COMP 150-4: Human Factors in Security and Privacy

Qualitative Methods

Lecture 4

Prof. Daniel Votipka
Spring 2021

(some slides courtesy of Prof. Adam Aviv and Prof. Michelle L. Mazurek)

Administrivia

- HW0 is graded!
  - Lots of expected anxiety over COVID and course pace
  - Projects may be done individually
- Do not pay for any research
- No class next Tuesday
- Pitch Day will be 2/23.
  - Email me any ideas you'd like to pitch

What we did last time!

- Research questions and hypotheses
- Is my study valid?

Research Methods (a sample)

- Qualitative
  - Collection methods:
    - Interviews/Focus Groups
    - Diary Studies
    - Surveys (open text)
  - Analysis methods:
    - Thematic Coding

- Quantitative
  - Collection methods:
    - Surveys (closed text)
      - Likert Responses
    - Behavioral observation
    - Artifact analysis
    - Experimental design
  - Analysis methods:
    - Significance Testing
    - Regression

Research Methods (a sample)

- Qualitative
  - Collection methods:
    - Interviews/Focus Groups
    - Diary Studies
    - Surveys (open text)
  - Analysis methods:
    - Thematic Coding

- Quantitative
  - Collection methods:
    - Surveys (closed text)
      - Likert Responses
    - Behavioral observation
    - Artifact analysis
    - Experimental design
  - Analysis methods:
    - Significance Testing
    - Regression

Qualitative Data Collection and Analysis

- What is qualitative data?
  - Qualitative data is raw, potentially unstructured, textual data, e.g., interview transcripts, answers to open-ended questions

- Key questions:
  - What are the methods we can use to collect qualitative data?
    - *Interviews*
    - *Open Ended Survey Questions*
    - *Diary Studies*
    - *Focus Groups*
  - *How do we covert this raw, mostly unstructured data, into something more structured that can be analyzed?*
    - Inductive Themes: What are the common themes that occur in the data?
    - Quantitatively: How frequently do those themes occur?

Interviews/Focus Groups

Interviews and Focus Groups

- Direct discussion with concerned participants
  - A way to “go deep” on a topic
- Interviews: Single participant
  - Interviewer ask questions and follow up of a single participant
- Focus Groups: A group of participants
  - Interviewer ask questions of a group of participants
  - Participants in the group may also interact with each other on questions

Pros and Cons of Interviews

How might this compare to a survey?

- PRO: Extremely Flexible and Interactive
  - Even with structured questions, you can ask detailed follow ups
  - Force participants to fully explain their reasons
  - Get to the WHY not just the WHAT --- extremely RICH DATA SOURCE!
- PRO: (semi-)Structured Questions Can Change Iteratively
  - Preliminary analysis could lead to quick additions of questions to address emergent themes
  - (con) sometimes your script gets longer!
- Pro: Exploratory
  - Can identify themes or hypothesis that can be deductively evaluated later
- CON: High Effort Requirements
  - One interviewer to one participant --- That's a lot of time!
  - A one-hour long interview with 50 participants is months of work
- CON: Analysis
  - Transcription of interviews into raw text --- Expensive!
  - Converting raw text into themes/codes --- Time consuming!

Interview Best Practices

- Make the participant feel comfortable
- Avoid leading questions
  - Ordering is important!
- Support whatever the participant says
  - Don’t make them feel incorrect or stupid
- Know when to ask follow up questions
  - Have some example follow up questions planned out ahead of time, or keep track of them based on prior interviews
- Get a broad range of participants (hard!)

Interview Strategies

- Fully Structured
  - Rigid script with questions in well defined order
  - Survey administered by the researcher
  - Easier to analyze since every subject is asked the same question

- Unstructured Interview
  - Interviewer guide with a list of topics
    - Ask an opening question and go from there
    - Once there is a lull in the conversation, move to the next topic

- Semi-structured Interview
  - Rigid set of questions, but can ask follow ups
  - Ex: Can you please explain your answer more? What does ‘X’ mean?

Exercise: Write some interview questions

- Research Question:
  - **How do users gauge the trustworthiness (with respect to security and privacy) of browser extensions before installation?**

- Exercise: Come up with three-to-four interview questions in your group?
  - Semi-Structured – rigid questions, but allow for follow ups
    - Explore one or two follow up questions you might need to ask as well
  - **8 MINUTES!**

- Some ways to break this down:
  - Consider the RQ as a process: how do you/participant explore that process?
  - What is the “why” and “how” someone may engage with that process?
  - Are there other contextual information you might want to learn?

Exercise Report Out: Interview Questions

- Do you even use browser extensions?
  - Which ones did you use?
  - Why did you download them?
- What is a browser extension?

- Do you know what browser extensions are?
  - How did you learn about them?
  - Where did you get them?
- Have you removed any extensions?
  - If so, why?
- Have you read any reviews?
  - Have you decide not to instal something? Why?
- Do you read the privacy statements?

Diary Studies

What is a diary study?

- Rich longitudinal data (from a few participants)
  - In the field … ish
- Natural reactions and occurrences
  - Existence and quantity of phenomena
  - User reactions in the moment rather than via recall
- Lots of work for you and your participants
  - Lots of data to look at
  - Lots of time for the participant
- On paper vs. technology-mediated
  - Eg., have an app on the phone that prompts for input

Experience Sampling

- Kind of a prompted diary
- Send participants a stimulus when they are in their natural life, not in the lab
  - Ex: Every time a participant navigates to an http (rather than https) website, ask a question about how they trust or understand that website?
  - Ex: Have participants send an encrypted email and document their experience each time.

Group Exercise: Developing a Diary/ESM

- Can we come up with some security and privacy examples for a diary/ EMS?
  - <fill in>

Diary / ESM best practices

- When will an entry be recorded?
  - How often? Over what time period?
- How long will it take to record an entry?
  - How structured is the response?
- Pay well
  - Pay per response? But don’t create bias

Qualitative analysis

How to analyze in interview/qualitative data

- You may have hours (and hours) of recordings or pages and pages of notes/transcriptions --- how do you convert this into structured data?

- Text Analysis
  - Frequency of terms
  - Patterns of usage

- Categories
  - What primary themes are discussed?
  - *Validity?*

Qualitative Coding Schemes

- Coding: assigning categories and descriptors to blocks text

Why did you decide to adopt [this communication tool]?

“there is no point of using a chat service that not many people use”	**Adoption**
“The only call people if it’s urgent and I need a quick response. If it’s a longer conversation I might want a record of, I use email.”	**Context**
“I use WhatsApp since I can set up group chats”	**Features**

Obstacles to the Adoption of Secure Communication Tools. Abu-Salma et al. 2017

Qualitative Coding Schemes

- Coding: assigning categories and descriptors to blocks text

- A-Priori Coding
  - Use an established theory or hypothesis to guide the creation of codes
    - Ex: Explicitly looking for theme X and theme Y to compare how frequently they occur
  - Established theory could come from other-/prior-work on this or related topics

- Emergent Coding
  - Coding without a prior theory or guidance
  - Simply start by noting intercepting concepts or ideas
  - Continually refine them over time until a coherent model emerges

Grounded Theory

- Develop a set of well-grounded theories from the data

- Four stages:
  - Open Coding
  - Development of concepts
    - Axial Coding
  - Grouping concepts into categories
  - Forming a theory

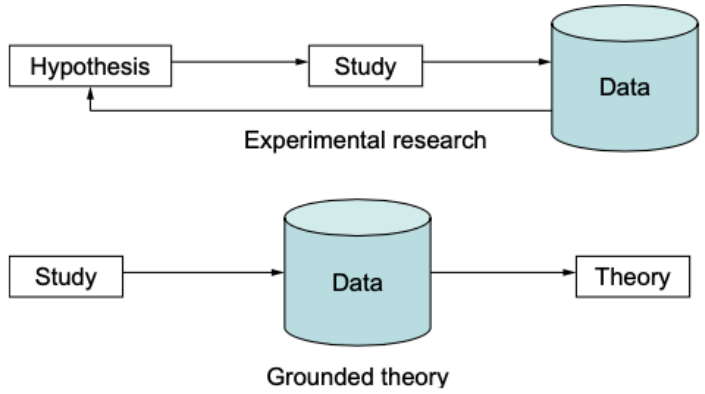


FIGURE 11.1

Experimental research compared with grounded theory.

- These stages can be repeated, paired with tweaking the survey instrument and collecting new data, until a theory is established and supported

- *When do you stop collecting data?* When saturation is reached.
  - When no knew codes/themes are emerging that inform your theory

Tools

- Spreadsheets!
- MaxQDA, Atlas.ti, Nvivo, Cassandra, QDAMiner, Dedoose, others
- Most of the good ones aren’t free
  - You’ll get a free copy of MaxQDA for this course!!
- Features to consider:
  - Granularity, multiple coders, auto stats, etc.

Demo: Reverse Engineering Processes

- Four stages:
  - Open Coding
  - Development of concepts
    - Axial Coding
  - Grouping concepts into categories
  - Forming a theory

An Observational Investigation of Reverse Engineers' Processes

Daniel Votipka, Seth M. Rabin, Kristopher Micinski*,
Jeffrey S. Foster†, and Michelle M. Mazurek
*University of Maryland; *Syracuse University; †Tufts University*
{dvotipka,srabin,mmazurek}@cs.umd.edu; kkmicins@syr.edu; jfoster@cs.tufts.edu

Abstract

Reverse engineering is a complex process essential to software-security tasks such as vulnerability discovery and malware analysis. Significant research and engineering effort has gone into developing tools to support reverse engineers. However, little work has been done to understand the way reverse engineers think when analyzing programs, leaving tool developers to make interface design decisions based only on intuition.

This paper takes a first step toward a better understanding of reverse engineers' processes, with the goal of producing insights for improving interaction design for reverse engineering tools. We present the results of a semi-structured, observational interview study of reverse engineers (N=16). Each observation investigated the questions reverse engineers ask as they probe a program, how they answer these questions,

Researchers, companies, and practitioners have developed an extensive array of tools to support RE [5–24]. However, there is limited theoretical understanding of the RE process itself. While existing tools are quite useful, design decisions are currently ad-hoc and based on each designer's personal experience. With a more rigorous and structured theory of REs' processes, habits, and mental models, we believe existing tools could be refined, and even better tools could be developed. This follows from recommended design principles for tools supporting complex, exploratory tasks, in which the designer should "pursue the goal of having the computer vanish" [25, pg. 19-22].

In contrast to RE, there is significant theoretical understanding of more traditional program comprehension—how developers read and understand program functionality—including tasks such as program maintenance and debugging [26–36].

Demo: MaxQDA and Interview Coding

Establishing validity

- Typical process:
  - One or two researchers develop codebook
  - Independently code samples
  - Measure agreement
  - Resolve disagreements to 100%

Reliability of Coding Process

- How do we know if the coding (or codebook) is reliably capturing the themes?
  - If we asked another coder to code the same data with the same codebook, would we get the same results?

- Inter-Rater Reliability
  - Amount of agreement between the coders
  - For ever case a code could be applied, do the coders agree to apply the same set of codes to the same segment of text?
    - Not applying a code also counts towards agreement

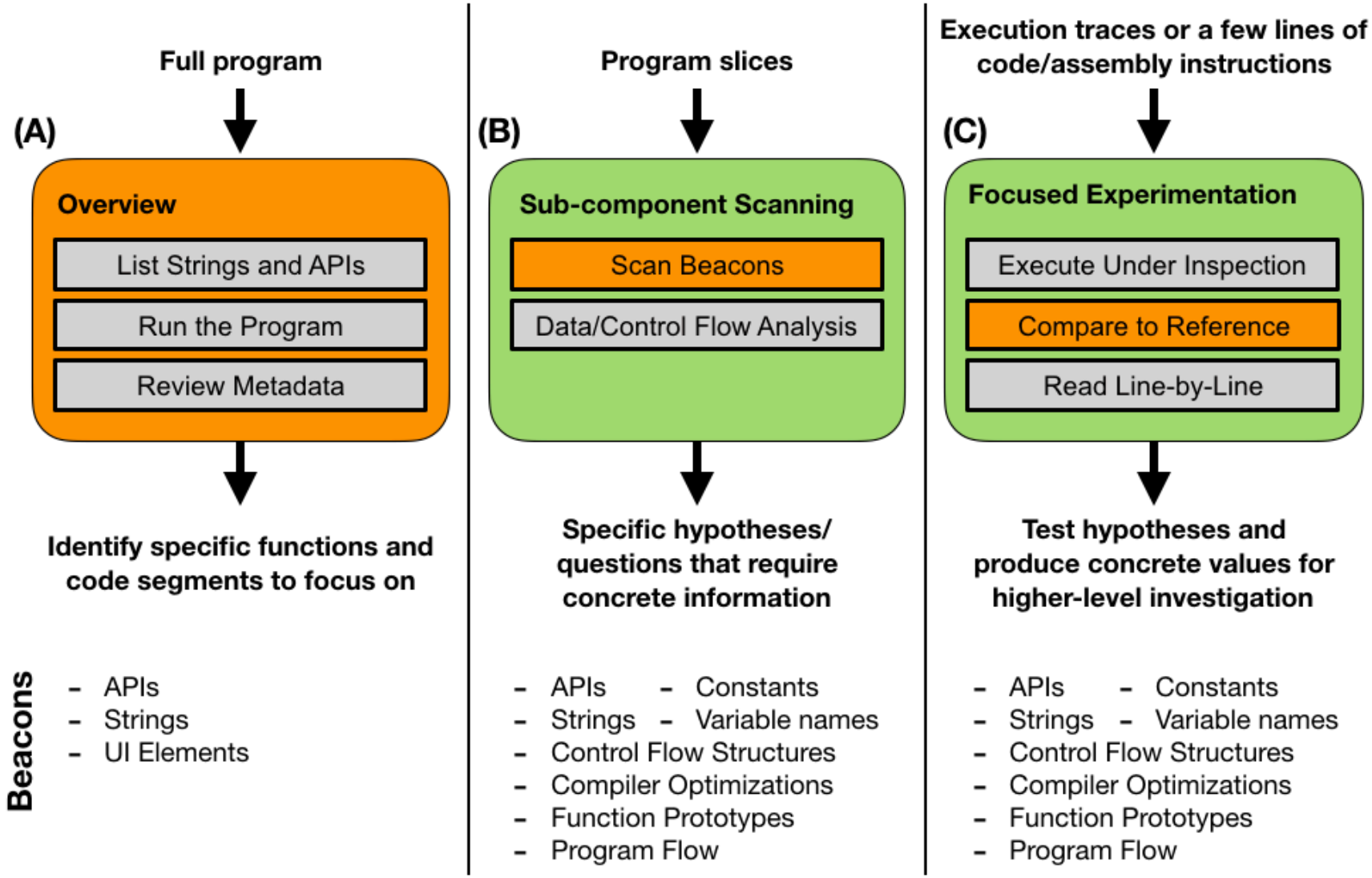
Measures of agreement

- Percent agreement
  - In most cases, overstates agreement due to chance
- Cohen's Kappa
  - Accounts for chance
  - Statistical independence of coder data
- Krippendorff's Alpha
  - Fundamentally measures disagreement
  - Works for multiple variables, non-exclusive, nominal/ordinal/interval, missing data, etc.

Reporting on qual analysis

- Describe coding process
  - How many researchers, rounds, etc.
- Provide initial agreement numbers
- Describe how conflicts were resolved

Demo: Axial Coding + Develop a Theory



Notes on sample size

- Goal to surface ideas, develop theory
  - Need further study to confirm, build on, measure frequency

What we did today!

- Qualitative methods
  - Goals
  - Interviews/Focus groups
  - Diary Studies
  - Qualitative Coding

What's next?

- Crash course in encryption
- PGP usability
  - "Why Johnny Can't Encrypt"
- Key Management
  - "This World of Ours"
  - "A Comparative Usability Study of Key Management in Secure Email"

Logistics + Questions?

- No class next Tuesday