# CMPE 235: User Evaluation of Technology

**Qualitative Research, Content Analysis and Grounded Theory**

Course contact: srikur@soe.ucsc.edu

## Reactive vs. Nonreactive Research

- **Nonreactive:** Subjects are unaware they are being studied
  - Unobtrusive measures
  - Often use naturalistic settings
  - 1. **Physical traces**
    - Erosion measures – selective wear
    - Accretion measures – deposits of something left behind.
    - Examples: floor wear analysis, restroom graffiti
  - 2. **Archival studies** – study existing data
    - Written and audio/visual records
  - 3. **Observation**
    - Mall visitor behavior

## Nonreactive advantages

- No subject confounds
- Assess actual behavior rather than self-report
- Safety
- Reliability
- Inexpensive
- Good for longitudinal data

## Nonreactive disadvantages

- No control
- Often don’t know anything about the subjects
- Sample may not be representative
- Secondary information may have bias
- Intervening variables – e.g., floor erosion
- Need for triangulation – looking at material from several different perspectives gives a more accurate view of it.

## Qualitative Analysis

- Data don’t exist unless they are written
- Transcription-taking oral conversation to words
- Organizing data for analysis
- Computer programs to assist: NVIVO, Atlas, N*DIST

## Steps

- Read, read, read
- Data reduction (coding and counting) → *emic* (ideas expressed by respondents) and *etic* (data expressed in evaluator’s language)
- Data refinement (constant comparison)
- Data display/notes (memos)
- Data selection (focusing on certain aspects)
- Writing (linking to original and rival hypotheses)

## Content analysis

- One way of analyzing qualitative data
- A technique used to study written material by breaking it into meaningful units, using carefully applied rules.
- Use objective and systematic coding to produce a quantitative description of the observed material.
  - Can analyze common myths
  - e.g., women are portrayed as inferior to men in the cowboy movies.
  - e.g., graffiti in toilets are heavy on pornographic comments
  - e.g., spam mails promise money

## Content analysis seeks to avoid confirmation bias

- the tendency to look for information that confirms our beliefs and ignore information that disconfirms our belief

## Can be used to quantify concepts

## Can also be used in a qualitative way.

## What can be studied

- Any written material
- Audio/visual information

## Useful for 3 types of research

- Problems involving a large volume of test
- Research from afar or in the past
- Revealing themes difficult to see with casual observation.
Steps in content analysis

1. Define problem
2. Select the media that will be used
3. Derive coding categories
4. Sampling strategy – (every 10th page, every other sentence?)
5. Choose the coders: Humans vs. computer
   - Human is useful for coding complex concepts
   - Computer removes subjectivity
   - Human costs more time and money
   - Multiple coders and interrater reliability are a must
6. Code the material
   - In vivo codes vs. conceptual constructs
   - Established vs. your own codes
7. Analyze the data

Getting started

What gets counted?
- # certain words, # pictures, senders/receivers
What is important for understanding themes?
- Explicit themes
- Number of times mentioned
- Amount of space dedicated
What is the coding unit of analysis?
- Word vs. paragraph vs. themes
What to analyze?
- Frequency
- Direction: Positive vs. negative; happy vs. sad
- Intensity: Strength of message, minor vs. major issues
- Space: Picture size, amount of time spent, etc.

Latent vs. manifest content

Manifest – overt, visible material
- How many times a word appears
- How many times someone is mentioned
- Highly reliable coding
- No judgment
Latent content – symbolic content; semantic analysis
- Ex. Level of violence
- Requires judgment
- Depends on coders prior knowledge, expectations, etc.
- Often required – writers portray meaning indirectly
- Lower reliability, increases with training
- Allows for more flexibility

Coding approaches

Common classes
- used by virtually anyone in society, e.g. age, gender, mother, father, etc
- essential in assessing whether certain demographic characteristics are related to patterns that arise from other coding
Special classes
- colloquial categories
- includes jargon of various professions, e.g. petty crime vs. serious crime
Theoretical classes
- those that emerge in the course of analyzing the data
category labels generally borrowed from special classes
- their substance is grounded in the data
- not immediately knowable until observers spend considerable time with the content

Inductive vs. Deductive Category

Deductive – reasoning from the general to the specific
Forming categories to score based on theoretical ideas.

Theoretical codes
- Set up categories based on theory or framework
- Advantages and disadvantages?
Inductive category formation
- Reason from the specific to the general
- Come up with categories from data
- Advantages and disadvantages?
Can obtain categories by using grounded theory
Trade-offs

► In vivo codes vs. conceptual constructs
  ▪ Actual words vs. terms constructed by professionals (obsessive workaholics)
► Established vs. your own codes
  ▪ Individuality in the data vs. being accused of circular reasoning
  ▪ Avoiding accusation: divide the data set in half, develop the code on one half, apply it to the other half

FACT

THEORY

induction

deduction

(read text)

revision

(revision)

(read text again)

TED 0

Grounded Theory

► Sociologists Glaser and Strauss’ study of dying in a California hospital.
► Found no real theory to test against this subject.
► Developed the methodology to give them a method of developing a social theory of dying in a hospital using only existing data gathering methods.
► A "loose", less structured method of developing and testing theory simultaneously.
► A system of describing society through abstract notions, rather like a play describes a fictional/documentary occurrence.
► A creative method of breaking away from highly structured analyses.

Distribution of Word-counts in <title>

► Spam more likely in pages with more words in title
Core philosophical principles of Grounded Theory

► Data is data. Qualitative is as good as quantitative.
► Theories are man-made and evolve.
► All data and existing theories are equal. Even supporting studies in refereed journals have equality with data freshly collected in the field.
► All researchers’ experiences are valuable.
► Useful when
  ▪ In an area where there is little or no theory in existence.
  ▪ You disagree with existing theories.
  ▪ You may not want to test existing hypothesis.
  ▪ You may want to mix qualitative and quantitative data.
  ▪ You may want to collect a broad range of data beyond more structured methods.

Forms of data and coding

► Usually involves qualitative data collection, particularly interviewing or observations
  ▪ Self reviews and field diaries are also used as data.
► Statistical information is usually regarded as secondary source data.
► Constant comparative method
  ▪ Observations are compared with one another and with the evolving inductive theory
► Concept mapping: prefers graphical depiction
► Codes
  ▪ Open coding. Identifies concepts and actors
  ▪ Axial coding. Connecting the concepts and actors (plot)
  ▪ Selective coding. Find the stories behind the connections between actors and their concepts (storyline)

Open coding:

SHAZIA: Yeah I'd rather have some, a few people that mean a lot to me than hundreds of people that I know so little about, that doesn't, its, its not, you can't even call it a friendship it's just like an acquaintance [INT: yeah], but I don't, I'm not really bothered about people that I'm just acquainted with and like just fellow students and stuff like that, I kind of push those kind of, cause I don't like people knowing exactly what I'm feeling and thinking all of the time, so I think if I've got these few people around me that are close to me I can confide with them and that's it, I don't have to, yeah I don't like other people knowing too much about me [INT: OK, OK], I don't know if it's because I don't really like them or because I keep them away for a reason, but yeah, so only a few people close to me.

Quality not quantity in friends
How much you know about people
Acquaintances vs. friends
Not bothered with acquaintances
Don't want people to know my thoughts
A few close confidants

Selective Coding

Selective Coding

Fig. 1. Concept map depicting a grounded theory of the flow experiences of Web users.
Content analysis: an example


1. Identify problem
   - Is American and Canadian television different at portraying crime?
2. Select media
   - Television news programs
3. Derive coding categories (manifest & latent contents)
4. Sampling strategies: equal 100 30-min news in 4 areas (Detroit, Toledo, Toronto and Kitchener)

Manifest Content

- Type of crime
  - Which crime was being reported on
  - Kept 28 categories of crime, not necessarily mutually exclusive
- Local or national story
  - Origin of study
- Length of story
  - Used stopwatch to measure exact time spent reporting each study
- Stage of crime: Pre-arrest, Arrest, Court, Disposition
- Live footage?
- Firearm reported?
- Was it the lead story?

Latent Content

- Reporting of motive
  - Implied or confirmed
  - Ex. Drug-related; gang-related
- Emotive presentation: 3 categories
  1. Presentation of fear:
     - Words were explicitly stated about fear, e.g., “be advised”; “on the run”
  2. Presentation of outrage or sympathy
     - Explicit statements made by reporters or interviewees
     - E.g., “tragedy”; “devastated”; “savage”; “horrifying”
  3. Sensationalism
     - Involving famous people
     - Comical stories
     - Dramatic arrests
     - Vivid descriptions – “bizarre”

Secondary analysis

- Conducting research on secondary sources of material
  - Primary sources – empirically arrived. You gather the information yourself
  - Secondary sources – analysis of someone else’s data
- Advantages
  - Efficiency – data has already been collected, cheap
  - Replication
  - Research question requires a large scale project
  - No time loss for entering and reformatting data
  - Ethical obligation to fully utilize data
- Disadvantages
  - Loss of context, especially if it’s not your data
  - Methodological blunder can’t be fixed