



# Content Analysis

Renee Meyers  
UWS Leadership Site  
February 2010

## Overview

- What is content analysis?
- Steps in content analysis
  - Initial preparation of data
  - Constructing a coding scheme
  - Preparing the data units
  - Coding the data
  - Analysis of the data

## Content Analysis

- Content analysis is a research technique (or method of inquiry) for systematic and replicable analysis of the content of communication, and for making inferences from that data to their context

## Initial Preparation of Data

## Initial preparation of data

- If data is “spoken,” (interviews, focus groups, videotaped groups, etc), it is best to have the data transcribed
  - Can code from spoken text, but much more difficult (typically)

## Developing the Coding Scheme

## Developing the Coding Scheme

- Coding scheme generates from a theory or past research findings (See handout 1, B)
- Use theory and research to develop a set of relevant categories

## Example: Coding Scheme Development

- Question: How do students argue their positions in group quiz discussions? How do they support their quiz answer choices?
  - Theory and research on argument
  - Research on classroom groups
  - Research on group quiz discussions
- See Handout 1B

## Coding Scheme Development

- Begin developing categories based on theory and research findings
- Test category scheme against data that is similar to, but not included in, final data set
- Revise coding scheme based on testing
  - Process is best done by two or more people to discuss similarities and differences
- Continue process until coding scheme is both exhaustive and exclusive

## Sample Category Scheme: Handout #2

### Coding Scheme for Evidence in Group Quiz Discussions

#### I. No Evidence Categories

1. Provide **claim without evidence** (I put that too, I put true, I put D, nonverbal agreement, Yeah;); a claim is verbally or nonverbally stated, but **no evidence** is provided at all.
2. Provides **neither claim nor evidence**; **does not speak** in regards to quiz question.
3. Indicate **lack of own knowledge or expertise** as source. Examples include (don't know, don't remember reading it, guessed, used process of elimination, don't care, don't remember from earlier exam, didn't read all the answers to the question, asks questions of other participants that show doesn't know for sure (but provides no additional evidence or information through these questions), etc.

#### II. Evidence Provided Categories

##### 4. Expert-Based Evidence—Use sources other than self

- Use/Cite **text authority** (It said in the book, I remember from the book, I don't remember reading that in the book, The text said that . . .etc.)
- Use/Cite **teacher's** authority/expertise (She said this in class, I asked her, etc.)
- Use/Cite **class discussion** as authority (We talked about this in class, When we did this in class, Remember when we did that activity in class, etc.)
- Use/Cite **another question on the exam** as evidence
- Use/Cite **class notes** as authority (I wrote it in my notes, I remember from my notes, etc.)

##### 5. Speaker-Based Evidence—Use "self" as source

- Use "own authority" (I know that is right, I remember that, I know this is true, I am sure I am right)
- Use "own sources" such as an example or illustration or definition or clarification to clarify/explain the concepts/theories/wording in test question, etc.

#### III. Other types of evidence—types of evidence NOT contained in the codes above.

## Determining the Coding Units

## Determining the Coding Units

- Need to determine your "unit of analysis"
  - WHAT will you define as your coding units?
  - You want sensible, but rich, units
  - Units need to fit research question, object of research interest, coding scheme (See handout 1A)
    - Words
    - Thought units (subject-verb; complete thought)
    - Sentences
    - Paragraphs
    - Simple turn
    - Complete turns-at-talk
    - Partnered turns-at-talk
    - Whole discussions
  - Your choice of "unit" is based on your research question, your coding scheme, what you know about your data

## Examples of Units

- Want to understand emotion in student discussions
  - Might choose turn-at-talk (see handout 1A)
- Want to study argument in decision-making discussion
  - Might choose thought-unit (because more than one argument can occur in larger units)
    - Handout 3
- Want to study conflict in online discussion
  - Might choose whole discussion, or partnered turns-at-talk

## Unitizing the data

- Most written data must be put into units
  - Develop rules for units (handout #3)
  - Train two (or more) unitizers to identify units
  - Practice until unitizers achieve reliability
  - When reliable, then each unitizer breaks written or spoken data into units independently
  - When complete, overall reliability is computed

## Example: Simple Turn Units

### S Text

C: I put D, because it's definitely not C.  
A: No, I put A.  
B: I put A.  
A: It's not, it's not D, sweetie.  
C: It isn't?  
D: It isn't?  
A: No, I don't believe so.  
C: Because accentuating, what does that mean? Does that mean to show more or to reveal?  
B: Yeah  
D: Yeah, to emphasize . . .  
C: To reveal the differences among groups, is that good?  
A: They [the authors of the book] were talking about how diversity in groups increases group, um, group cohesiveness and all that stuff.  
B: Right, and group creativity.  
D: I think you're right, A is a better answer.

## Example: Thought Units

### S Text

**Ca:** I put D/  
**Cb:** because it's definitely not C.  
A: No, I put A.  
B: I put A.  
A: It's not, it's not D, sweetie.  
C: It isn't?  
D: It isn't?  
A: No, I don't believe so.  
**Ca:** Because accentuating, what does that mean?/  
**Cb:** Does that mean to show more/  
**Cc:** or to reveal?  
B: Yeah  
D: Yeah, to emphasize . . .  
**Ca:** To reveal the differences among groups./  
**Cb:** is that good?  
A: They [the authors of the book] were talking about how diversity in groups increases group, um, group cohesiveness and all that stuff.  
**Ba:** Right/  
**Bb:** and group creativity.  
**Da:** I think you're right./  
**Db:** A is a better answer.

## Coding the Units into Categories

## Coding Process

- Two (or more) coders unfamiliar with research question. Train them on coding scheme
  - Read through relevant past research and coding scheme to familiarize with categories
  - Develop coding rules, Handout 4
  - Practice coding data similar to, but not included in, final data set (Handout 1, C & D)
    - Check codes with each other (and trainer)
    - Discuss differences
    - Develop coding rules, routines, definitions (handout #5)
    - Practice more

## Coding the Units into Category Scheme

- Coders must be reliable . . . must apply the categories similarly
  - 80% reliability is considered standard prior to final coding
- Must practice coding, refine scheme, practice coding, refine scheme, etc. until reach standard of reliability
- Once coders are reliable, they code the “real” data independently
- Final reliability is checked between coders (Handout 1, E)

## Sample Coded Data (Turn Units)

S	L#	C1	C2	Text
C:	1.0	II	5	I put D, because it's definitely not C.
A:	2.0	I	1	No, I put A.
B:	3.0	I	1	I put A.
A:	4.0	I	1	It's not, it's not D, sweetie.
C:	5.0	I	1	It isn't?
D:	6.0	I	1	It isn't?
A:	7.0	I	1	No, I don't believe so.
C:	8.0	II	5.2	Because accentuating, what does that mean? Does that mean to show more or to reveal?
B:	9.0	I	1	Yeah
D:	10.0	II	5.2	Yeah, to emphasize . . .
C:	11.0	II	5.2	To reveal the differences among groups, is that good?
A:	12.0	II	4.1	They [the authors of the book] were talking about how diversity in groups increases group, um, group cohesiveness and all that stuff.
B:	13.0	II	4.1	Right, and group creativity.
D:	14.0	I	1	I think you're right, A is a better answer.

## Statistical Analysis of Coded Data

## Statistical Analysis

- Codes are entered in statistical package (SPSSX, SAS, Excel, etc)
- Statistics depend on research question
  - Frequencies
  - Differences in categories (chi-square)
    - See handout #5

## Summary

- What is content analysis?
- Steps in content analysis
  - Preparing the data
  - Constructing a coding scheme
  - Preparing the units
  - Coding the data
  - Analysis of the data