Coding as an initial step of qualitative data analysis

- Texts
  - Linking raw data to the research question
    - Consolidating raw data
      - Identifying patterns in the data
        - Integrating patterns
          - Answer to the research question: theoretical explanation

Research question and prior theory
Coding

As Part of a Grounded Theory approach


Glaser 1978, 1992

As an independent analytical method

e.g.
- Miles & Huberman 1994,
- Patton 1990

Strauss/Corbin 1990

Different approaches to theory
Identification and Processing of relevant information

by indexing raw data
**Coding**

Procedure (initial step):

1. We read a text sequence (e.g. a line, a paragraph)

2. We interpret it

3. We decide which code to assign (indexing a paragraph, line, word ..)
What is a code?

Codes are categories (keywords, phrases, mnemonics, numbers) which are assigned to a text segment.

Where do the codes come from?

- From the empirical data
- From theoretical considerations

Battle field of Grounded Theorists
Where do the codes come from?

- From the empirical data
- From theoretical considerations

Why is this an important question?

- **principle of Openness**
  The empirical research process must be open for unexpected information

- **principle of theory-guidedness**
  We must proceed from the existing theoretical knowledge; only this way we can contribute to it
Where do the codes come from?

- From the empirical data
- From theoretical considerations

This is totally forbidden! (Glaser/Strauss 1967)

Open Coding, then Theoretical coding = codes arrived from epistemology, general theory are allowed (Glaser 1978)

Examples: limit, extent, goal, social norms

Open Coding, then Axial coding = relating of categories by using a general model of action (Strauss/Corbin 1990)

Axial coding is bad!, it forces categories on the data (Glaser 1992)
Where do the codes come from?

From the empirical data

From the theoretical considerations

Axial Coding

- Intervening conditions
- Causal conditions
- Action/interaction
- Phenomena
- Context
- Consequences
Where do the codes come from?

From the empirical data

From theoretical considerations

Start with theory!
(Miles & Huberman 1994)

Our recommendation:

- Start with theory
- Use middle range theory

<table>
<thead>
<tr>
<th>External Context</th>
<th>EC (PRE) (DUR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC: Demographics</td>
<td>EC-DEM</td>
</tr>
<tr>
<td>In county, school personnel</td>
<td>ECCO-DEM</td>
</tr>
<tr>
<td>Out county, nonschool personnel</td>
<td>ECEXT-DEM</td>
</tr>
<tr>
<td>EC: Endorsement</td>
<td>EC-End</td>
</tr>
<tr>
<td>In county, school personnel</td>
<td>ECCO-END</td>
</tr>
<tr>
<td>Out county, nonschool personnel</td>
<td>ECEXT-END</td>
</tr>
<tr>
<td>EC: Climate</td>
<td>EC-CLIM</td>
</tr>
<tr>
<td>In county, school personnel</td>
<td>ECCO-CLIM</td>
</tr>
<tr>
<td>Out county, nonschool personnel</td>
<td>ECEXT-CLIM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Internal Context</th>
<th>IC (PRE) (DUR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IC: Characteristics</td>
<td>IC-CHAR</td>
</tr>
<tr>
<td>IC: Norms and Authority</td>
<td>IC-NORM</td>
</tr>
<tr>
<td>IC: Innovation History</td>
<td>IC-HIST</td>
</tr>
<tr>
<td>IC: Organization Procedures</td>
<td>IC-PROC</td>
</tr>
<tr>
<td>IC: Innovation-Organization Congruence</td>
<td>IC-FIT ...</td>
</tr>
</tbody>
</table>
What is a good code?

1. It should be sufficiently precise

2. It should enable a large part of the data material to be subsumed under it

3. It should be relevant for the research question

(Kelle/ Kluge 1999)
Research Question:
How does the performance based block funding of Australian universities affect the content of research?

Funding rules of the university = all rules of the university and its units which prescribe the distribution of funding for research purposes.
“The qualitative methods literature still lacks a systematic explication, presentation and discussion of coding technologies. Usually, coding strategies are handed over from one researcher generation to the next orally as an institute’s tradition.”
(Kelle/ Kluge 1999, translated by GL)

1. Indexing of textual units

2. Synopsis of all textual units which have certain categories and possibly other characteristics in common

3. Identification of structures and patterns in the data material which can lead to new categories and subcategories

(Kelle/ Kluge 1999)
Further Strategies after the initial coding:

a) collect all text segments that are tagged with the same code and compare these segments

b) co-occurrences of codes (collect all text segments that are tagged by the same two codes, three codes)

c) Reorganise the data by Axial coding or Selective coding (identifying core categories, find relations between codes)