

18th February 2020

CSCM39/CSDM001: Human Computer Interaction

Dr Siyuan Liu

Department of Computer Science

siyuan.liu@swansea.ac.uk

Office Hour: Thursday 2-4pm

PURPOSE OF



QUANTITATIVE RESEARCH METHODS

Explaining phenomena by collecting numerical data that are analysed using mathematically based methods (in particular statistics), e.g., controlled experiment, surveys.



QUALITATIVE RESEARCH METHODS

Discovering why and how people behave in the way that they do to provide in-depth information about human behaviour, e.g., observations, field studies, focus groups, interviews.

NUMBERS ARE NOT ALWAYS MEANINGFUL

'Er... Good morning, O Deep Thought,' said Loonquawl nervously, 'do you have... er, that is...'

'An answer for you?' interrupted Deep Thought majestically, 'Yes, I have.'

'Though I don't think,' added Deep Thought, 'that you're going to like it.'

'Doesn't matter!' said Phouchg. 'We must know it! Now!'

... 'Alright,' said Deep Thought.

'The answer to the Great Question...'

'Yes!...!'

'Is...' said Deep Thought, and paused.

'Yes...!'

'Is...'

'Yes...!!!!...?'

'Forty-two,' said Deep Thought, with infinite majesty and calm.

...

'Forty-two!' yelled Loonquawl. 'Is that all you've got to show for seven and a half million years' work?'

'I checked it thoroughly,' said the computer, 'and that quite definitely is the answer. I think the problem, to be quite honest with you, is that you've never actually known what the question is.'

Benefits of qualitative methods

- Used in situations where little is known
- Useful for explanations (asking questions “what?”, “how?” and “why?”)
- Provides the base for quantitative research – develop a theory that can be evaluated quantitatively
- Complementary to quantitative research – can use in “Discussion”

METHODS



OPEN-ENDED
QUESTIONNAIRES



INDIVIDUAL
INTERVIEWS



FOCUS
GROUPS



PARTICIPANT
OBSERVATION



DAIRY
STUDIES



NATURALLY
OCCURRING DATA



ROLE-PLAY &
SIMULATION



CASE
STUDIES

Interviews and Focus Groups

Ask the users

- Direct feedback from interested individuals is fundamental to human-computer interaction research
 - What should a new tool do?
 - Does a proposed design do what it should do? If not, what should be changed or revised?
- Surveys can be useful , but are limiting: being broad but not deep
- An alternative approach: deep but not broad
- Two forms
 - Interviews with individuals
 - Focus groups involving multiple users at one time

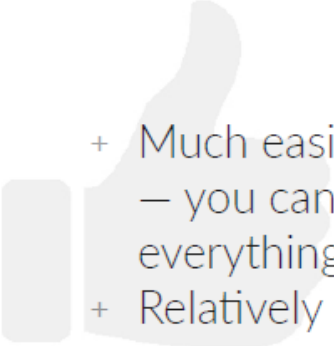
Applications of interviews

- Initial exploration
- Requirements elicitation
- Evaluation and subjective reactions

TYPES OF INTERVIEWS

STRUCTURED

A tightly-defined schedule of questions in a set order

- 
- + Much easier for the interviewer — you can be sure that you get everything covered
 - + Relatively easy to analyse
 - Does not allow the interviewee to follow their train of thought
 - May not allow a full exploration of the topics

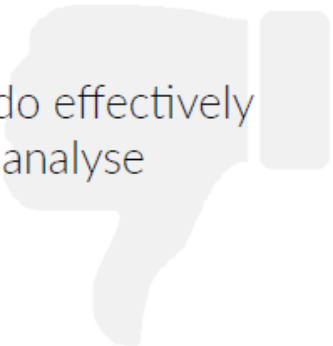
SEMI-STRUCTURED

A general set of questions but you can let interviewee direct the conversation

- + Allows exploration of the issues the interviewee is interested in and knows
- + Provides greater depth of information
- Requires skill to do effectively
- More difficult to analyse

UNSTRUCTURED

A list of topics, start with one, let the conversation flow, if it falters, choose another topic from your list

- 
- + Allows full exploration of the issues the interviewee is interested in and knows
 - + Provides greatest depth of information
 - Requires skill to do effectively
 - More difficult to analyse

How to choose?

- Fully structured interviews are most appropriate when you hope to compare responses across individuals: these comparisons are often most useful for evaluations aimed at understanding user response to designs or systems.
- Unstructured and semistructured interviews can be most appropriate when you are looking to dig deeper, in search of critical comments, design requirements, and other insights: when you are unfamiliar with a problem domain or set of users, when you do not even know which questions to ask.

Interviews vs. focus groups

- Interviews take time
 - Often 1 hour or more
 - Several hours for analyzing notes
- Focus groups
 - More people in less time
 - Not large
 - 8-12 people
 - 5-7 participants for an in-depth conversation
 - Two or more groups will increase your chance of success
 - Generally semistructured or unstructured

Types of questions

- Closed-ended questions: Specific answers
 - Yes/No, multiple choice, true/false, Likert-scale questions, e.g., “On a scale of 1-10, 10 being best, how did you like the web page?”
 - Easy to analyze
 - Discourage elaboration and further comments: “Did you like the design of the home page” or “What do you think about this home page?”
- Open-ended question: Invite elaboration, discussion
 - “What do you think about this home page?”
- Ask users to complete a sentence:
 - “My favorite web browser feature is...”
- Conceptual mapping: Draw pictures or layouts to describe understanding of a situation or problem

Other guidelines

- Simple questions – no jargon
- Avoid compound questions with multiple parts
 - Not "“What were the strengths and weaknesses of the menu layout and the toolbar?”"
 - Ask two separate questions instead.
 - Or four...
- Avoid judgmental phrasing or tone
 - “what do you think of...” rather than “did you like...”

Preparing for interviews

- Pilot test – with colleagues and participants
 - Find questions that are hard to understand
 - Potential length of an interview
- Write an interview guide listing what to do and when
 - Helpful for focus groups
 - Particularly good if multiple researchers are involved
- Backups
 - Batteries for laptop, audio recorder, extra paper, etc.

Recording the responses

- Always have a notebook on hand
- Written notes can be useful for recording interviewee responses and interviewer comments
 - Non-verbal responses, concerns, body language cues
 - Most effective for simple, closed questions
 - Summarize written notes as soon as possible after the interview

Audio and video recordings

- Complete, but at the cost of difficulty of transcription and interpretation
 - Transcription can take many hours
 - Massive amounts of content
- Video is logistically harder, but gets useful information
- Consider audio along with still pictures
- Respect privacy and anonymity

During the interview

- You're the host:
 - Build rapport
 - Be friendly, respectful, nonjudgmental
 - Listen carefully
- Outline
 - Briefly introduce research goals
 - Complete paperwork (informed consent), tell participants if you are recording the session and how.
 - Simple questions first, hard questions later

During the interview

- Be adaptable and flexible
 - If your interview is not fully structured
 - But try to keep things on track
- Explain why you are asking each question
- Define terms, avoid jargon
- Ask for clarification

Challenges of focus groups

- Manage the room. Be prepared to deal with
 - Digressions
 - Arguments
- Give everyone a chance to talk
 - Address them directly
 - “Joan, what do you think about...?”

Promoting discussion

- What if they won't talk?
- Fully-structured – not much to do
- Otherwise
 - Rephrase questions
 - Dig deeper into specifics
- Use physical props, note cards, and other probes to stimulate feedback
- Focus groups – ask for dissenting or concurring feedback

Closing it out

- Ask for any final comments
- Provide more detail about research goals
- Brief summary of findings
- Turn off recording devices
 - Interviewees might make additional useful comments
 - Ask before including these comments in analysis
- Say “thanks!”
- Reflect and summarize notes immediately

Electronically mediated interviews

- Online chat, conference calls, one-to-one phonecalls, emails
- Pros
 - Easy, inexpensive
 - Reach more people with less effort
 - Potentially powerful screen, audio capture
- Cons
 - Lack of face-to-face contact
 - Fewer non-verbal cues
 - Pacing can be harder

Data analysis

- Do it as soon as possible
 - Sit down and summarize notes immediately afterward
- Avoid bias and reliance on preconceived notions
- Fully-structured, closed-ended: tabulate answers
- Open-ended questions require qualitative coding
 - Transcribe audio
 - Written summaries

Reporting results

- Be as specific as possible
 - Not “most respondents”
 - Instead, “7 out of 10 respondents”
- Use quotes or paraphrases from respondents
 - But don't use participant name
 - Use identifiers (Subject 3) or pseudonyms

Diary

What is a diary?

- A diary is a document created by an individual who maintains regular recordings about events in their life, at the time that those events occur
- Diaries are useful for recording information that is fluid and changes over time, such as user mood
- If recall after-the-fact will be hard, inaccurate, or subject to biases, diaries might be a useful data collection method
- Many users already keep informal diaries in social networking, without realizing it

Why use diaries in HCI?

- Diaries fill the gap between observation in naturalistic settings, fixed laboratory settings, and surveys
- Users may have different reactions when being observed, and observers may not always understand what is going on
- If interested in collecting data that is fluid and changes over time (rather than factual data), such as mood, feeling, perception, time , or response, surveys can lead to biased data due to biases in recall

Why use diaries in HCI?

- Diaries are good for recording user-defined incidents
 - When users intended to perform a task, but decided not to do so
 - When users feel that they have learned something new
 - When users decide to contribute to projects actively for the first time
- Diaries are also good at researching situations where users move around and don't stay in one place
 - Mobile phones, GPS devices, hand-held technology

Challenges with diaries

- Users sometimes are not introspective and are not even aware of the specifics of what they are doing and therefore may have trouble recording it in a diary entry
- Users may not follow through and record (via paper or electronic) a sufficient number of entries
- Time recording may still be less accurate for time diaries than for controlled laboratory setting or automated data collection
- Generally harder to recruit users for a diary study than for something less intrusive, like a survey

| Advantages | Disadvantages |
|---|---|
| Good for understanding how individuals utilize technology in nonworkplace, noncontrolled, or on-the-go settings | Participants are sometimes not introspective and not aware of the specifics of what they are doing; they may therefore have trouble recording it in a diary entry |
| Good for understanding the “why” of user interaction with a technology or any technology phenomenon | Participants may not follow through and record a sufficient number of entries |
| More accurate time recording than in a survey | Time recording may be less accurate than in a controlled laboratory setting or automated data collection |
| Good for collecting data that is fluid, and changes over time (such as time, mood, perception, or response) | Generally harder to recruit participants for a diary study than for a less intrusive study, such as a survey |
| The limited gap between an event happening and it being recorded can help limit the impact of individual personality on interpretation of what occurred | Since data is both qualitative and quantitative, data analysis may take a long time |
| Good for collecting user-defined data (e.g., when a user intended to perform an action but did not do so) | Hard to strike a balance between a frequent-enough series of diary entries and infringement on daily activities (user participation may then trail off) |

Participants for a diary study

- Determine in advance who appropriate participants/users are
 - Demographic, education, computer experience, job responsibility
- Try to get a representative group of participants, but it is more important to have users who can provide useful insight
- Potential diarists must not only meet demographic requirements but also possess three qualities
 - Understand the purpose of maintaining the diary
 - Be motivated to keep a regular and accurate record
 - Competence in using the technology that is the subject of the diary and the method used to record the diary

Types of diaries

- Feedback diary: the data recorded in the diary is itself the purpose of the research
 - Users make entries when a certain event or threshold occurs, or on a stated time basis
- Elicitation diary: the users record only basic information about important events occurring in their day
 - These data points are used as prompts for expansion at a later time

Feedback diary

- How often should an entry be made?
 - What event, time, or threshold triggers the need for the user to make a diary entry?
 - Make a diary recording when users feel frustrated with an interface
 - Make a diary recording at a set time every day, e.g., 9pm.
- Feedback diaries can be structured or unstructured
 - Likert scales, multiple-choice questions, closed-ended questions, checkboxes, and time recording
 - A personal reflection
 - How are you feeling right now about our computer?
 - How do you think an activity could be improved?

Elicitation diary

- Data points recorded in an elicitation diary are quick and simple
 - Pictures
 - Short snippets of text
 - Short audio clips
 - Combination of the above
- Then, later (in an interview, on a web site, or other format), users expand on their entries, describing in more detail

Data collection

- How will the diaries be recorded?
 - Paper?
 - Electronic?
 - Voice recording?
 - Pictures?
 - Smart phones? Tablets?
- If a specific technology is being studied in a diary study, you may want to use a different, common technology for the diary recordings
 - Use whatever is most natural for the participants

When to record an entry?

- Participants should be given information about
 - What activities are of interest?
 - What events, incidents, or feelings should result in a diary entry?
 - How often should diary entries be made?
 - How detailed should the entries be?
- Make sure NOT to pay participants based on the number of diary entries

When to record an entry?

- Two weeks are often an appropriate length of time for a diary study
- If diary reports are turned in during the study period, researchers should monitor the incoming reports, check on who is not reporting diary entries, or if the entries are not providing useful data
- Reminders and feedback can be sent during the period of the diary study

Analysis of diaries

- Transfer all records to an easy-to-analyze electronic format
- Do statistical analysis on quantitative data
- Prepare and examine qualitative data, potentially do a content analysis
- Follow-up interviews with participants who keep an elicitation diary
- Even if the diary is not an elicitation diary, you can contact the participants to ask for clarification of diary entries

Case Study

Case study

- In-depth study of a specific instance (or a small number of instances) within a specific real-life context
- Goals
 - Exploration: understanding novel problems or situations, often with the hopes of informing new designs
 - Explanation: developing models that can be used to understand a context of technology use
 - Description: documenting a system, a context of technology use, or the process that led to a proposed design
 - Demonstration: showing how a new tool was successfully used

An example: Observing Sara

- The researchers were interested in understanding how a blind person might use a variety of assistive technologies to accomplish tasks and to recover from task failures using workarounds.
- By examining the use of a range of technologies in the user's (Sara's) home, the investigators expected to see common types of failures and workaround strategies.
- They also expected that the choice of implementing features in hardware or software might influence user interactions, including failures and responses to those failures.

An example: Observing Sara

- They used a series of semistructured interviews to collect the observations that form the basis of the case study
- In a series of 6, 2-hour sessions in her home, Sara demonstrated how she used technologies such as tactile wristwatches and screen readers; discussed early memories of using various objects and her reactions to them; and imagined improved designs for various objects or tasks. Notes, audio recordings, interviewer reactions, and photographs from these sessions provided the raw data for subsequent analysis. Insights and theories based on early observations were shared with the subject for validation and clarification.
- Twelve tasks were recorded

Types of case study

- Intrinsic or instrumental
 - The intrinsic studies describe cases that are of interest to a particular situation
 - The instrumental case studies work towards developing a broader understanding
- Single case or multiple cases
 - Multiple cases are possible
 - Similarities between cases might support generalization
 - Only limited generalization possible on 2-3 cases
- Embed or holistic
 - The inclusion of multiple units of analysis within a single case is referred to as an embedded case study
 - Holistic studies that address only one unit in each case.

Components of a case study design

- Questions: the goals of your study
- Hypotheses or propositions: statement of what you expect to find
- Units of analysis: the granularity of your study
- A data analysis plan

Data sources

- Documentation and archival records
 - Understand past and current practices
- Interviews
 - Understand perceptions, concerns, needs, and other user reactions
- Observation
 - Understand what people do in circumstances of interest
- Physical artefacts
 - How people bridge the gap between computer work and the rest of their lives, e.g., paper notes stuck to the edge of computer monitors

Observation

Observation

- Observational data is also very useful in overcoming discrepancies between what people say and what they actually do and might help you uncover behaviour of which the participants themselves may not be aware.
- Basic types
 - Quick and dirty observation
 - Observation in the lab
 - Observation in field studies

How to observe?



NOTES
&
PHOTO CAMERA



CHECKLIST
OF
BEHAVIOURS



AUDIO RECORDING
&
PHOTO CAMERA



VIDEO

Naturally Occurring Data

YOU DON'T ALWAYS
HAVE TO COLLECT
THE DATA YOURSELF...

Naturally occurring data

- Some written sources of data that might help you answer your research question include:
 - Political speeches
 - Clinical records
 - Policy reports
 - Media debate programmes
 - Online newspaper readers' comments
 - Blogs
 - Discussion forums
 - Web pages
 - Social media sites

Analysing Qualitative Data

Goals and stages of qualitative analysis

- Goal: turn unstructured data into a detailed description about the important aspects of the situation or problem
- Stages:
 - Identify components of the substance
 - Identify major themes and ideas that describe the context, activities, and other perspectives that define the problem of interest
 - E.g., challenges faced by senior citizens, the interaction behavior in a specific context
 - Study properties and dimensions of each component
 - Understand and make inference about the substance

Content analysis

- A more specific view: a systematic, replicable technique for compressing many words of text into fewer content categories based on explicit rules of coding
- A broader view: any technique for making inferences by objectively and systematically identifying specified characteristics of messages
- In-depth analysis that searches for theoretical interpretations that may generate new knowledge.

Content categories

- Media content can be any material in printed publications, or any other types of recording.
- Audience content is feedback directly or indirectly collected from an audience group

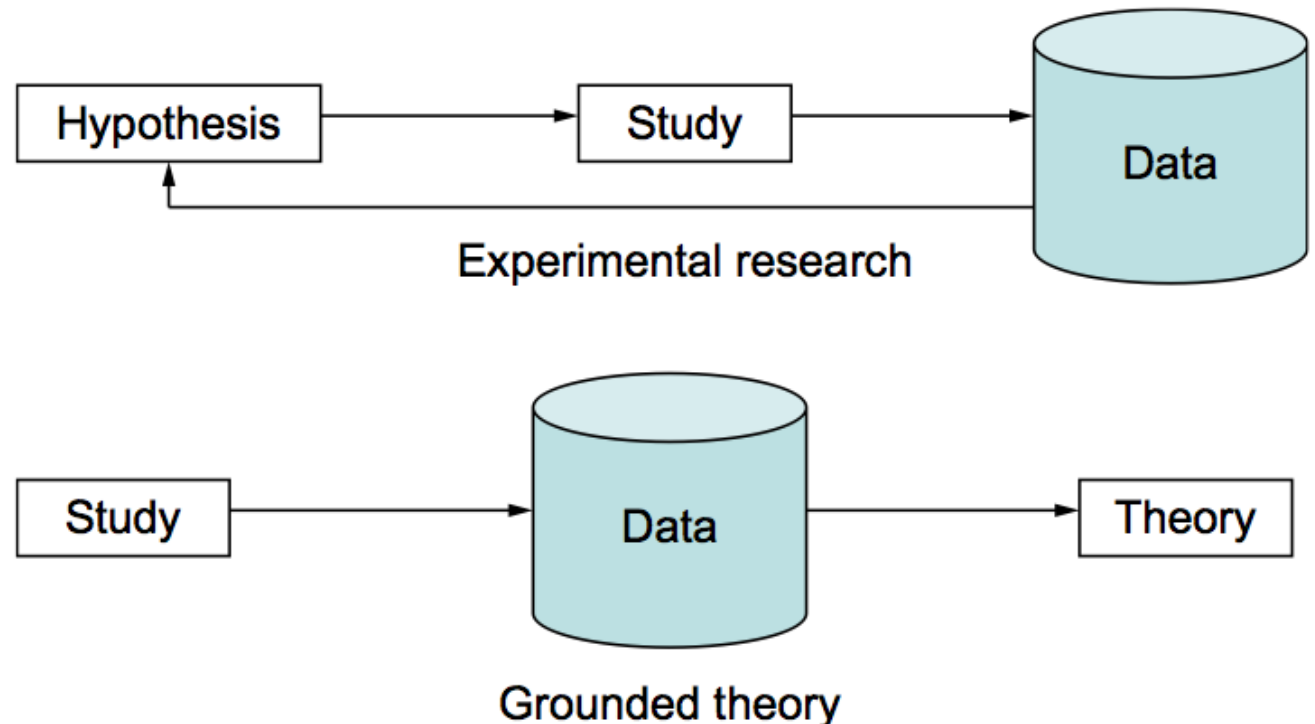
| Category | Subcategory | Examples |
|------------------|--------------|---|
| Media content | Publications | Books, journals, newspapers, brochures |
| | Broadcasting | TV programs, radio programs |
| | Websites | News, web portals, organizational websites, blogs |
| | Others | Films, music, photos |
| Audience content | Text | Notes from interviews, focus groups, or observations or diaries or surveys, text posts on social media |
| | Multimedia | Video- or audio-recording of interviews, focus groups, observations, or user studies, pictures or video recordings posted on social media |

Analyzing text content

- Coding: a process that involves assigning categories and descriptors to blocks of text
- Emergent coding
 - Conducted without any theory or model
 - Identify any interesting concepts or ideas in data
- A priori coding
 - Involves the use of established theory or hypothesis to identify categories

Grounded theory and emergent coding

- If you are working on a new topic that has very limited literature to build on, you may not be able to find established theories that allow you to develop the coding categories in advance. In this case, the emergent coding approach, based on the notion of grounded theory, is appropriate.
- An inductive research method



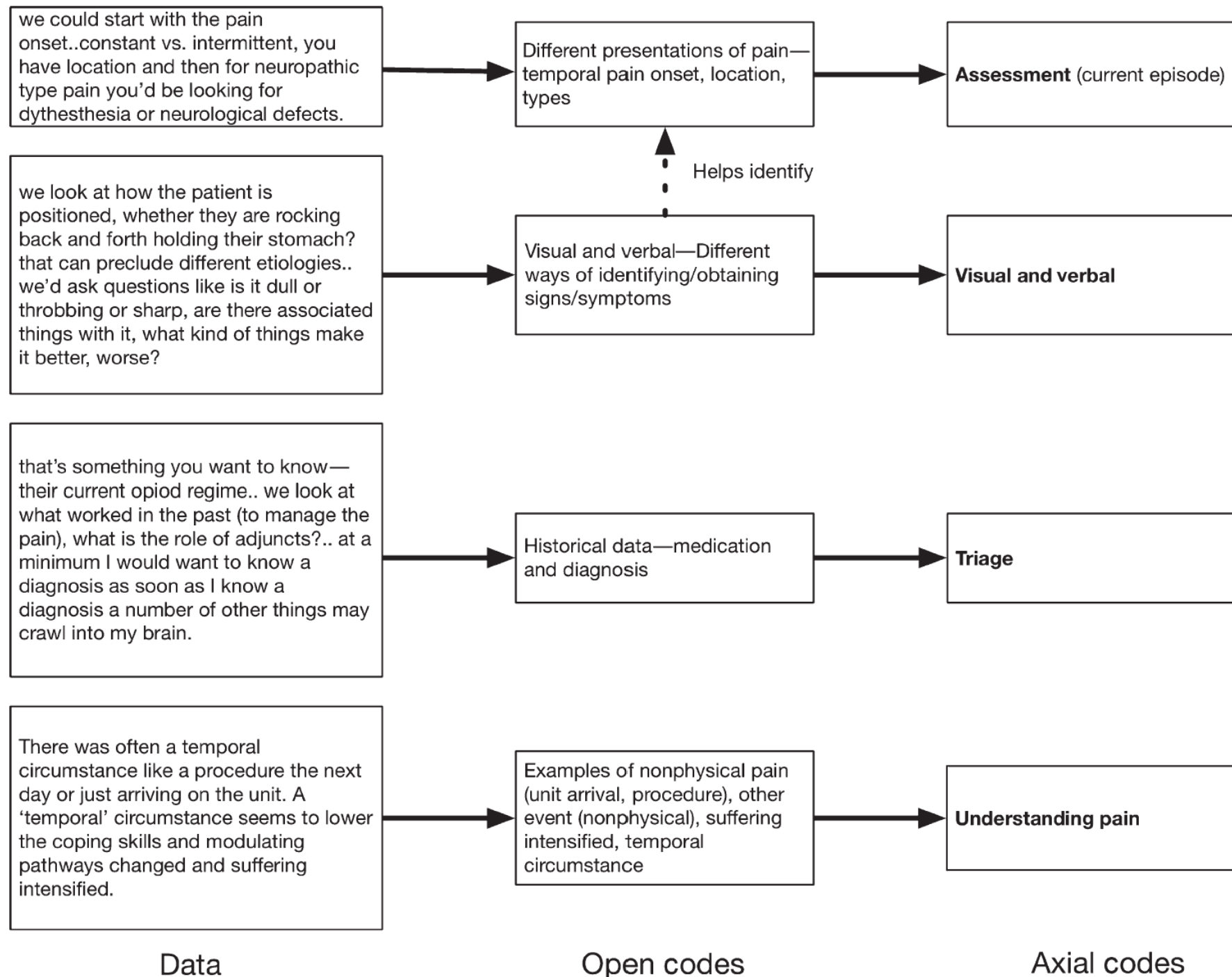
Procedures of grounded theory

- Open coding:
 - Analyze the text and identify an interesting phenomena in the data. Each unique phenomenon is given a distinctive name or code.
- Development of concepts
- Grouping concepts into categories
- Formation of a theory

My son just sits there and sobs when the computer does not do what he wants. He becomes irritated and keeps pushing the Enter button when the web page loads slowly.

Procedures of grounded theory

- Open coding:
 - Analyze the text and identify an interesting phenomena in the data. Each unique phenomenon is given a distinctive name or code.
- Development of concepts
 - Collections of codes that describe similar contents are grouped together to form higher level “concepts”, which can then be grouped to form “categories”
 - Key step: axial coding
- Grouping concepts into categories
- Formation of a theory
 - Creating inferential and predictive statements about the phenomena recorded in the data.



Grounded theory

- Advantages

- A systematic approach to analyzing qualitative, mostly text-based, data
- Generating theory out of qualitative data that can be backed up by ample evidence of the coding
- Interplay between data collection and analysis

- Disadvantages

- Researcher can be overwhelmed by the details of the data
- The theory generated is hard to evaluate
- Findings may be subject to bias

A priori coding and theoretical frameworks

- Benefits of theoretical frameworks
 - Design stage:
 - Frame research questions
 - Decide research approach (i.e., survey, interview, focus group, etc.)
 - Identify concepts and questions
 - Analysis stage
 - Identify categories or coding items
 - Explain the finding

Coding the text

- When the data set is not large, which is typically true for interviews, focus groups, or observations, it is recommended to read the text from beginning to end before starting to do any coding: immerse you into the life and experience of the participants and get a general, unbiased idea of the data set before focusing on any specific aspects.
- Main steps:
 - Look for key items
 - Ask questions about the data
 - Make comparisons at various levels

Look for key items

- Specific types of “statements” are more likely to carry valuable information.
A particle list of such statements

| Statement | Examples |
|--------------|---|
| Objectives | Use computers for educational purposes |
| Actions | Enter a password, chat online |
| Outcomes | Success or failure, whether the objective is achieved |
| Consequences | Files unintentionally deleted, a specific application abandoned |
| Causes | Limited memory, dated equipment |
| Contexts | User is computer savvy, user works with classified information |
| Strategies | Avoid specific tasks, multimodal interaction |

Ask questions about the data

- Sensitizing questions: to better understand the meaning of the data
 - What is happening?
 - What did the user click?
 - How did the user reach the specific web page?
- Theoretical questions: help make connections between concepts and categories
 - What is the relationship between the two factors?
 - How does the interaction change over time?

Make comparisons

- Between different coding category
 - E.g., if you are investigating the difficulties that older people experience when using computers, you can compare the frequency with which each capability is reported.
- Between different participant group
 - E.g., is the diversity related to age, educational background, or community and family support?
- Between existing data and previous literature
 - Do your findings align with the existing literature or is it contradictory?

Ensure high quality analysis

- **Validity:** in terms of qualitative research, validity means that we use well-established and well-documented procedures to increase the accuracy of findings.
- **Reliability:** the consistency of results. If different researcher working on a common data set come to similar conclusions, those conclusions are said to be reliable.

Validity

- Constructing a multi-faceted argument in favor of your interpretation of the data
 - Constructing a database: all the materials that you collect and create during the course of the study, including notes, documents, photos, tables, procedures and products of your analysis, including summaries, explanations, and tabular presentations
 - Data source triangulation: the use of multiple data sources to support an interpretation
 - Interpretation should account for as much as possible of the data
 - Alternative interpretations may also help

Reliability

- The ultimate goal of reliability control is to ensure that different people code the same text in the same way
- Reliability checks span two dimensions:
 - Stability (*intra-coder reliability*)
 - Examines whether the same coder rates the data in the same way throughout the coding process
 - If the coder is asked to code the same data multiple times, is the coding consistent time after time?
 - Reproducibility (inter-coder reliability)
 - Examines whether different coders code the same data in a consistent way
 - If two or more coders are asked to code the same data, is there coding consistent?

Reliability measures

- Percentage of agreement among coders:

$$\% \text{agreement} = \frac{\text{the number of cases coded the same way by multiple coders}}{\text{the total number of cases}}$$

- Several coders would agree with each other for a certain percentage of cases even when they just code the data by chance.
- Cohen's Kappa:

$$K = \frac{P_a - P_c}{1 - P_c}$$

where P_a represents the percentage of cases on which the coders agree and P_c represents the percentage of agreed cases when the data is coded by chance.

Reliability check

- A survey of senior citizens, asking them to describe the primary causes of the difficulties that they encounter when using computers.
- We identify three major categories of causes: difficulties due to physical capabilities, difficulties due to cognitive capabilities, and difficulties due to perceptual capabilities.
- Two coders code the data independently.

| | | Coder 2 | | | |
|---------|----------------|-------------|-------------|-------------|----------------|
| | | Physical | Cognitive | Perceptual | Marginal total |
| Coder 1 | Physical | 0.26 (0.14) | 0.07 (0.08) | 0.04 (0.15) | 0.37 |
| | Cognitive | 0.04 (0.07) | 0.12 (0.04) | 0.01 (0.07) | 0.17 |
| | Perceptual | 0.09 (0.18) | 0.02 (0.10) | 0.35 (0.18) | 0.46 |
| | Marginal total | 0.39 | 0.21 | 0.40 | 1.00 |

Reliability check

| | | Coder 2 | | | |
|---------|----------------|-------------|-------------|-------------|----------------|
| | | Physical | Cognitive | Perceptual | Marginal total |
| Coder 1 | Physical | 0.26 (0.14) | 0.07 (0.08) | 0.04 (0.15) | 0.37 |
| | Cognitive | 0.04 (0.07) | 0.12 (0.04) | 0.01 (0.07) | 0.17 |
| | Perceptual | 0.09 (0.18) | 0.02 (0.10) | 0.35 (0.18) | 0.46 |
| | Marginal total | 0.39 | 0.21 | 0.40 | 1.00 |

$$P_a = 0.26 + 0.12 + 0.35 = 0.73$$

$$P_c = 0.37 \times 0.39 + 0.17 \times 0.21 + 0.46 \times 0.40 = 0.36$$

$$K = \frac{0.73 - 0.36}{1 - 0.36} = 0.58$$

Reliability check

| Interpretation | Kappa range |
|--------------------------|----------------------|
| Poor or slight agreement | $K \leq 0.20$ |
| Fair agreement | $0.20 < K \leq 0.40$ |
| Moderate agreement | $0.40 < K \leq 0.60$ |
| Satisfactory agreement | $0.60 < K \leq 0.80$ |
| Near-perfect agreement | $K > 0.80$ |

When the value of Kappa is below 0.60, the reliability of the analysis is questionable.

Analyzing multimedia content

- The supporting techniques and methods are less mature compared to text data analysis
- Approaches:
 - Manual analysis:
 - Highly labor intensive and time consuming
 - More accurate
 - Completely automated analysis:
 - Faster, less amount of work
 - Highly inaccurate
 - Partially automated approach
 - Combines the advantages of the manual process and the completely automated process

Summary

- Familiarising yourself with your data
- Generating codes
- Reviewing
- Checking validity and reliability
- Producing the report