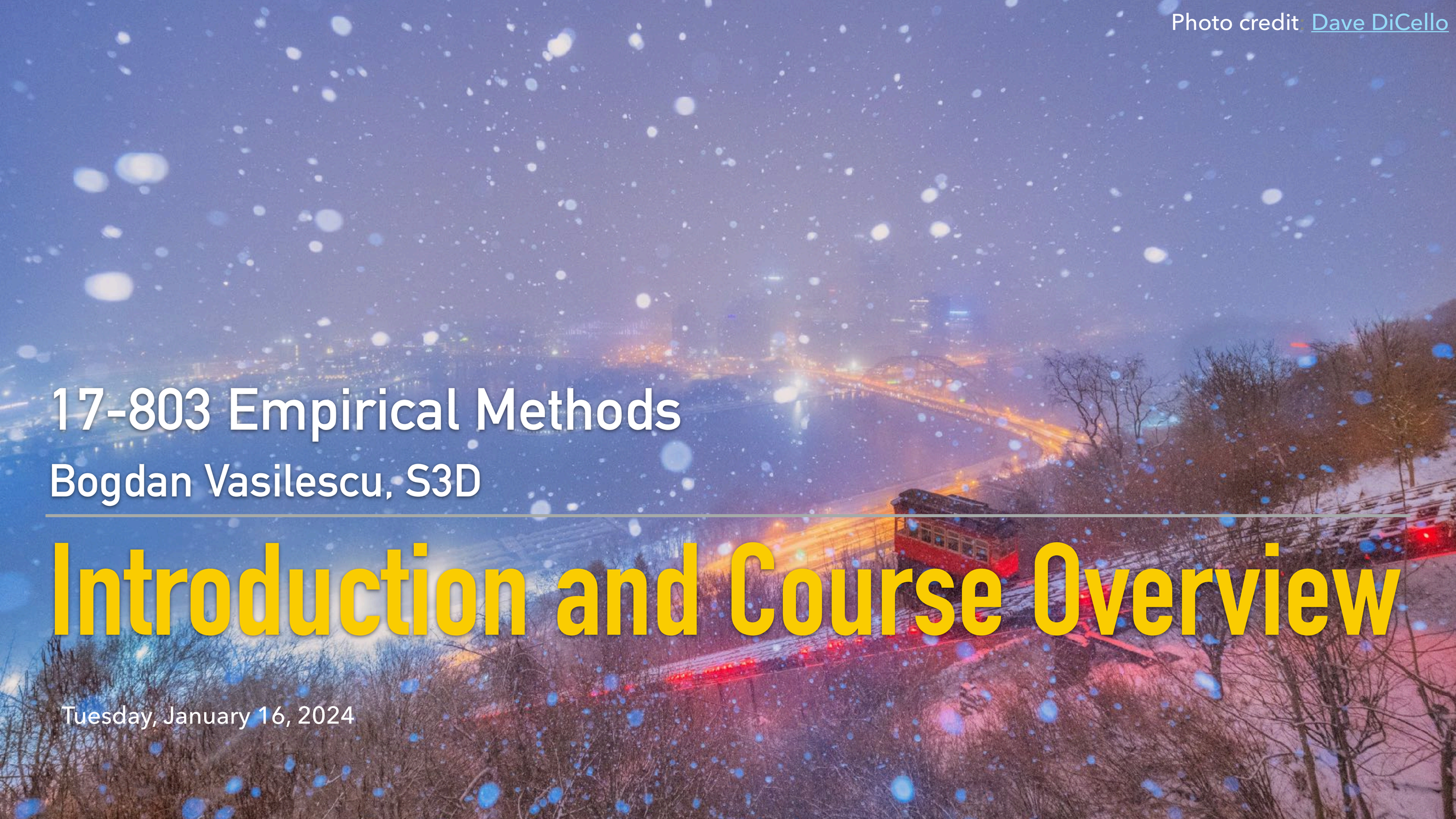


Photo credit: Dave DiCello

17-803 Empirical Methods

Bogdan Vasilescu, S3D

Introduction and Course Overview

Tuesday, January 16, 2024

Welcome to This Rare Course!

- The “modern” version of this course has only been offered a handful of times!
  - Several offerings of the “classical” version by Jim Herbsleb and Marcelo Cataldo in the past.
  - The current version starting Spring-2021 is much changed.
- Very few courses like this have been offered anywhere!
  - Notable exceptions (tell me if you know of more):
    - Steve Easterbrook & Barbara Barbosa Neves, “Empirical Research Methods for CS” at UofT (-2014)
    - Sara Kiesler (prev) / Laura Dabbish, “Applied Research Methods” at CMU (2015+) (Fall ‘19 syllabus)
    - Peggy Storey, “Empirical Software Engineering: Bridging Research and Practice” at UVic (2020+)
    - Shurui Zhou, “Empirical Software Engineering” at UofT (2021+)
  - Particularly notable exception:
    - Carl Bergstrom & Jevin West, “Calling Bullshit: Data Reasoning in a Digital World” at UW (2019+)

Outline for Today

- Introduction to the topic
- Introduction to each other
- Course logistics

Is This Your Research Plan?

- You are trying to understand how software engineers / designers / … work and what challenges they face.
- You have identified some challenges (e.g., working more productively) and are looking to inform the design of possible solutions.
- You have created a new algorithm / tool / process / programming language / system / … and are looking to evaluate it.
- …

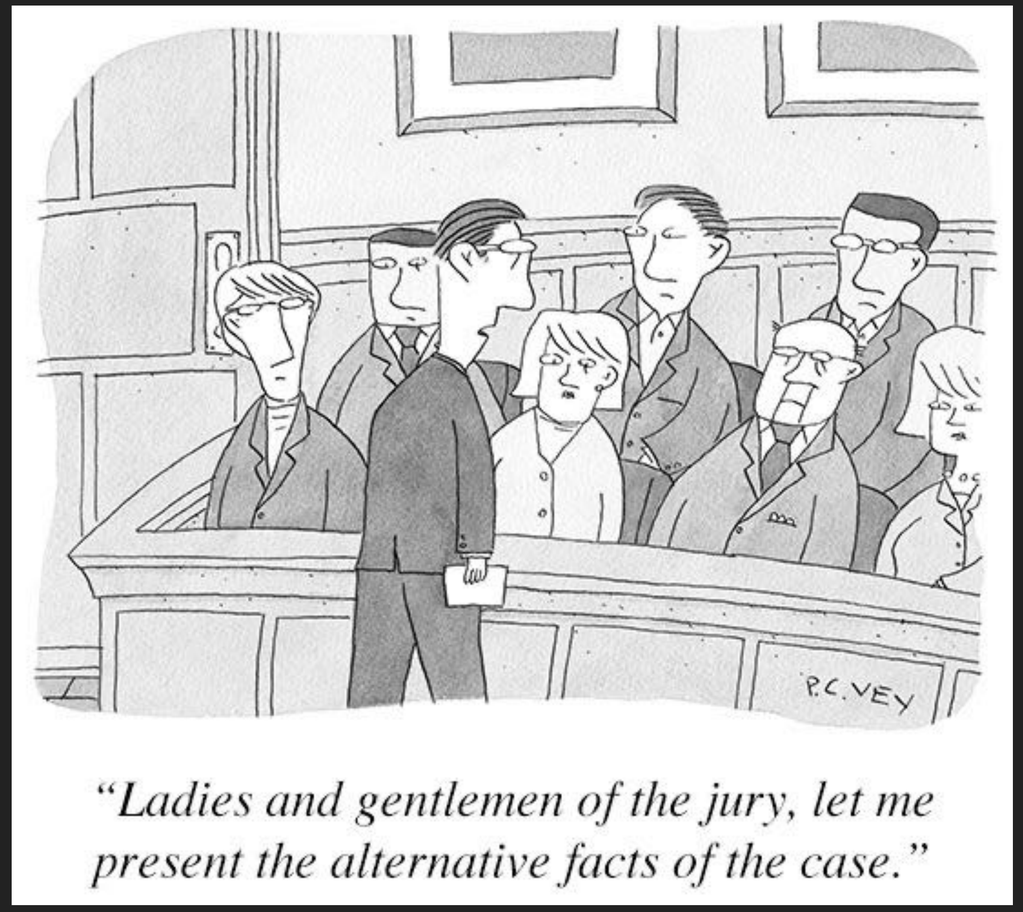
Is This Your Research Plan?

- You are trying to understand how software engineers / designers / … work and what challenges they face.
- You have identified some challenges (e.g., working more productively) and are looking to inform the design of possible solutions.
- You have created a new algorithm / tool / process / programming language / system / … and are looking to evaluate it.
- …

How to collect data?
How to analyze it?

How to collect evidence that approach A is better than B?

How to draw conclusions?
What constitutes sufficient evidence?



"Ladies and gentlemen of the jury, let me present the alternative facts of the case."

No matter your reasons for doing a study,

How To Validate Your Claims?

* Maybe the most common reason why papers get rejected

Summary of This Course: “It Depends”

This Is Where Empirical Methods Come in

- A diversity of methods are available.
- They're used in many forms and phases of research.
  - Understand problem
  - Current practice
  - Demonstrate utility of solution
- Each method has its own standards and techniques for rigor.

This Is Where Empirical Methods Come in

- A diversity of methods are available.
- They're used in many forms and phases of research.
  - Understand problem
  - Current practice
  - Demonstrate utility of solution
- Each method has its own standards and techniques for rigor.

Bad news:

- All methods are flawed

This Is Where Empirical Methods Come in

- A diversity of methods are available.
- They're used in many forms and phases of research.
  - Understand problem
  - Current practice
  - Demonstrate utility of solution
- Each method has its own standards and techniques for rigor.

Bad news:

- All methods are flawed

Good news:

- When applied correctly, they can be useful.

Selection of Methods Depends on:

- Approach to Research
- Nature of Contribution
- Specific Research Question
- State of Knowledge
- . . .

One Way To Think About Research Approaches and Methods

Research Exists Across Multiple Continua. Two Examples:

Research approaches

Inductive ⟷ Deductive

Research methodologies

Qualitative ⟷ Quantitative

The (Objectivist) Deductive Approach to Research

- Traditional form of research, also referred to as *the scientific method*, or *empirical science*
- Assumptions:
  - (1) there is an external reality (i.e., a real world that exists independent of the researcher)
  - (2) reality can be understood by collecting objective, unbiased data about that reality
- Knowledge builds by developing increasingly better understandings of, and insights into, the causal workings of the world

"You're fifty-seven years old. I'd like to get that down a bit."

The (Objectivist) Deductive Approach to Research

- Reductionistic in nature (top-down):
  - From general, abstract conceptualizations to observable and measurable data in some context
- Research questions: testing a cause-and-effect relationship underpinning a phenomenon
- Typical approach:
  - Start with some abstract conceptualization (theory)
  - Derive a hypothesis
  - Collect data, test the hypothesis
  - Findings may falsify, support, refine, challenge, or extend the conceptualization
  - Make necessary revisions, perform additional tests
- Common method: experiments

"You're fifty-seven years old. I'd like to get that down a bit."

The (Subjectivist) Inductive Approach to Research

- Assumptions:
  - (1) reality is socially and experientially constructed (i.e., reality exists because individuals and social groups share interpretations and understandings of reality);
  - (2) to understand reality, researchers need to explore the meanings constructed by individuals and groups.
- Constructivist in nature (bottom-up):
  - From specific data to a general or abstract conceptualization of the phenomenon (theory).
- Knowledge is subjective.
- Collecting data from a multitude of perspectives gives a richer and more nuanced understanding of the phenomenon.

"I'm neither a good cop nor a bad cop, Jerome. Like yourself, I'm a complex amalgam of positive and negative personality traits that emerge or not, depending on circumstances."

The (Subjectivist) Inductive Approach to Research

- Research questions: explore phenomena to increase our understanding of them.
- Typical approach:
  - Start with a desire to understand or explain a particular phenomenon.
  - Collect data of and/or about this phenomenon.
  - Search for patterns across the data to generate an understanding of the phenomenon.
- Common methods: interviews, focus groups, observations.

"I'm neither a good cop nor a bad cop, Jerome. Like yourself, I'm a complex amalgam of positive and negative personality traits that emerge or not, depending on circumstances."

The Types of Beliefs Held by Researchers ("Philosophical Worldviews") Influence the Practice of Research and Choice of Methods

Positivist (or "Post-positivist")

- Knowledge is objective
- "Causes determine effects/outcomes"
- Reductionist: study complex things by breaking down to simpler ones
- Prefer quantitative approaches
- **Verifying (or Falsifying) theories**

Constructivist / Interpretivist

- Knowledge is socially constructed
- Truth is relative to context
- Theoretical terms are open to interpretation
- Prefer qualitative approaches
- **Generating "local" theories**

Objectivist Deductive

Subjectivist Inductive

Four Philosophical Worldviews Are Commonly Encountered

Positivist (or "Post-positivist")
- Knowledge is objective
- "Causes determine effects/outcomes"
- Reductionist: study complex things by breaking down to simpler ones
- Prefer quantitative approaches
- **Verifying (or Falsifying) theories**

Constructivist / Interpretivist
- Knowledge is socially constructed
- Truth is relative to context
- Theoretical terms are open to interpretation
- Prefer qualitative approaches
- **Generating "local" theories**

Advocate / Critical Theorist
- Research is a political act
- Knowledge is created to empower groups/individuals
- Choose what to research based on who it will help
- Prefer participatory approaches
- **Seeking change in society**

Pragmatist
- Research is problem-centered
- "All forms of inquiry are biased"
- Truth is what works at the time
- Prefer multiple methods / multiple perspectives
- **Seeking practical solutions to problems**

Q: Which Do You Subscribe to?

Positivist (or "Post-positivist")
- Knowledge is objective
- "Causes determine effects/outcomes"
- Reductionist: study complex things by breaking down to simpler ones
- Prefer quantitative approaches
- **Verifying (or Falsifying) theories**

Constructivist / Interpretivist
- Knowledge is socially constructed
- Truth is relative to context
- Theoretical terms are open to interpretation
- Prefer qualitative approaches
- **Generating "local" theories**

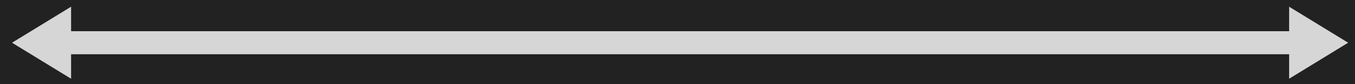
Advocate / Critical Theorist
- Research is a political act
- Knowledge is created to empower groups/individuals
- Choose what to research based on who it will help
- Prefer participatory approaches
- **Seeking change in society**

Pragmatist
- Research is problem-centered
- "All forms of inquiry are biased"
- Truth is what works at the time
- Prefer multiple methods / multiple perspectives
- **Seeking practical solutions to problems**

Selection of Methods Depends on:

- Approach to Research
- Nature of Contribution
- Specific Research Question
- State of Knowledge
- . . .

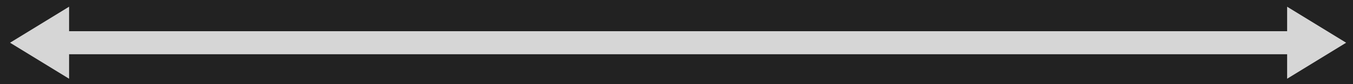
Another Few Ways of Thinking About How Methods Are Related

Qualitative  Quantitative

Very little prior knowledge:
"What's going on?"

Extensive prior knowledge:
"Precisely how is this different?"

Another Few Ways of Thinking About How Methods Are Related

Qualitative  Quantitative

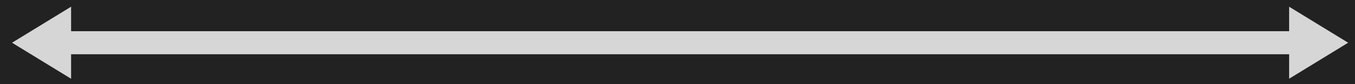
Very little prior knowledge:
"What's going on?"

Extensive prior knowledge:
"Precisely how is this different?"

Rich

Precise

Another Few Ways of Thinking About How Methods Are Related



Very little prior knowledge:
"What's going on?"

Extensive prior knowledge:
"Precisely how is this different?"

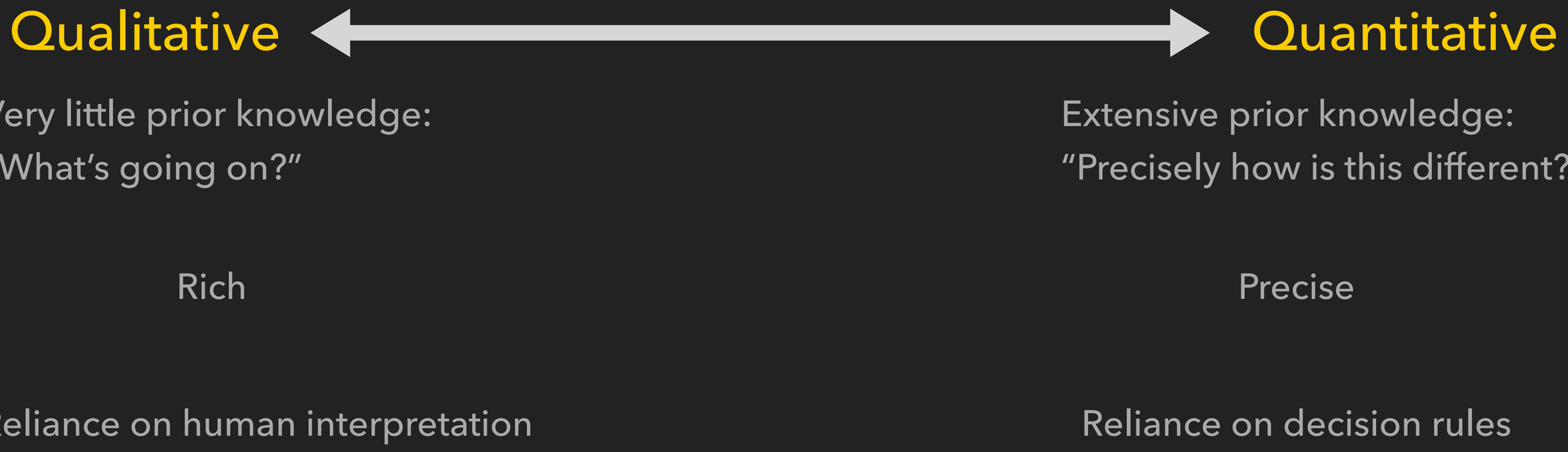
Rich

Precise

Reliance on human interpretation

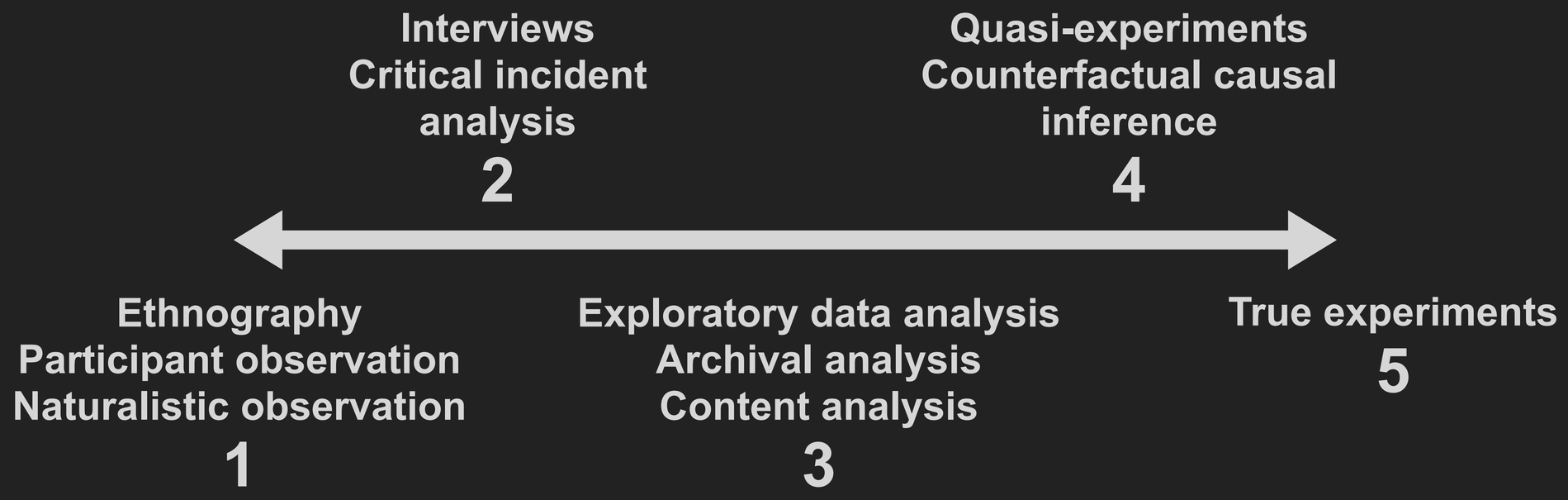
Reliance on decision rules

Another Few Ways of Thinking About How Methods Are Related



- Often most effective to use methods in combination or in sequence (rich to precise)

Points on the Spectrum



RESEARCH ARTICLE Open Access

Communication and relationship dynamics in surgical teams in the operating room: an ethnographic study

Birgitte Tørring[1,2*], Jody Hoffer Gittell[3], Mogens Laursen[1,4], Bodil Steen Rasmussen[1,5] and Erik Elgaard Sørensen[1,6]

Abstract

Background: In surgical teams, health professionals are highly interdependent and work under time pressure. It is of particular importance that teamwork is well-functioning in order to achieve quality treatment and patient safety. Relational coordination, defined as "communicating and relating for the purpose of task integration," has been found to contribute to quality treatment and patient safety. Relational coordination has also been found to contribute to psychological safety and the ability to learn from mistakes. Although extensive research has been carried out regarding relational coordination in many contexts including surgery, no study has explored how relational coordination works at the micro level. The purpose of this study was to explore communication and relationship dynamics in interdisciplinary surgical teams at the micro level in contexts of variable complexity using the theory of relational coordination.

Methods: An ethnographic study was conducted involving participant observations of 39 surgical teams and 15 semi-structured interviews during a 10-month period in 2014 in 2 orthopedic operating units in a university hospital in Denmark. A deductively directed content analysis was carried out based on the theory of relational coordination.

Results: Four different types of collaboration in interdisciplinary surgical teams in contexts of variable complexity were identified representing different communication and relationship patterns: 1) *proactive and intuitive communication*, 2) *silent and ordinary communication*, 3) *inattentive and ambiguous communication*, 4) *contradictory and highly dynamic communication*. The findings suggest a connection between communication and relationship dynamics in surgical teams and the level of complexity of the surgical procedures performed.

Conclusion: The findings complement previous research on interdisciplinary teamwork in surgical teams and contribute to the theory of relational coordination. The findings offer a new typology of teams that goes beyond weak or strong relational coordination to capture four distinct patterns of relational coordination. In particular, the study highlights the central role of *mutual respect and presents proposals for improving relational coordination in surgical*

An Ethnographic Study of Copy and Paste Programming Practices in OOPL

Miryung Kim[1] Lawrence Bergman[2] Tessa Lau[2] David Notkin[1]

Department of Computer Science & Engineering
University of Washington[1]
{miryung, notkin}@cs.washington.edu

IBM T. J. Watson Research Center[2]
{bergmanl, tessalau}@us.ibm.com

Abstract

Although programmers frequently copy and paste code when they develop software, implications of common copy and paste (C&P) usage patterns have not been studied previously. We have conducted an ethnographic study in order to understand programmers' C&P programming practices and discover opportunities to assist common C&P usage patterns. We observed programmers using an instrumented Eclipse IDE and then analyzed why and how they use C&P operations. Based on our analysis, we constructed a taxonomy of C&P usage patterns.

This paper presents our taxonomy of C&P usage patterns and discusses our insights with examples drawn from our observations. From our insights, we propose a set of tools that both can reduce software maintenance problems incurred by C&P and can better support the intents of commonly used C&P scenarios.

locate code duplicates and refactor existing duplications to a unit of programming language abstraction [1][2][3][6][12][14][15]. However, in practice, a substantial amount of duplicated code is still present in many software systems [6][12]. Our understanding of how and why code clones are created is very limited.

Earlier studies have formed a few informal hypotheses about how C&P is performed by programmers to reuse code [17][18]. However, existing work has not focused specifically on solving the possible problems that can be incurred by C&P during software evolution.

The main purpose of our work is to investigate common C&P usage patterns and associated implications as a first step toward understanding and solving such problems. We believe that understanding when and how C&P is used will also reveal limitations in programming language designs and the lack of software engineering tool support to cope with common usage patterns.

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: 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

Social Coding in GitHub: Transparency and Collaboration in an Open Software Repository

Laura Dabbish, Colleen Stuart, Jason Tsay, Jim Herbsleb
School of Computer Science and Center for the Future of Work, Heinz College
Carnegie Mellon University
5000 Forbes Ave., Pittsburgh, PA 15213
{dabbish, hcstuart, jtsay, jdh}@cs.cmu.edu

ABSTRACT
Social applications on the web let users track and follow the activities of a large number of others regardless of location or affiliation. There is a potential for this transparency to radically improve collaboration and learning in complex knowledge-based activities. Based on a series of in-depth interviews with central and peripheral GitHub users, we examined the value of transparency for large-scale distributed collaborations and communities of practice. We find that people make a surprisingly rich set of social inferences from the networked activity information in GitHub, such as inferring someone else's technical goals and vision when they edit code, or guessing which of several similar projects has the best chance of thriving in the long term. Users combine these inferences into effective strategies for coordinating work, advancing technical skills and managing their reputation.

functionality. Users can articulate an interest network of people or things by defining a set of individuals or artifacts (like blogs or RSS feeds) to pay attention to. In doing so, users immediately subscribe to a stream of events and actions other individuals take. Thus the social web provides an unprecedented level of transparency in the form of visibility of others' actions on public or shared artifacts. The question remains, however, what benefits this transparency provides, particularly in the large scale (i.e. across a community).

Previous work on awareness has explored the value of activity information for small groups [7, 12]. This work has found that notifying members of actions on shared artifacts helps them maintain mental models of others activities [11] and avoid potential coordination conflicts [20]. However, activity awareness has largely been examined in the context of well-defined small groups within organizations. Online,

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, checking interpretations

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

How Work From Home Affects Collaboration: A Large-Scale Study of Information Workers in a Natural Experiment During COVID-19

Longqi Yang, Sonia Jaffe, David Holtz, Siddharth Suri, Shilpi Sinha, Jeffrey Weston, Connor Joyce, Neha Shah, Kevin Sherman, CJ Lee, Brent Hecht, Jaime Teevan

Microsoft Corporation
Correspondence: Longqi.Yang@microsoft.com

ABSTRACT

The COVID-19 pandemic has had a wide-ranging impact on information workers such as higher stress levels, increased workloads, new workstreams, and more caregiving responsibilities during lockdown. COVID-19 also caused the overwhelming majority of information workers to rapidly shift to working from home (WFH). The central question this work addresses is: can we isolate the effects of WFH on information workers' collaboration activities from all other factors, especially the other effects of COVID-19? This is important because in the future, WFH will likely to be more common than it was prior to the pandemic.

We use difference-in-differences (DiD), a causal identification strategy commonly used in the social sciences, to control for unobserved confounding factors and estimate the causal effect of WFH. Our analysis relies on measuring the difference in changes between those who WFH prior to COVID-19 and those who did not. Our preliminary results suggest that on average, people spent more time on collaboration in April (Post WFH mandate) than in February (Pre WFH mandate), but this is primarily due to factors other than WFH, such as lockdowns during the pandemic. The change attributable to WFH specifically is in the opposite direction. This reversal shows the importance of using causal inference: a simple analysis would have resulted in the wrong conclusion. We further find that the effect of WFH is moderated by individual remote collaboration experience prior to WFH. Meanwhile, the medium for collaboration has also shifted due to WFH: instant messages were used more, whereas scheduled meetings were used less. We discuss design implications -- how future WFH may affect focused work, collaborative work, and creative work.

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: 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

Does Tweeting Improve Citations? One-Year Results From the TSSMN Prospective Randomized Trial

Jessica G. Y. Luc, MD, Michael A. Archer, MD, Rakesh C. Arora, MD, PhD, Edward M. Bender, MD, Arie Blitz, MD, MBA, David T. Cooke, MD, Tamara Ni Hlci, MB, MCh, Biniam Kidane, MD, MSc, Maral Ouzounian, MD, PhD, Thomas K. Varghese, Jr, MD, MS, and Mara B. Antonoff, MD

Division of Cardiovascular Surgery, Department of Surgery, University of British Columbia, Vancouver, British Columbia, Canada; Division of Thoracic Surgery, Department of Surgery, SUNY-Upstate Medical University, Syracuse, New York; Section of Cardiac Surgery, Department of Surgery, Max Rady College of Medicine, University of Manitoba, St Boniface Hospital, Winnipeg, Manitoba, Canada; Department of Cardiothoracic Surgery, Stanford University, Palo Alto, California; Division of Cardiac Surgery, University of Cincinnati Medical Center, Cincinnati, Ohio; Section of General Thoracic Surgery, UC Davis Health, Sacramento, California; Department of Cardiothoracic Surgery, Morriston Hospital, ABMU, Swansea, United Kingdom; Section of Thoracic Surgery, Department of Surgery, University of Manitoba, Winnipeg, Manitoba, Canada; Division of Cardiovascular Surgery, Peter Munk Cardiac Centre, Toronto General Hospital, University of Toronto, Toronto, Ontario, Canada; Division of Cardiothoracic Surgery, Department of Surgery, University of Utah, Salt Lake City, Utah; Department of Thoracic and Cardiovascular Surgery, University of Texas MD Anderson Cancer Center, Houston, Texas

EDUCATION

Background. The Thoracic Surgery Social Media Network (TSSMN) is a collaborative effort of leading journals in cardiothoracic surgery to highlight publications via social media. This study aims to evaluate the 1-year results of a prospective randomized social media trial to determine the effect of tweeting on subsequent citations and nontraditional bibliometrics.

Methods. A total of 112 representative original articles were randomized 1:1 to be tweeted via TSSMN or a control (non-tweeted) group. Measured endpoints included citations at 1 year compared with baseline, as well as article-level metrics (Altmetric score) and Twitter analytics. Independent predictors of citations were identified through univariable and multivariable regression analyses.

Results. When compared with control articles, tweeted articles achieved significantly greater increase in Altmetric scores (Tweeted 9.4 ± 5.8 vs Non-tweeted 1.0 ± 1.8, $P < .001$), Altmetric score percentiles relative to articles of similar age from each respective journal (Tweeted 76.0 ± 9.1 percentile vs Non-tweeted 13.8 ± 22.7 percentile, $P < .001$), with greater change in citations at 1 year (Tweeted +3.1 ± 2.4 vs Non-Tweeted +0.7 ± 1.3, $P < .001$). Multivariable analysis showed that independent predictors of citations were randomization to tweeting (odds ratio [OR] 9.50; 95% confidence interval [CI] 3.30-27.35, $P < .001$), Altmetric score (OR 1.32; 95% CI 1.15-1.50, $P < .001$), open-access status (OR 1.56; 95% CI 1.21-1.78, $P < .001$), and exposure to a larger number of Twitter followers as quantified by impressions (OR 1.30, 95% CI 1.10-1.49, $P < .001$).

Conclusions. One-year follow-up of this TSSMN prospective randomized trial importantly demonstrates that tweeting results in significantly more article citations over time, highlighting the durable scholarly impact of social media activity.

(Ann Thorac Surg 2021;111:296-301)

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: 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 people can manipulate
- What resources does it take to do it well?
  - Access to appropriate subjects
  - Statistics, measurement instruments

All Methods Are Limited

This course teaches strategies to overcome weaknesses

Course Overview

Coordinates

- Instructor: Bogdan Vasilescu, S3D

- Class meets twice a week, Tuesdays / Thursdays 3:30-4:50pm,
  - WEH 4708
- Communication: Canvas & email & Slack
  - In-person office hours: by appointment
- Materials:
  - Most public on course webpage: https://bvasiles.github.io/empirical-methods
  - Assignments / private on Canvas

Guiding Principle: Prioritize Supporting Each Other as Humans

Format

- Split lecture - seminar style
  - Lecture/discussion of a new method, summarizing the material (Bogdan)
  - In-class hands-on practice with the method (everyone)
  - Presentations of the example papers (students) – sign up sheet

- Most work happens on your own, outside class
  - Lots of curated readings, best done before each class
    - Method descriptions, how-tos, standard of rigor (the “theory”)
    - Example research papers applying that method (the “practice”)

Two Types of Homework Assignments

- (1) A mixed-methods research project using empirical component(s) from the course
  - Deliverables:
    - Short kick-off presentation (~1 month in)
      - Research question(s), overview of study design, rough plan for data collection and analysis
    - Final report (end of semester)
      - Thorough literature review, detailed description of methods, results, discussion

Notes:

- Must combine qualitative and quantitative methods
  - But it's okay if not equally complete
- Grading emphasis on motivation, study design, soundness of methodology, quality of writing
- Individual / pairs both ok
- Best if aligned with your current research
  - Alternatively, talk to me, I have ideas!

Two Types of Homework Assignments

- (2) Occasional small assignments throughout the semester, to get hands-on experience with some of the methods.
  - For example:
    - Do and write up a literature review
    - Design an interview protocol or a survey instrument
    - Conduct two interviews and analyze the data using an LLM / ChatGPT
    - Apply a particular statistical analysis technique, e.g., segmented regression, to an existing dataset
    - Study how FCEs vary with demographics and tenure at CMU (data from https://cmucourses.com)
    - Critique a research paper

Grading

- 50% research project
  - 10% kick-off presentation
  - 10% final presentation
  - 30% final report
- 40% other homework assignments
- 10% in-class presentations & participation
  - Occasional Canvas quizzes

Tentative Schedule

Tue, Jan 16 - Introduction
Thu, Jan 18 - Formulating research questions
Tue, Jan 23 - The role of theory
Thu, Jan 25 - Literature review
Tue, Jan 30 - Conducting interviews
Thu, Feb 1 - TBD
Tue, Feb 6 - Exemplar interview papers
Thu, Feb 8 - Qualitative data analysis
Tue, Feb 13 - Survey design
Thu, Feb 15 - In-class activity: qualitative coding & thematic analysis
Tue, Feb 20 - Project proposal presentations
Thu, Feb 22 - Experimental design (part I)
Tue, Feb 27 - Experimental design (part II)
Thu, Feb 29 - Experimental design (part III)
~~Tue, Mar 5~~ - ~~Spring break, no class~~
~~Thu, Mar 7~~ - ~~Spring break, no class~~
Tue, Mar 12 - Mixed-methods designs
Thu, Mar 14 - Intro to regression modeling
Tue, Mar 19 - Diagnostics, factors, std coefficients
Thu, Mar 21 - Simpson's paradox, exemplar papers, in-class activity
Tue, Mar 26 - Interrupted time series design
Thu, Mar 28 - In-class activity: interrupted time series analysis
Tue, Apr 2 - Difference in differences
Thu, Apr 4 - CausalImpact, instrumental variables
Tue, Apr 9 - Agree to disagree
~~Thu, Apr 11~~ - ~~Carnival, no class~~
Tue, Apr 16 - Social network analysis
Thu, Apr 18 - Social network analysis
Tue, Apr 23 - Final presentations (part I)
Thu, Apr 25 - Final presentations (part II)

What This Class Is and Isn't About

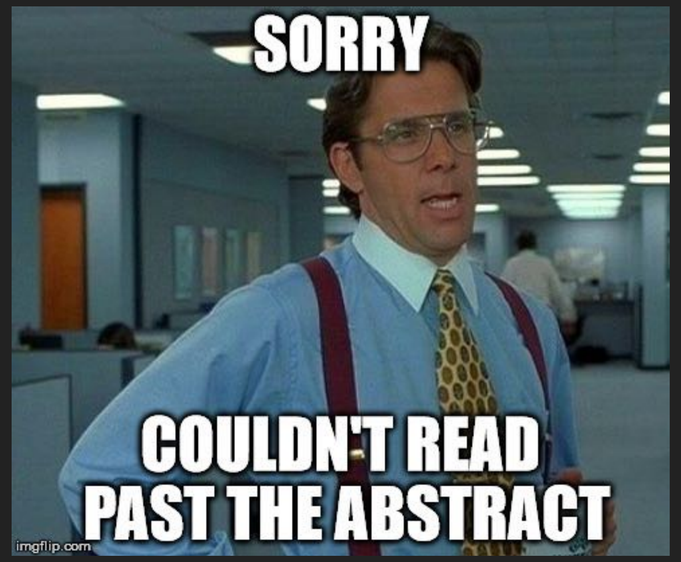
- This class is not about software engineering.
  - In fact, this is more a science class than a class about any CS discipline.
  - But because I am most familiar with SE, you will see many readings from SE venues.
    - No prior knowledge of SE or SE research experience is needed to digest these.

What This Class Is and Isn't About

- This class is ~~secretly~~ about communication.
  - You will practice:
    - Articulating what the problem is
    - Articulating why the problem is important
    - Articulating your vision and plan for solving the problem
    - Disseminating (your) results
      - Oral presentations throughout the semester
      - Blog posts?
      - A research project report at the end
- Effective (oral & written) communication is:
  - Necessary for successful research
  - Much harder than you think

What This Class Is and Isn't About

- This class is ~~secretly~~ about peer review.
  - You will read and critique (including in writing) many research papers throughout the semester



What This Class Is and Isn't About

- This class is ~~secretly~~ about developing a healthy dose of skepticism.
  - All fields of science use (the same) empirical methods
  - By learning about empirical research in CS you'll also get better at recognizing strengths and weaknesses in any scientific paper
    - Tune your scientific BS meter!

Anecdotal evidence reliable? One man says "yes".

A STUDY CONDUCTED YESTERDAY by a man on himself concluded that self-reported anecdotal evidence is, in fact, both reliable and relevant.

The landmark study, conducted by Mark Mattingly of Virginia Beach in his apartment, concluded with 100% accuracy that data collected from personal experience can disprove other data conducted by reputable scientific institutions, thereby proving once and for all that "statistics can't be trusted".

In a press release Mr. Mattingly took aim at his detractors saying that "…this study shows what I've been telling people on the internet for years: all your fancy evidence and statistics don't mean nothing in the real world.".

A frequenter of internet forums, comment sections, and social media, Mr. Mattingly recounts that he was inspired to undertake the study when someone reportedly kept insisting that he provide evidence for his claims. "I think everyone's entitled to an opinion, and that my opinion is worth just as much as anyone else's" Mr. Mattingly said.

Academic types have criticised the study, and papers who are publishing it, saying that it lacks everything and makes no sense. When shown the study, Emeritus Professor James Albrecht of Carnegie Mellon University looked all confused and hopeless before making pining, guttural sounds.

Mr. Mattingly in his apartment looking all smug.

Mr. Mattingly has responded saying that this is just the first of many studies he intends to conduct, and that a meta-analysis of people who have opinions and anecdotal experiences independent of controls, methodological rigor, blinding and peer review are soon to be published, adding further weight to his initial findings.

***Published Saturday 22 February 2014 by** yourlogicalfallacyis.com/anecdotal* *Photo: Weasello*

Skeptics Are Welcome!

Iain Cruickshank,
SC PhD, West Point

"I tend to prefer formal methods and **really did not like the survey and interview portions of the class** [as] I would **never need these methods in the real-world**.

So, I just started a new job this fall, and [...] I got assigned to [do X]. As I pondered this, I realized, albeit begrudgingly, that this would be a great opportunity to do a survey. So, **having just arrived in the real-world from doing a Ph.D., one of my first studies has been to design and implement a survey**...

I thought you might appreciate hearing that story, both because I am no-kidding using the skills you taught me (whether I liked it or not) and because the irony is pretty rich."

Who are you?
What is your research?
What would make this course valuable to you?

Summary:

All Methods Are Flawed

Selection of Methods Depends on:

- Philosophical Worldview
- Approach to Research
- Nature of Contribution
- Specific Research Question
- State of Knowledge
- . . .

Credits

- Graphics:
  - New Yorker magazine cartoons
  - “This if fine” meme by K.C. Green
  - Steve Easterbrook slides
- Content:
  - Chapter 1 - Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.
  - Varpio, L., Paradis, E., Uijtdehaage, S., & Young, M. (2020). The distinctions between theory, theoretical framework, and conceptual framework. Academic Medicine, 95(7), 989-994.
  - Young, M., Varpio, L., Uijtdehaage, S., & Paradis, E. (2020). The spectrum of inductive and deductive research approaches using quantitative and qualitative data. Academic Medicine, 95(7), 1122.

Bonus Slides

Q: What's the Relationship?
Does Inductive → Qualitative? Deductive → Quantitative?

A: No. These Perspectives Are Frequently (and Falsely) Conflated

Quantitative (numerical) data	Qualitative (non-numerical) data
Exploratory factor analysis[1]: A set of statistical techniques, typically applied to assessment- or survey-generated data, that identifies the underlying theoretical constructs (i.e., "factors") of the phenomena of interest, which researchers then name inductively.	**Traditional grounded theory[2]:** A research methodology aimed at generating a local theory using qualitative data collected with participants. The researcher acts as a tabula rasa (blank slate), making sense of a phenomenon without biasing that interpretation.
Structural equation modeling[3]: A set of approaches to analysis whereby several statistical models are built based on theory and then tested with an appropriate dataset. The model that best fits the data is considered superior.	**Ethnography[4]:** This approach often combines existing exploratory theories about human social behavior and collects data to observe and describe the culture of a group.
Bayesian approaches to analysis[5]: These analytic approaches use previous knowledge, available data, or beliefs, to acknowledge the a priori likelihood of findings. The strength of prior evidence influences analysis.	**Constructivist grounded theory[6]:** This approach analyzes data with the goal of contributing to previous understandings of a phenomenon, building explicitly on others' work.
Analysis of hypothesis-driven experimental studies[1]: Among the most familiar, these approaches result in a test of significance, which leads to the rejection of, or failure to reject, the null hypothesis (that the findings are due to chance).	**Deductivist content analysis[7]:** Qualitative data are analyzed using a predetermined theory or theoretical framework, to find examples of constructs and support or challenge them.

Young, M., Varpio, L., Uijtdehaage, S., & Paradis, E. (2020). The spectrum of inductive and deductive research approaches using quantitative and qualitative data. Academic Medicine, 95(7), 1122.