

# Introduction to HCI

## Qualitative Data Analysis

**Prof. Narges Mahyar**  
**UMass Amherst**

[nmahyar@cs.umass.edu](mailto:nmahyar@cs.umass.edu)

Courses, projects, papers, and more:

<http://groups.cs.umass.edu/nmahyar/>

# Today

- Type of data [10min]
- Qualitative analysis techniques [20min]
- In class activity [20min]
  - Coding
- Project discussions [20min]

# Learning goals

- Explain types of data that HCI methods provide
- Explain what it means to triangulate in data gathering and analysis
- Understand theme-based approaches to analysis
- Understand thematic analysis as one of the most common forms of analysis and describe how to do thematic analysis in detail
- Understand why reliability and validity are important in qualitative research
- Make and justify strategic decisions in evaluation planning

# Types of data that HCI methods provide

- **Qualitative:**

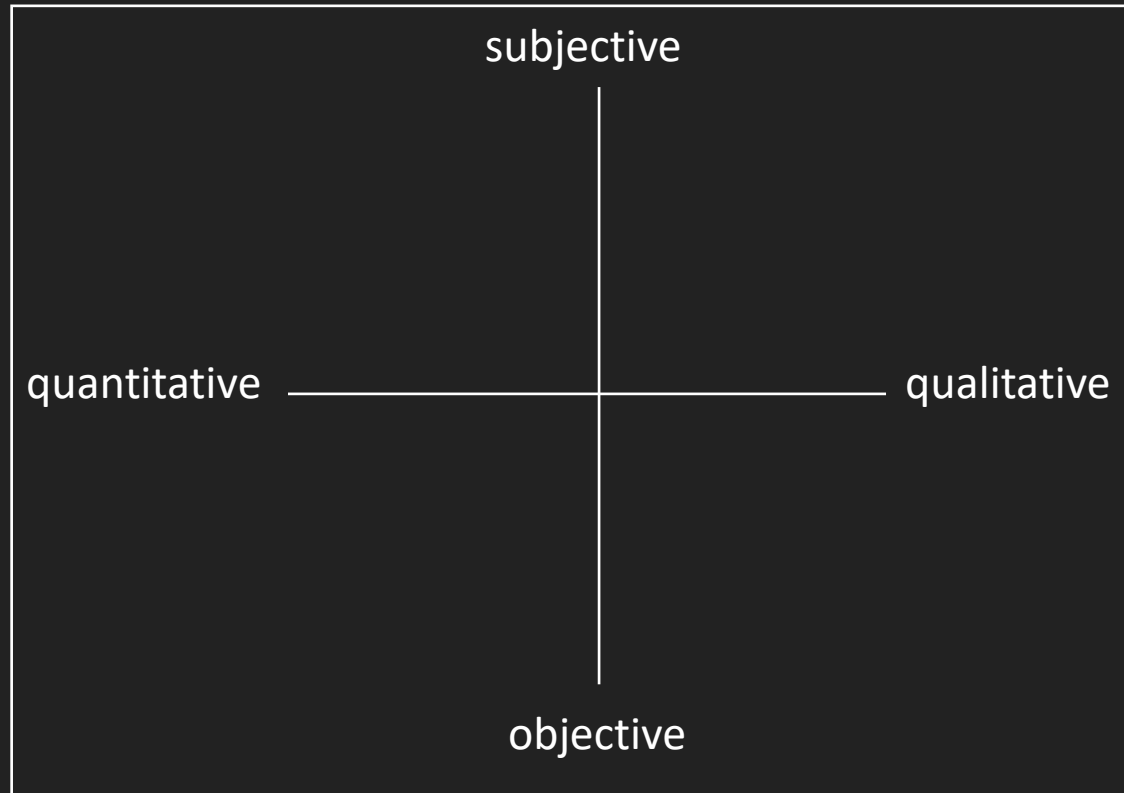
- *Users describe/report X, to extent they are aware*
- *You observe X, that users may not be fully aware of*
- *Where X can be: behaviors, processes, usability challenges...*

- **Quantitative:**

- Measure task performance with existing tools / methods:
  - e.g., Speed, errors, dead-ends, learning curves for novices ...
- Numerical data from user-reported answers: e.g., # of emails/day
- Counting observed occurrences: e.g. # of times looked at instruction

# Activity:

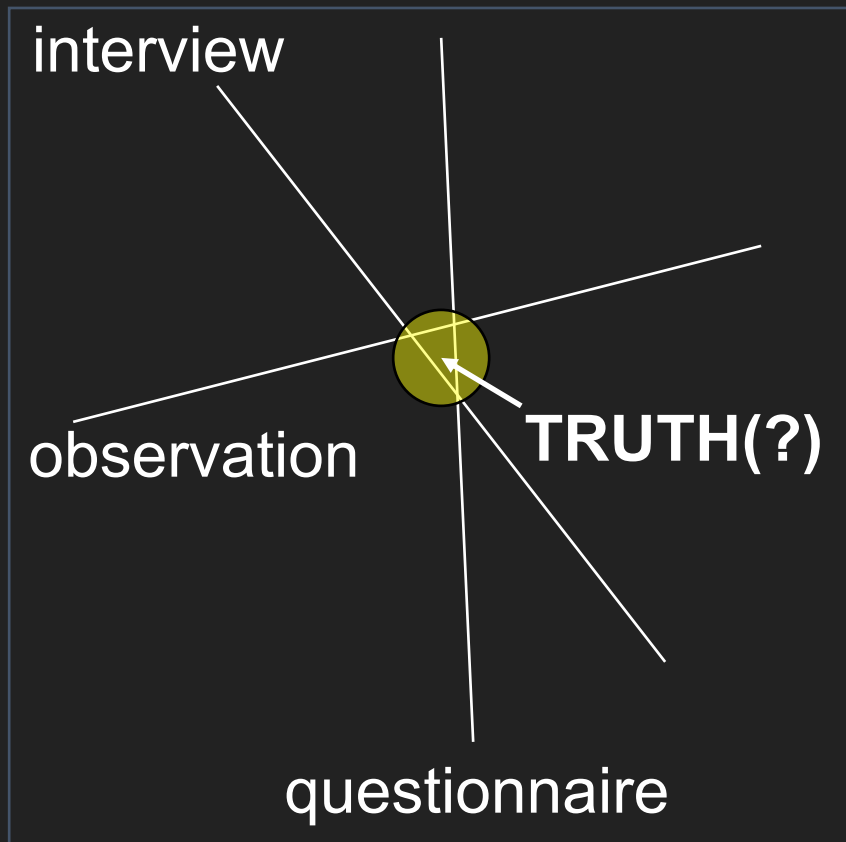
types of data that HCI methods provide



- controlled observations – of time to complete task?
- unstructured observations – of observed steps to complete a task?
- unstructured interviews – of user telling stories?
- questionnaire – self report numbers of times do something?

# Triangulation

- a strategy to enhance validity:  
use the *multiple perspectives* available from complementary sources



Use multiple:

- data sources  
people, places, times
- data collection methods
- researchers/evaluators

# Analyzing & interpreting data

**Qualitative data** – interpreted to tell a “story”

- Categories, themes, patterns, etc.

**Quantitative data** – presented as values, tables, charts and graphs

- Counts (e.g., Summary of total # of errors)
- Simple statistical analysis (e.g., Averages)
- Advanced statistical analysis (e.g., Linear regression)  
... more on quantitative coming later this term.

# Methods for qualitative analysis

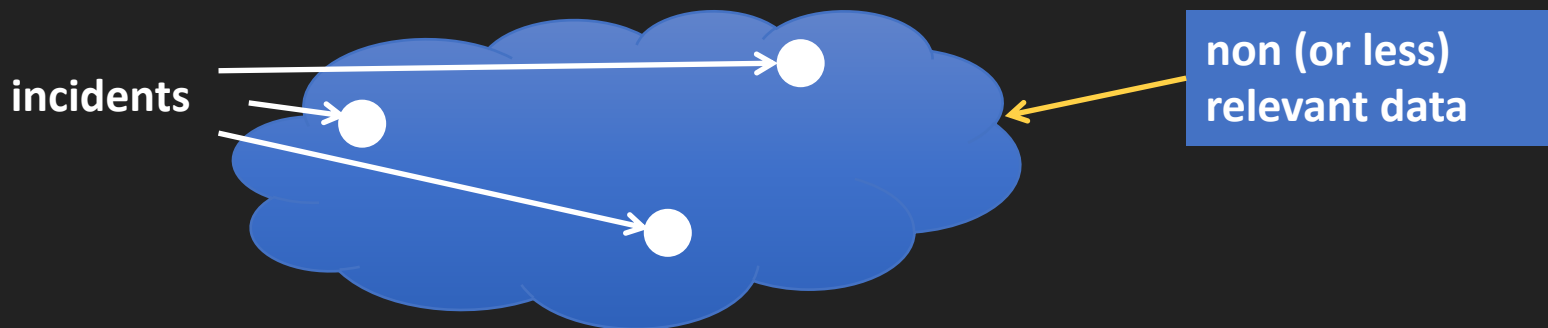
- From simpler (less effort) to more advanced (more effort):
  - Identify critical incidents
  - Simple categorization
  - Categorization - themes, patterns
- In all cases your aim is to interpret the data in ways that encapsulate and document your understanding.
- Level of effort depends on your goal.
- Many methods often used in combination.



# Qualitative analysis methods

## *finding critical incidents*

- **Identify and focus on the most significant incidents**
  - Efficient when you have lots of data
  - Incidents can be either desirable or undesirable
- NOT about summarizing all incidents that occur
  - **More like finding tiny gold nuggets in buckets of sand**
  - Appropriate for usability studies, where qualitative is not primary analysis



# Qualitative analysis methods

## *categorizing data*

- Typically used on transcripts (observations, interviews, etc.)
  - At non-granular level of detail to find stories or themes
  - At fine-grain level of detail focusing on words, sentences, gestures, etc. (E.G., Discourse analysis)
- First data are 'coded' according to a scheme of categories
  - Can be predetermined, or arise from the data
    - At a high level: ***affinity diagram***

## Example of coded transcript mid- to low-level of detail

own home  
alone / lonely  
When you move into your own home, you're alone. There is no bustle of people around the  
miss company  
house. I miss having someone to chat to when I get home. I put the TV or some music so  
background noise  
there's some background noise. the silence makes me feel so alone. Sometimes I will be sat  
lonely  
watching trash TV and thinking I should be out doing something rather than watching this  
wasteful time / inactive  
rubbish. I read a lot but sometimes I am too tired and just want to veg out. But it's been good  
doing  
bored / depressed  
to move out of mum and dad's as it's not healthy to rely on them as they won't last forever. I  
unhealthy to be dependent  
become independent and made my own decisions. It's good they still there when I need them.  
independence  
support  
distance  
conflict  
It's good to have some distance as when I was at home I was arguing a lot with my dad and  
moving out  
that was made me decide it was time to go.

Very informative, but time consuming!

# Thematic analysis

- Thematic analysis is one of the most common forms of analysis in qualitative research.
- It is a method for identifying, analyzing and reporting patterns (themes) within data. It minimally organizes and describes your data set in (rich) detail.

Virginia Braun & Victoria Clarke (2006) Using thematic analysis in psychology, *Qualitative Research in Psychology*, 3:2, 77-101

# Qualitative analysis methods

## *patterns and themes*

- **Can be revealed in many ways:**
  - Through the process of conducting the study
  - Use of tools and techniques (e.g., Affinity Diagrams)
- **Can support many types of user study goals**
  - E.G., Understanding behavior, culture, places or situations where events occur, breakdowns, user characteristics, etc.
- **Very flexible and widely used**
  - Can be reported as findings or inform more analysis

# Inductive vs. theoretical thematic analysis

- Inductive or 'bottom up' way
  - Data-driven; process of coding the data without trying to fit it into a preexisting coding frame, or the researcher's analytic preconceptions
- Theoretical or deductive or 'top down' way
  - Analyst driven; driven by the researcher's theoretical or analytic interest in the area

# A closer look at note taking in the co-located collaborative visual analytics process



**Narges Mahyar**, Ali Sarvghad, and Melanie Tory, “A closer look at note taking in the co-located collaborative visual analytics process,” *IEEE VAST 2010*. [TOP 4 papers]

# Exploratory user study

- 27 participants (9 groups of 3)
- 2 set of tasks
- 4 groups: tabletop
- 4 groups: wall display
- 1 group: both



## Data gathered

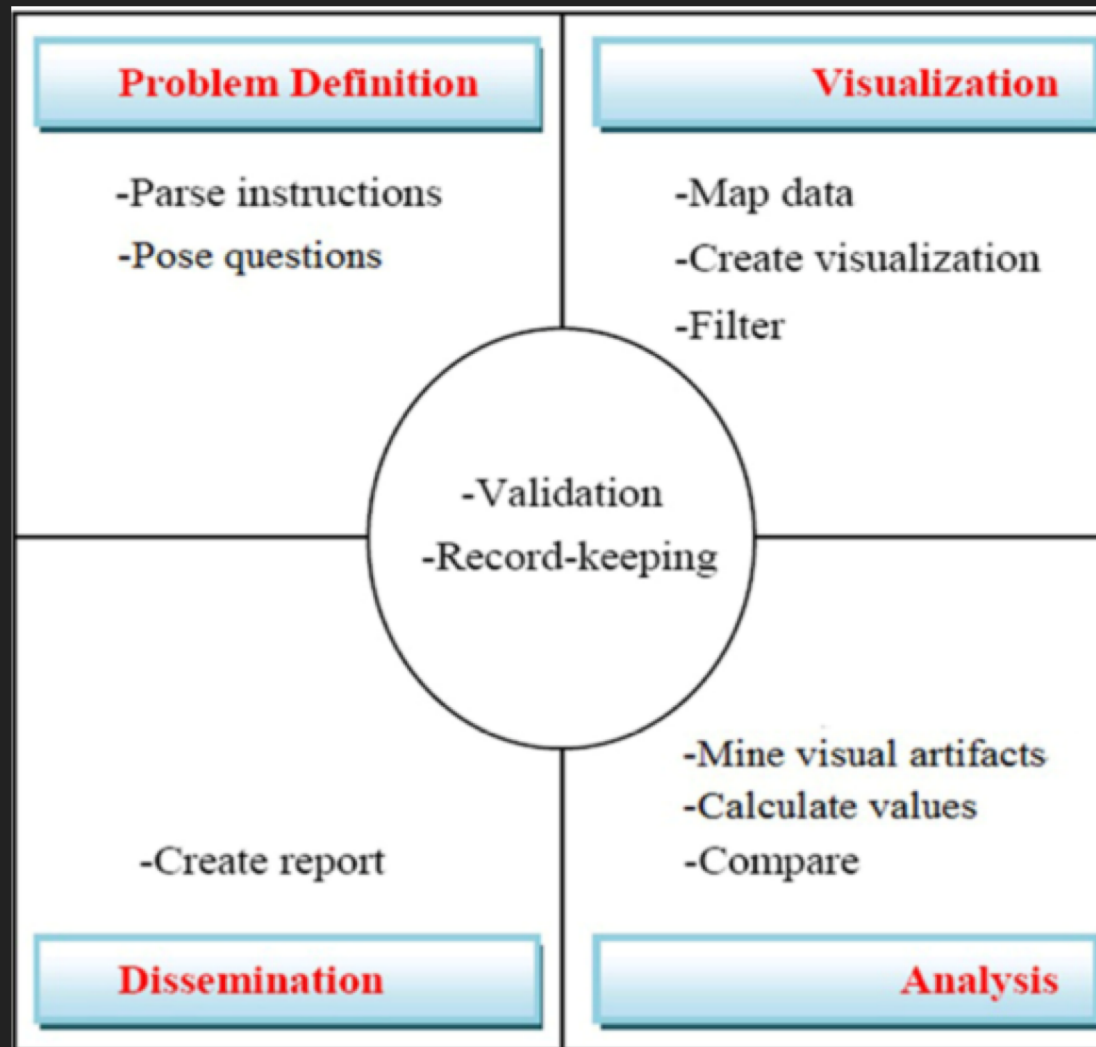
- ~630 min of video
- screen logs were captured (~ 70 min per session)
- ~20 min of interviews per session

# Record-keeping actions

**Table 1.** Number of note-taking and chart-saving actions by each group. Shaded groups relied heavily on saved charts for analysis.

Group		1	2	3	4	5	6	7	8	9
Display used		Table				Wall				Both
Task 1	Number of note-taking actions	5	2	4	3	0	13	18	0	5
	Number of charts saved	0	0	0	2	2	0	4	0	2
	Designated note-taker	N	N	N	N	N	Y	N	N	Y
	Number of note-takers	2	1	1	3	0	1	3	0	1
	Number of times all took notes simultaneously	0	0	0	1	0	0	8	0	0
Task 2	Number of note-taking actions	8	4	7	7	20	8	11	9	6
	Number of charts saved	4	22	8	12	3	8	7	2	2
	Designated note-taker	N	N	N	N	Y	N	N	Y	Y
	Number of note-takers	3	1	3	3	3	3	3	1	2
	Number of group notes	0	0	0	1	1	0	0	1	1

**Narges Mahyar**, Ali Sarvghad, and Melanie Tory, “A closer look at note taking in the co-located collaborative visual analytics process,” *IEEE VAST 2010*. [TOP 4 papers]



**Narges Mahyar**, Ali Sarvghad, and Melanie Tory, "A closer look at note taking in the co-located collaborative visual analytics process," *IEEE VAST 2010*. [TOP 4 papers]

# Doing thematic analysis: a step-by-step guide

1. Familiarizing yourself with your data
2. Generating initial codes
3. Searching for themes
4. Reviewing themes
5. Defining and naming themes
6. Producing the report

# 1. FAMILIARIZING YOURSELF WITH YOUR DATA

- Transcribing data, reading and re-reading the data, noting down initial ideas.
- **Note:**
  - writing is an integral part of analysis
  - analysis is not a linear process
  - read through the entire data set before you begin your coding

## 2. GENERATING INITIAL CODES

- Coding interesting features of the data in a systematic fashion across the entire data set, collating data relevant to each code.
  - your coded data differ from the units of analysis (your themes), which are (often) broader.
- coding will, to some extent, depend on whether the themes are more 'data-driven' or 'theory-driven'

# Coding manually

Code your data by writing notes on the texts you are analyzing, by using highlighters or colored pens to indicate potential patterns, or by using 'post-it' notes to identify segments of data. Key advice for this phase is:

(a) Code for as many potential themes/patterns as possible

(b) Code extracts of data inclusively / i.e., Keep a little of the surrounding data if relevant

4 Creativity: Where's The Inspiration?  
5 Honestly, I find it **tough to be creative in a classroom** sometimes. Not that I lack a sense  
6 or want to be creative, I get tons of ideas jumbled in my head and many notebooks filled with  
7 stuff. But I find that students just get into their routines and day-in-the-easiness of a class that  
8 doing or trying **something creative can sometimes be a forlorn hope**. **It's not really a habit**  
9 **and it's not a task**. It doesn't require much thought, they can **chat about other class things**,  
10 and feel that it's an **easy A**. That's a tough challenge indeed for any teacher. To just have a  
11 **"done and turned in"** seems like that is the only goal in my classes right now. **I do not agree and**  
12 **it is not the most effective way of learning**. I have struggled throughout my past field placements  
13 with that. Was are teachers being too predictable and too easy? Just because it is not AP or  
14 Honors **does not mean we have to be that "basic" with them**.  
15 There is almost no "creativity" in my placements. Book Work with a worksheet and  
16 long tedious vocabulary lists with information on the test book, after a basic PowerPoint with line of  
17 art, discussions or deep questioning is the pattern. The students are going to really get a quick  
18 and I hope successful taste of my different method of teacher this next week. Next Week A  
19 Reflection will be over how all that went.  
20 The best moment of this past week? Hmm, I guess the PowerPoint over the Home front  
21 in America during World War One. My **Co-op teacher** did use some **good propaganda material**  
22 but the questions attached to the lesson were **very easy** and **not Big Picture enough**. **The students**  
23 **did not see how** really this method of Control and Influence was vital during the war and at  
24 home.  
25 Most Difficult? Well, I was asked 5 minutes before hand to start a class. **Idiotism**  
26 **teacher had a Meeting with a parent** I over Women's roles during WWI. Wow, **it's** **rough**  
27 **watching someone else's lesson with someone else's PowerPoint**, and **teaching that class in the**  
28 **first time**. I at least feel I did pretty much OK with managing the class. I've tested the. But this  
29 was only initial fun. I mean I was ready for it. **"Be Prepared, Be Ready, and Be Organized"** is  
30 my **Teacher's motto**. I guess he's right about that. **Teachers have to always be ready**. Always.  
31  
— Co-op teacher stuff.  
1. questions of the book.  
very easy & not Big P.P.  
easy  
2. left 5 min before and ...  
— Views of teachers  
— Views of students  
— Analysing for social studies  
— Context of schools  
— Emergent codes  
1. tough to be creative, try  
to be creative in my place  
2. "done and turn" is not the  
most effective way of learning  
3. No: too predictable & too easy  
— 1. like being busy and easy  
work.  
2. they like to chat about  
off-topic things.  
3. have it more and turning  
No "creativity" Book work,  
to do it vocabulary after a  
basic P.P. - few discussions

# Coding with computer software

- If using computer software, you code by tagging and naming selections of text within each data item.
- NVivo: <http://www.qsrinternational.com/nvivo-product>
- ATLAS.ti: <http://atlasti.com>
- Saturate App: <http://www.saturateapp.com>



### 3. Searching for themes

Collating codes into potential themes, gathering all data relevant to each potential theme.

- Analyze codes and consider how different codes may combine to form an overarching theme.
- It may be helpful at this phase to use ***visual representations*** to help you sort the different codes into themes. You might use tables, or mind-maps, or write the name each code.

## 4. Reviewing themes

- Checking if the themes work in relation to the coded extracts and the entire data set, generating a thematic ‘map’ of the analysis.
- **Goal:**
  - to ascertain whether the themes ‘work’ in relation to the data set.
  - to code any additional data within themes that has been missed in earlier coding stages.
- **Note:**
  - Data within themes should cohere together meaningfully, while there should be clear and identifiable distinctions between themes.

## 5. Defining and naming themes

- Identify the 'essence' of what each theme is about
- Ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells, generating clear definitions and names for each theme.
- Identify whether or not a theme contains any sub-themes
- Give names to the themes

## 6. Producing the report

- The final opportunity for analysis.
- Selection of vivid, compelling extract examples, final analysis of selected extracts, relating back of the analysis to the research question and literature, producing a scholarly report of the analysis.
- It is important that the analysis (the write-up of it, including data extracts) provides a concise, coherent, logical, non-repetitive and interesting account of the story the data tell / within and across themes.

# Reliability in qualitative research

- Reliability and validity are fundamental concerns of the qualitative researchers
- Transparency of technique
  - Carefully documenting all their steps so that they can be checked by another researcher
- Reliability checker
  - Organizing an independent assessment of transcripts by additional skilled qualitative researchers and comparing agreement between the raters.
  - Can be done statistically (called inter-rater reliability) or qualitatively by discussing disagreements

## In class activity [20 min]

- Coding a transcript
- Check the website for the activity description

## On deck...

- Please check the website for upcoming deadlines

# Extra slides



# Advantages of thematic analysis

- Flexibility.
- Relatively easy and quick method to learn, and do.
- Accessible to researchers with little or no experience of qualitative research.
- Results are generally accessible to educated general public.
- Can usefully summarize key features of a large body of data, and/or offer a 'thick description' of the data set.
- Can highlight similarities and differences across the data set.
- Can generate unanticipated insights.
- Allows for social as well as psychological interpretations of data.

# Pitfalls to avoid when doing thematic analysis

1. Failure to actually analyze the data
  - Thematic analysis is not just a collection of extracts strung together with little or no analytic narrative.
2. Using of the data collection questions (such as from an interview schedule) as the 'themes' that are reported.
3. A weak or unconvincing analysis
  - where the themes do not appear to work, where there is too much overlap between themes, or where the themes are not internally coherent and consistent.
4. A mismatch between the data and the analytic claims that are made about it.
  - the claims cannot be supported by the data