ECE1724H S2: Empirical Software Engineering

Interviews

Shurui Zhou

The Edward S. Rogers Sr. Department
of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

Planning Checklist

- [x] Pick a topic
- [x] Identify the research question(s)
- [x] Check the literature
- [x] Identify your philosophical stance
- [ ] Identify appropriate theories
- [ ] Choose the method(s)
- [ ] Design the study
  - Unit of analysis?
  - Target population?
  - Sampling technique?
  - Data collection techniques?
  - Metrics for key variables?
  - Handle confounding factors
- [ ] Critically appraise the design for threats to validity
- [ ] Get IRB approval
  - Informed consent?
  - Benefits outweigh risks?
- [ ] Recruit subjects / field sites
- [ ] Conduct the study
- [ ] Analyze the data
- [ ] Write up the results and publish them
- [ ] Iterate

Identify Appropriate Theories

Where do theories come from?

The Role of Theory

- A scientific theory identifies and defines a set of phenomena; makes assertions about their nature and the **causal** relationships between them.
- **Positivism**: science - verifying theories by testing hypotheses derived from them.
  - Strong predictive power
  - Generalized models of cause-and-effect as basis
- **Constructivism**: science - seeking local theories that emerge from (and explain) the data.
  - Strengthens an understanding of complex situations
  - Categorizations and analogies

A simpler definition

A (good) Theory is the best explanation of all the available evidence

The Use of Theory

- Quantitative tests theories as explanations
- Qualitative studies may generate the theory
- Mixed methods studies may have no theories at all or a theoretical framework in which both quantitative and qualitative data are collected

Agenda for Today

- Paper reading presentation (HW2 - Theory)
- Interviews
  - Setting up
  - Conducting
  - Examples


8%
45%
17%
$
2%
28%

Heard It through the GITvine: An Empirical Study of Tool Diffusion across the *npm* Ecosystem

Hemank Lamba
Carnegie Mellon University
Pittsburgh, PA, USA

Asher Trockman
Carnegie Mellon University
Pittsburgh, PA, USA

Daniel Armanios
Carnegie Mellon University
Pittsburgh, PA, USA

Christian Kästner
Carnegie Mellon University
Pittsburgh, PA, USA

Heather Miller
Carnegie Mellon University
Pittsburgh, PA, USA

Bogdan Vasilescu
Carnegie Mellon University
Pittsburgh, PA, USA

Paper : https://cmustrudel.github.io/papers/lamba2020diffusion.pdf

Slides: https://docs.google.com/presentation/d/e/2PACX-1vTF9v_LHPChHrnztvctLNsCOUNpWN_QkrBREWyP6WDNiB8N7MTWj4F5y6M9ylrjxbm5oNu6CWqBkt_3/pub?start=false&loop=false&delayms=3000&slide=id.g9e982a32d1_0_75

Video: https://www.youtube.com/watch?v=3bn_BzO09mQ&feature=youtu.be

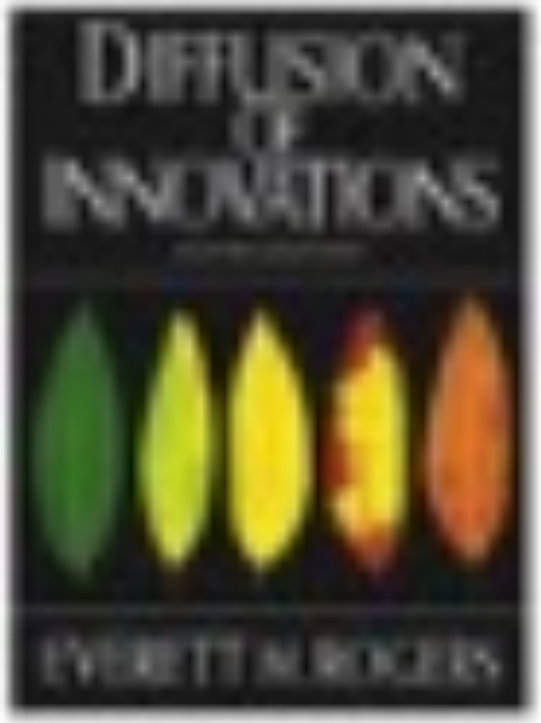
This Study ....

Quantitatively model tool diffusion factors

How visible are these tools ?

→ We compare different information diffusion channels presented by the platform.

Observability



Agenda for Today

- Paper reading presentation (HW2 - Theory)
- Interviews
  - Setting up
  - Conducting
  - Examples

Planning Checklist

- [x] Pick a topic
- [ ] Identify the research question(s)
- [ ] Check the literature
- [ ] Identify your philosophical stance
- [ ] Identify appropriate theories
- [ ] Choose the method(s)
- [ ] Design the study
  - Unit of analysis?
  - Target population?
  - Sampling technique?
  - Data collection techniques?
  - Metrics for key variables?
  - Handle confounding factors
- [ ] Critically appraise the design for threats to validity
- [ ] Get IRB approval
  - Informed consent?
  - Benefits outweigh risks?
- [ ] Recruit subjects / field sites
- [ ] Conduct the study
- [ ] Analyze the data
- [ ] Write up the results and publish them
- [ ] Iterate

Qualitative vs Quantitative

- Often:
  - Words (qual) vs Numbers (quant)
  - Open-ended questions (qual interview questions)

    vs

    Closed-ended questions (quant hypotheses)

Some common contrasts between quantitative and qualitative research	
Quantitative	**Qualitative**
Numbers	Words
Point of view of researcher	Point of view of participants
Research distance	Researcher close
Theory testing	Theory emergent
Static	Process
Structured	Unstructured
Generalization	Contextual understanding
Hard, reliable data	Rich, deep data
Macro	Micro
Behaviour	Meaning
Artificial settings	Natural settings

Bryman, 2008

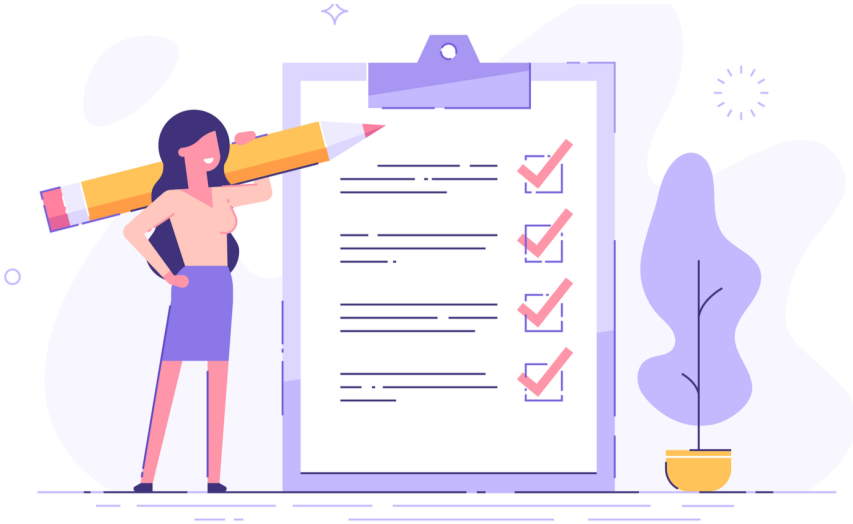
Data Collection Procedures – Qual. Study

- Identify the purposefully select individuals and sites for the study
- Indicate the number of sites and participants to be involved in the study
- Select the type(s) of data to be collected
  - Qualitative observations
  - Qualitative interviews
  - Qualitative documents
  - Qualitative audio-visual materials

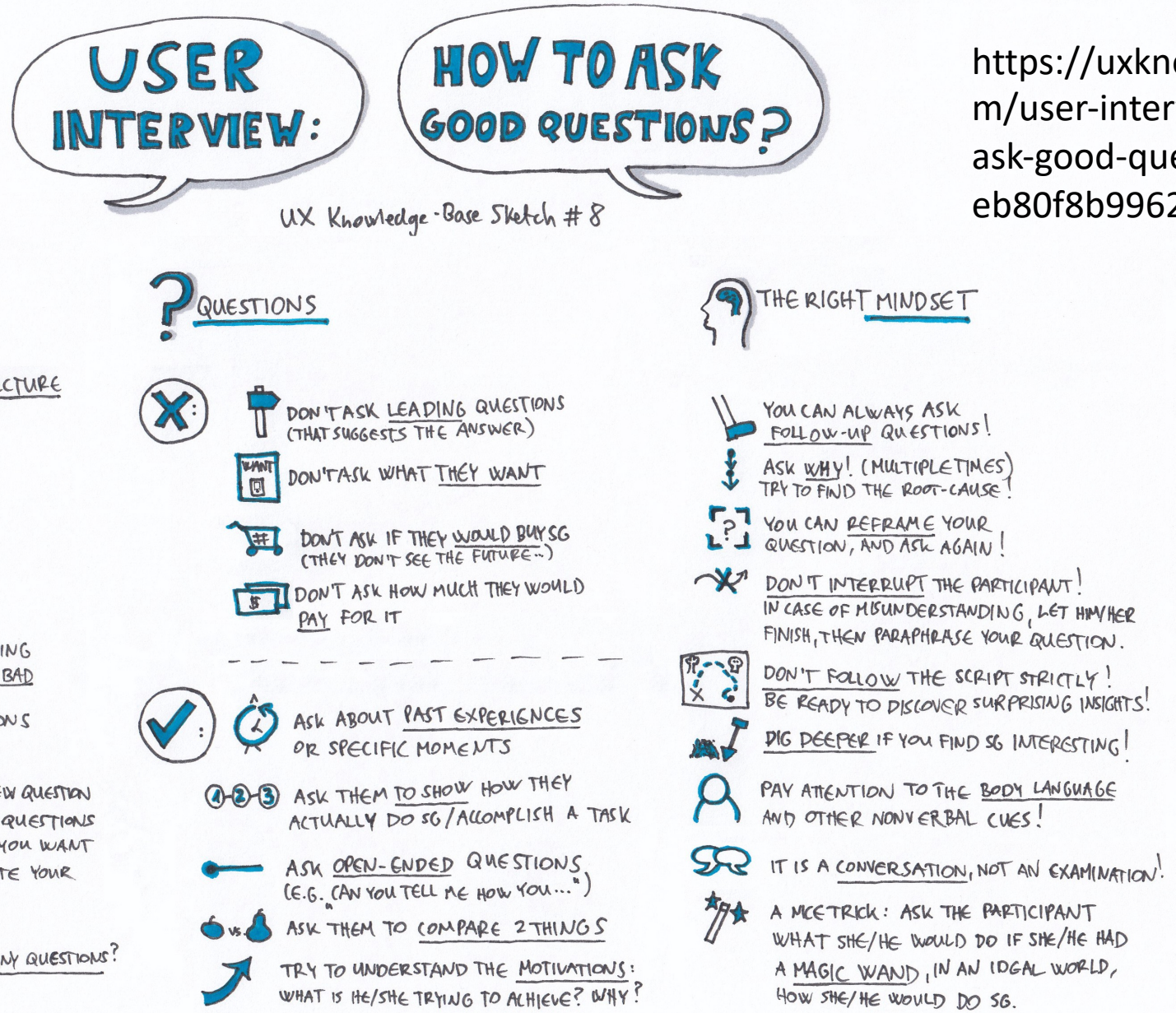
Creswell. Research Design.

Software Development process: Collecting requirements

- **Survey**: measure topics of interest in a controlled, consistent manner; easy to administer across large groups
  - Identify target population, their attitudes and preferences
  - Validate assumptions or facts
- **Interview:** More expensive, but could have follow-up questions to resolve ambiguity



UX design



https://uxknowledgebase.com/user-interview-how-to-ask-good-questions-eb80f8b99627

Interview

Interviews

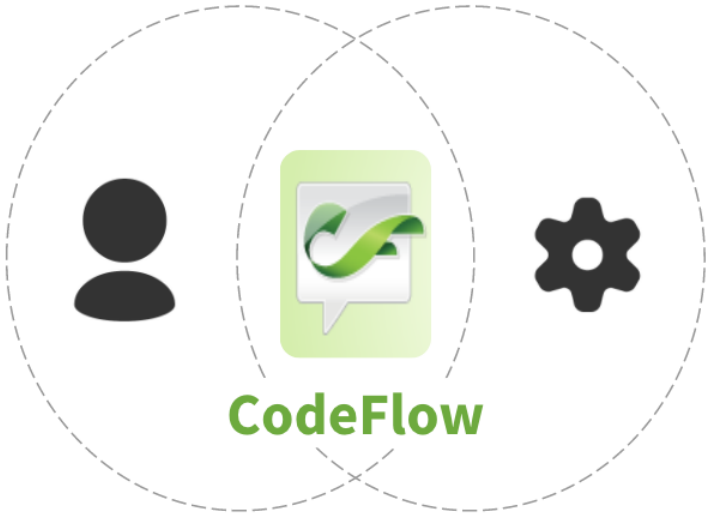
- (Most) common method of data gathering in qualitative research
- A variety of forms of qualitative research interview

(and assumptions that underlie their use)

- Types of interviews:
  - ‘semi-structured’ -- list of questions (open-ended and closed-ended) or topics
  - ‘un-structured’ -- list of prompts

Why Interviews?

Chris Bird @Microsoft Research



https://www.cabird.com/

Socio-Technical

CodeFlow: Improving the Code Review Process at Microsoft

A discussion with Jacek Czerwonka, Michaela Greiler, Christian Bird, Lucas Panjer, and Terry Coatta

Joint Optimization – Code Review

CodeFlow: Improving the Code Review Process at Microsoft, Czerwonka et al. 2018.

https://www.cabird.com/

Chris Bird @Microsoft Research

- “I was investigating code review latency at Microsoft when I found something odd in the data for a team in Bing: Many of their code reviews were signed off in just minutes (sometimes under a minute) after the code review was created.
- “I meticulously looked at the data collection code, I conducted a number of statistical tests on the data based on guesses that I had. Nothing
- “As a last resort, I contacted one of the developers on the team and scheduled an interview. She explained that the reason for the lightning fast reviews was that they often conduct code reviews in person with two or three reviewers huddled around the developer's screen as they explain the change. Once the reviewers were happy, the author would create the request in the review system and the reviewers would immediately sign off on the review.
- “In just a few minutes, she had answered a question that I hadn't been able to answer after hours of testing hypotheses on data. You can learn things in an interview that you would never have thought of yourself.”

General PROs

- Interviews allow rich engagement and follow up questions
- Collect historical data that is not recorded anywhere
- Elicit opinions and impressions in richer detail than people would provide through written communication
- Information from interviews can be triangulated with other data sources
- Interviews can be used to clarify things that have already happened (especially following an observation).

How to Break an API: Cost Negotiation and Community Values in Three Software Ecosystems

We interviewed developers in Eclipse, R/CRAN, and Node.js/npm about how they distribute the costs of breaking change. Community values, practices, and tools differ significantly among ecosystems. With Eclipse valuing stability, R/CRAN valuing ease and rapid access for end users, and Node.js/npm valuing ease for developers, they each adopted distinct practices and policies."

Why Not Interviews?

General CONs

- Usually small sample size
- The time required for each individual interview
- The challenge of finding appropriate interviewees and scheduling a time that works for all parties
- Potential bias introduced by the interviewer during the interview (word choice, tone of voice, or even body language can potentially affect responses)
- The time required for transcription and subsequent analysis.

Goals and characteristics

- Goal: to see the research topic from the perspective of the interviewee, and to understand how and why they come to have this particular perspective.
- Characteristics:
  - low degree of structure imposed by the interviewer
  - preponderance of open questions
  - focus on specific situations and action sequences in the world of the interviewee rather than abstractions and general opinions

Relationship between interviewer and interviewee

Interviewee as research subject:

- minimize the impact of inter-personal processes on the course of the interview
- typically quant research

Interviewee as research participant:

- actively shaping the course of the interview rather than passively responding to the interviewer's pre-set questions
- typically qual research

Steps

1. Defining the research question (see previous lecture)
   - Typical focus is on how participants describe and make sense of particular elements of their lives.
   - Goal is *not* to quantify individual experience
   - Avoid reflecting your own presuppositions or biases
2. Creating the interview guide (protocol)
3. Recruiting participants
4. Carrying out the interviews

2. Interview guide

- Not formal schedule of questions to be asked word-for-word in a set order
- Instead, list:
  - topics the interviewer should attempt to cover
  - probes which may be used to follow-up responses and elicit greater detail from participants
  - from open to more specific
- Guide can evolve after each interview:
  - adding probes / topics not originally included, that emerged spontaneously in interviews
  - dropping or re-formulating those which are incomprehensible or consistently fail to elicit relevant responses

Questions:
Simply, easy, short, no jargon

Interview Guide – Formulating questions

1. **Decisions about Question Content**

   - **Is question necessary? How will it be used?**
   - **Are several questions necessary? Do not put multiple ideas in the same question.**
   - **Do respondents have information to answer question?**
   - **Will respondents be willing to answer question?**

Paul Goodman's (2005) Building Effective Interviewing Skills.

Interview Guide – Formulating questions (Cont.)

2. Decisions about Wording

- **Is the language unclear or difficult to understand?**
- **Is the language specific enough in terms of when, where, etc.?**
- **Is the question leading or biased in one direction?**
- **Does the wording include unstated assumptions?**

Paul Goodman's (2005) Building Effective Interviewing Skills.

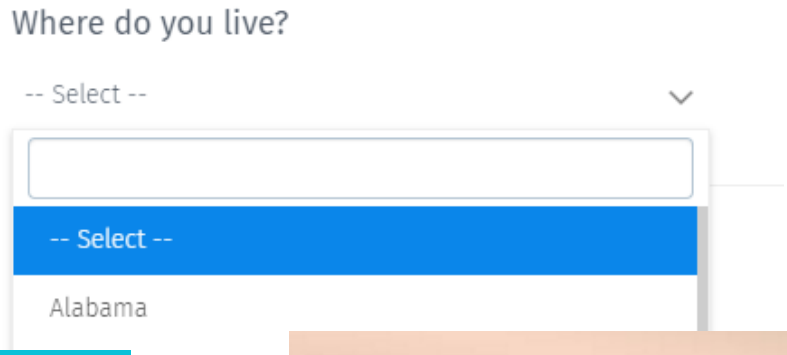
Interview Guide – Formulating questions (Cont.)

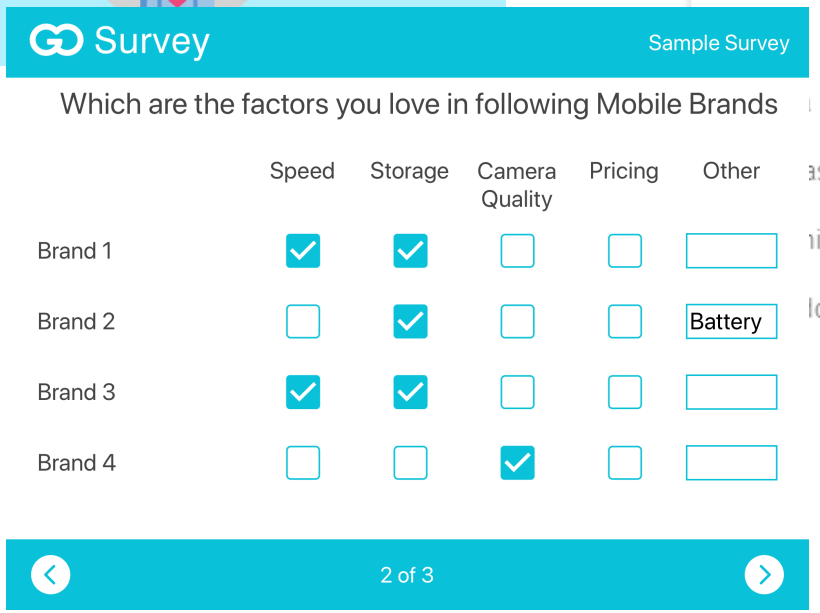
3. Decisions about Form of Response

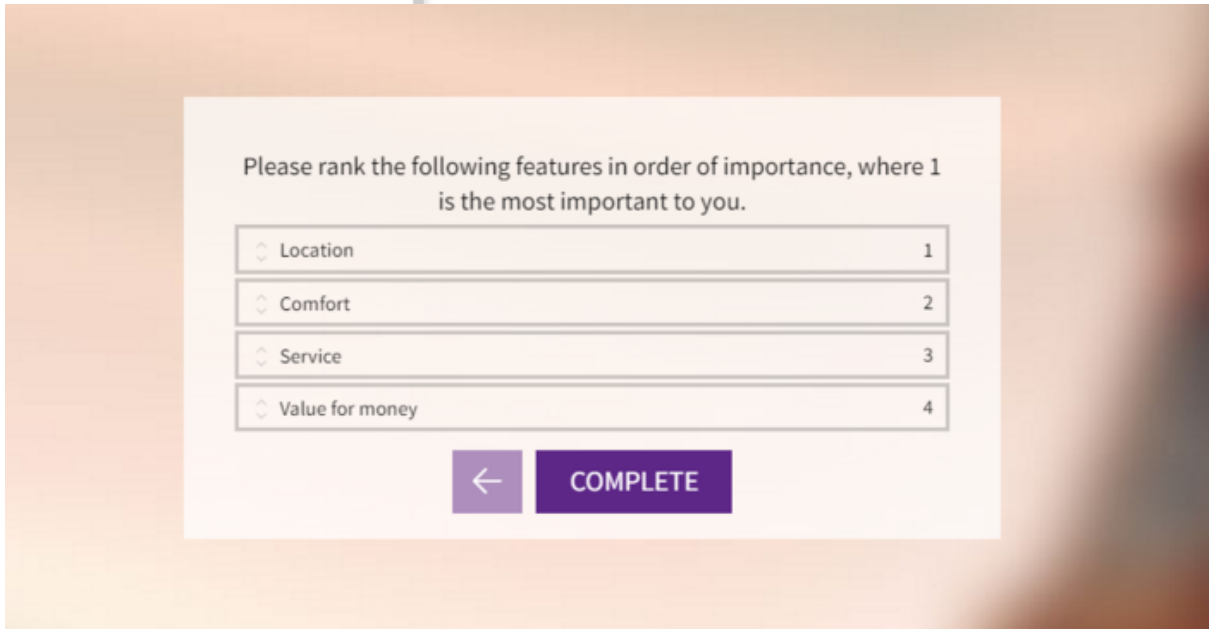
- **What is the form for response – open end, Likert scale?**
- **What is the form of any fixed response in terms of numbers of categories, labels?**
- **Are the response categories symmetrical or asymmetrical?**
- **Are the labels clear, understandable, etc.?**
- **See #2.**

Paul Goodman's (2005) Building Effective Interviewing Skills.

Closed-ended Questions







Closed-ended Questions

- **Nominal scales** provide interviewees with a list of categories from which to select their answer (e.g., White, Black or African American, American Indian, Asian, Native Hawaiian or Pacific Islander)
- Good practices –
  - Solicit response options in a pilot study
  - Randomize order, if concerned about order effects
  - Avoid bias from unequal response options
  - Check all that apply vs. forced-choice

Example: Unequal response options

How likely are you to share your location to meet friends after work?

- Absolutely never
- Sometimes
- Occasionally
- Once or more a week
- Everyday

Is it easy or difficult to distinguish between these three categories?

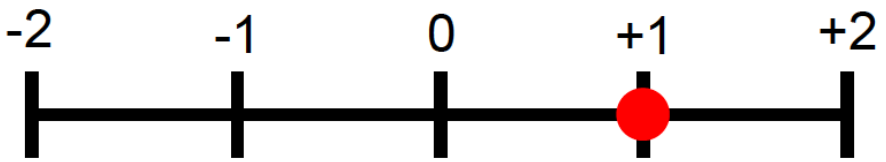
If difficult, why?

Cowley, Youngblood. “Subjective response differences between visual analogue, ordinal and hybrid response scales,” *Human Factors and Ergonomics Society Annual Meeting*, 53(25): 1883-1887, 2009.

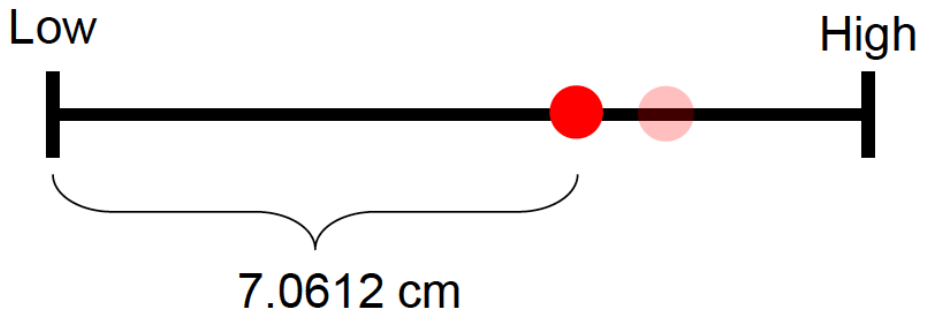
Ordinal Scales

- Ordinal or interval scales ask interviewees to choose a “level” of the variable of interest

Numbered Scale:
(choose your number)



Visual Analogue Scale:
(mark your level)



How do you use an ordinal scale?

- *Understand nuanced opinions.* Do respondents “agree” or “strongly agree” with a stance on an issue?
- *Uncover perceptions.* Do respondents find a particular statement “false,” “mostly false,” “mostly true,” or “true”?
- *Measure relative performance.* Is a certain employee “more productive,” “just as productive,” or “less productive” than other employees?
- *Gauge sentiment.* Is a customer “very satisfied,” “satisfied,” “dissatisfied,” or “very dissatisfied” with a recent purchase?

Employee Satisfaction Survey Template

How meaningful is your work?

- Extremely meaningful
- Very meaningful
- Moderately meaningful
- Slightly meaningful
- Not at all meaningful

How challenging is your job?

- Extremely challenging
- Very challenging
- Moderately challenging
- Slightly challenging
- Not at all challenging

Open-ended Questions

Exploratory in nature, and offer the researchers rich, qualitative data. In essence, they provide the researcher with an opportunity to gain insight on all the opinions on a topic they are not familiar with.

Open-ended Questions

- **Definition and designation questions**
  - What-is asks to develop definitions of things
  - Who identifies the responsible agent
  - What-kinds-of ask for possible types and exemplars

- **Process, event and exception questions**
  - How-to ask how an action is performed
  - When asks about timing constraints, pre-and post-conditions
  - What-if asks about failures or unexpected events
  - Follow-on questions result from answers from previous questions

- What do you like most about our new product?
- What changes would most improve our product?

1. How often do you attend events in this neighborhood?

- Extremely often
- Quite often
- Moderately often
- Slightly often
- Not at all often

2. If you do not attend events in this neighborhood, why not?

Interview Guide – Formulating questions (Cont.)

4. Decisions about Question Order

- **Do the initial questions easily launch the interview?**
- **Is there any order effects in terms of bias, priming, in the order of the questions?**

Paul Goodman's (2005) Building Effective Interviewing Skills.

Steps

1. Defining the research question (see previous lecture)
   - Typical focus is on how participants describe and make sense of particular elements of their lives.
   - Goal is *not* to quantify individual experience
   - Avoid reflecting your own presuppositions or biases
2. Creating the interview guide (protocol)
3. Recruiting participants
4. Carrying out the interviews

3. Recruiting participants

- Considerations:
  - Amount of time and resource available
  - Diversity of expected views
    - Representativeness of the sample
    - Depends on the design for which the interviews are being used; maximum variation is good default
  - Confidentiality
  - Avoid interviewing ”easy access” subjects
    - Need to have enough distance from interviewees so that they take nothing for granted

Recruiting participants (cont.)

- Have answers ready for:
  - Why are you doing this study?
  - What do you get out of it?
  - Will I receive the data?
  - How do I know it will be confidential?
  - What does university X / company Y get out of this?
  - Who is paying for this?
  - What will I get out of this?
  - How long will this take?
  - Are you doing other parts of the company?
  - Why did you select my group / me?

- *Targeting:* Reaching the right audience
- *Screening:* Making sure you recruit people in a reliable and unbiased way
- *Incentivizing:* Offering incentives which encourage uptake by the right audience
- *Sustaining:* Making sure you can keep doing research for the long-term.

Dear *[name]*,

We have found your fork (*kory75/Marlin_360*) that has done some changes that have not been merged to the main project (*MarlinFirmware/Marlin*). We have designed a prototype that analyzes forks' commits and identifies potential reusable features implemented in these commits. We are looking for participants for our study and we would like to ask you if you would accept to have a discussion (over Skype) that should take no more than 60 minutes. In this discussion, we will show you the tool and ask a few questions (e.g., does the tool identify correctly the implemented feature?).

My name is Shurui Zhou and I am a Ph.D. student at the Institute for Software Research at Carnegie Mellon University. We have collected 235 forks from Marlin project that have commits have not been merged to the main branch, and we obtained your contact information from your Github profile.

If you agree to discuss with us, I would like to schedule a meeting with you according to your convenience.

The study description is pasted below. Please let me know if you're interested!

Thank you very much,

Shurui Zhou
Institute for Software Research
Carnegie Mellon University
shuruiz@cs.cmu.edu

Helping Developers Help Themselves: Automatic Decomposition of Code Review Changesets

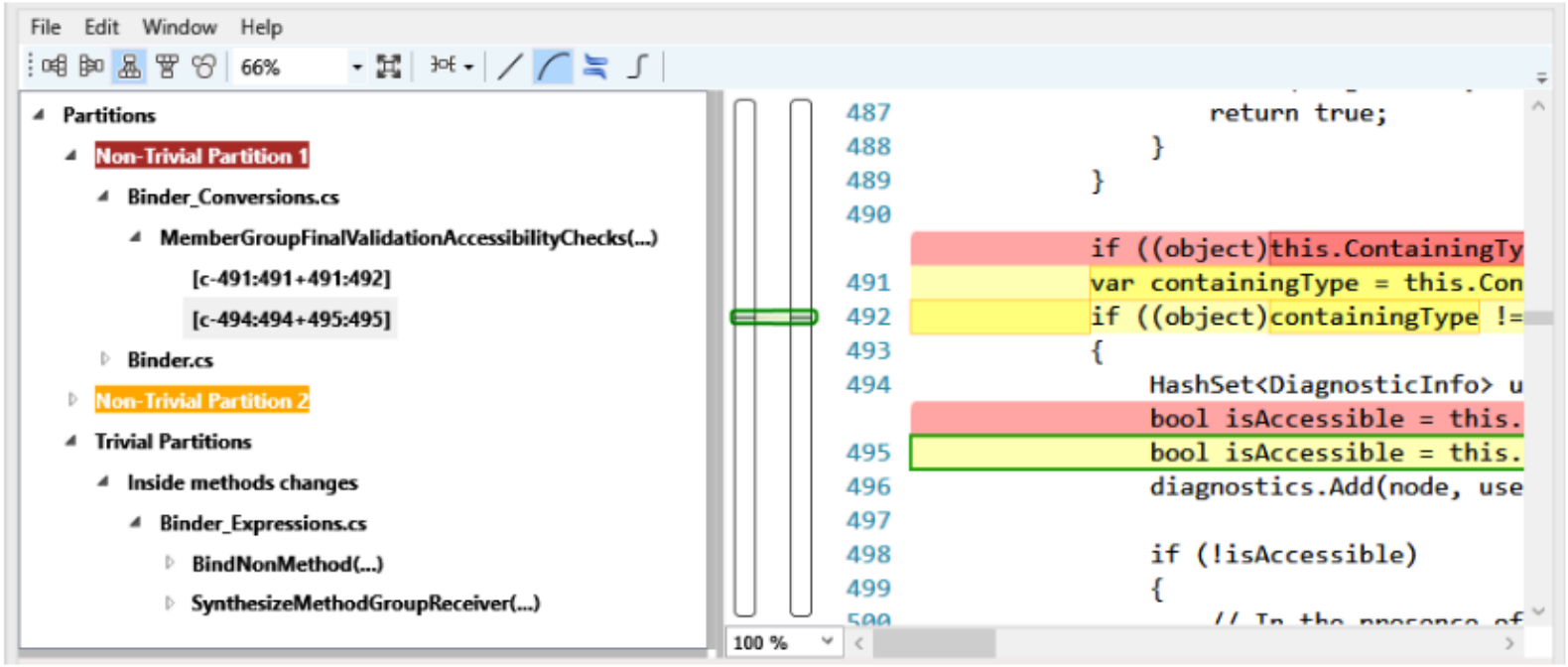
Mike Barnett
Microsoft Research
Redmond, WA, USA
mbarnett@microsoft.com

Christian Bird
Microsoft Research
Redmond, WA, USA
cbird@microsoft.com

João Brunet
Federal University of Campina Grande
Campina Grande, Paraíba, Brazil
joao.arthur@computacao.ufcg.edu.br

Shuvendu K. Lahiri
Microsoft Research
Redmond, WA, USA
shuvendu@microsoft.com

- a firehouse research study design

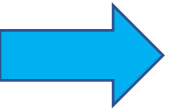


How Many Participants are Enough?

- Two Criteria for enough:
  - Sufficiency: Are there sufficient numbers to reflect the range of participants and sites that make up the population so that others outside the sample might have a chance to connect to the experiences of those in it?
  - Saturation of Information: a point in a study at which the interviewer begins to hear the same information reported.

Steps

1. Defining the research question (see previous lecture)
   - Typical focus is on how participants describe and make sense of particular elements of their lives.
   - Goal is *not* to quantify individual experience
   - Avoid reflecting your own presuppositions or biases
2. Creating the interview guide (protocol)
3. Recruiting participants
4. Carrying out the interviews

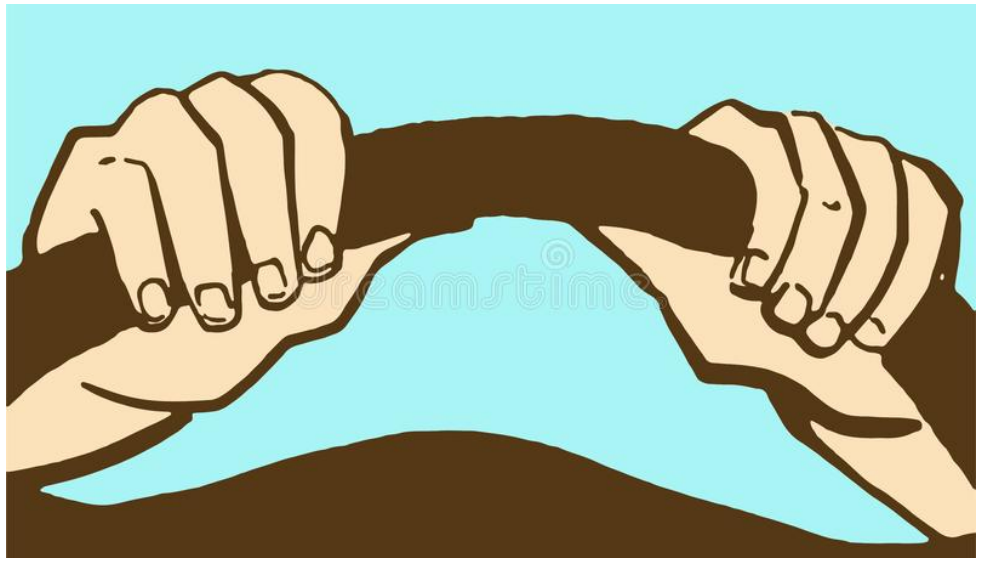


EXAMPLE

https://www.youtube.com/watch?feature=oembed&v=9t-_hYjAKww

Criteria

- Steering
- Open
- Critical
- Interpret
- …



A real-life example: evaluating the Calderdale and Kirklees Out of Hours Protocol for Palliative Care

RQ: From the perspectives of General Practitioners and District Nurses, how effective has the Out of Hours Protocol been in improving the provision of out of hours care for community palliative care patients – in particular in the areas of communication, carer support, specialist medical support and drug/equipment availability?

King, N. (2004). Using interviews in qualitative research. In C. Cassell & G. Symon (Eds.), Essential Guide to Qualitative Methods in Organizational Research (pp. 11-22). Loondon: Sage.

Table 2.1 Extract from the district nurse interview guide

Support issues

4 To what extent is the out of hours provision addressing the needs of the patients?

Probe: *In what ways is it/is it not?*

5 Do the patients have enough support?

Probe: *Who provides support?*

Probe: *What else (if anything) could be done?*

6 Do you feel that the patients and/or carers know who to contact in a crisis?

Probe: *If YES, how? If NO, why not?*

7 In your opinion, what – if anything – has been the effect of the protocol on continuity of care?

Ask each participant

13 Give an example (if you can) of a case where poor communication led to problems for a patient/carer.

14 Give an example (if you can) of a case where good communication halted potential problems.

King, N. (2004). Using interviews in qualitative research. In C. Cassell & G. Symon (Eds.), Essential Guide to Qualitative Methods in Organizational Research (pp. 11-22). London: Sage

Practical issues

- Avoid:
  - Multiple questions:
    - ‘Why did you join this open-source project, and do you think it has brought benefits to your programming experience?’
  - Leading questions:
    - ‘So you felt that using this tool improved your productivity?’
  - Assuming that the answer to a question is so obvious that it need not be asked:
    - ‘Whether, and to what extend, are you concerned about your privacy online?’
  - Imposing your perception:
    - ‘So what you’re really saying is …’
  - Ending the interview on a difficult, threatening or painful topic
    - E.g., finish by giving the interviewee the opportunity to make any comments about the subject at hand which have not been covered in the rest of the interview: ‘What else, if anything, should I have asked?’

Practical issues (Cont.)

- Do:
  - Be flexible:
    - Topic order may change during interview
  - Open with a question which can be answered easily and without potential embarrassment or distress
    - E.g., requests for factual or descriptive information
    - Relax and build rapport first
  - Listen more, talk less
    - Control mechanism: listen to yourself in recordings / check length of your paragraphs in transcripts
  - Follow up on what the participant says
    - Clarifications, details, stories
    - Trust your instincts, explore emerging directions
  - Ask participants to reconstruct, not remember, their experience:
    - 'What happened?' vs 'Do you remember what happened?'
    - Ask for concrete details

Difficult interviews

- Uncommunicative interviewee
  - Symptom: monosyllabic answers
  - Helps to be clear about required time and anonymity
  - Phrase questions as open as possible
  - Use silence
- Over-communicative interviewee
  - Symptom: repeatedly straying far from your questions without adding anything of significant interest
  - Politely interrupt the digression at a natural pause or break and refer back to an earlier point:
  - ‘That’s very interesting. Could we go back to what you were saying earlier about […], I’d like you to tell me more about that’

Difficult interviews (Cont.)

- High-status interviewees:
  - Don’t appear over-familiar
  - Nor overly nervous or submissive
  - Instead, be respectful but confident of the worth of what you are doing and of *your* own expertise.
- Emotionally charged topics
  - Give people the time they require to answer your questions
  - Avoid non-verbal cues that might be taken as indicating impatience
  - Skip question, return later if possible

Requirement Elicitation [Chef Co-pilot web app]

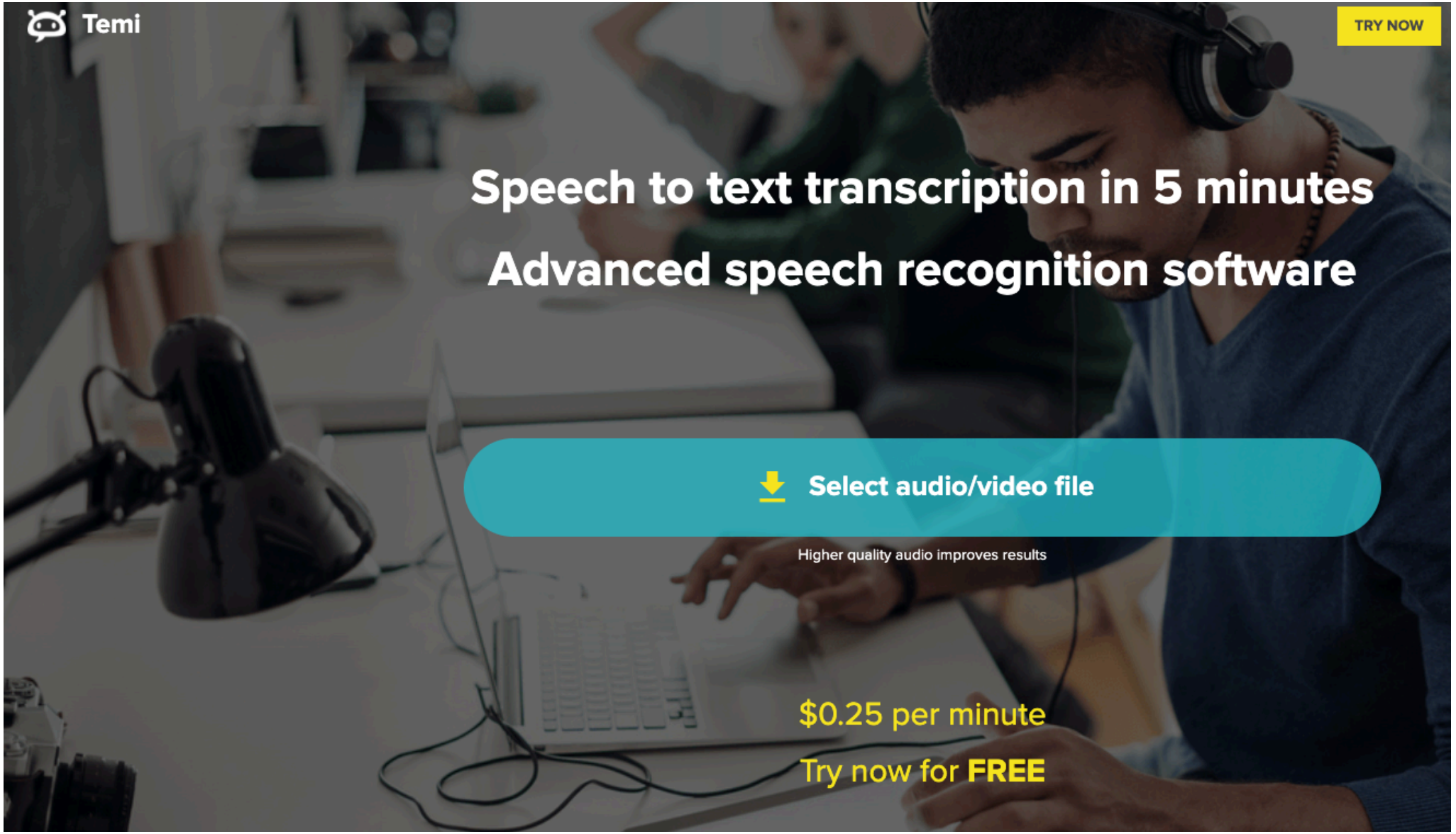
interview audio sample 1

interview audio sample 2

Recordings and transcripts

- Record the interview, capture everything verbatim
- But still take copious notes during the interview.
  - Much easier to refer to notes than find a particular place in a recording.

- Whatever can go wrong, will go wrong


Temi
TRY NOW
Speech to text transcription in 5 minutes
Advanced speech recognition software
Select audio/video file
Higher quality audio improves results
$0.25 per minute
Try now for FREE

References

• King, N. (2004). Using interviews in qualitative research. In C. Cassell & G. Symon (Eds.), Essential Guide to Qualitative Methods in Organizational Research (pp. 11-22). London: Sage.

• Seidman, I. (2012). Interviewing as qualitative research: A guide for researchers in education and the social sciences: Teachers college press. (Ch 4, Ch 6).

• Paul Goodman's (2005) Building Effective Interviewing Skills.

• Chris Bird’s (2016) interview guide. Published in Perspectives on Data Science for Software Engineering

Learn by **DOING**.

HW3 （see details on Quercus）

Think about a research idea that related to your own area and you need to collect information from stakeholders through interview.

1) decide what your purpose is and write a sentence describing it.

2) develop an interview protocol. The protocol can be short, focusing on exactly what you are interested in. You should anticipate short interviews, perhaps 15-20 minutes at most.

3) conduct two interviews.

4) be prepared to tell the class what you learned, how the interviews went, any problems or lessons you can share. In future classes, we will learn more structured ways of analyzing qualitative data such as interview transcripts.

Examples

- Grinter, Rebecca E., and Leysia Palen. “ Instant messaging in teen life .” Proceedings of the 2002 ACM conference on Computer supported cooperative work . 2002.
- Chattopadhyay, Souti, et al. " What's Wrong with Computational Notebooks? Pain Points, Needs, and Design Opportunities ." Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems . 2020.