

Closed Coding

Analyzing Qualitative Data - Tutorial @ VIS17

Melanie Tory

*A **code** in qualitative inquiry is most often a word or short phrase that symbolically assigns a summative, salient, essence capturing, and/or evocative attribute for a portion of language-based or visual data*

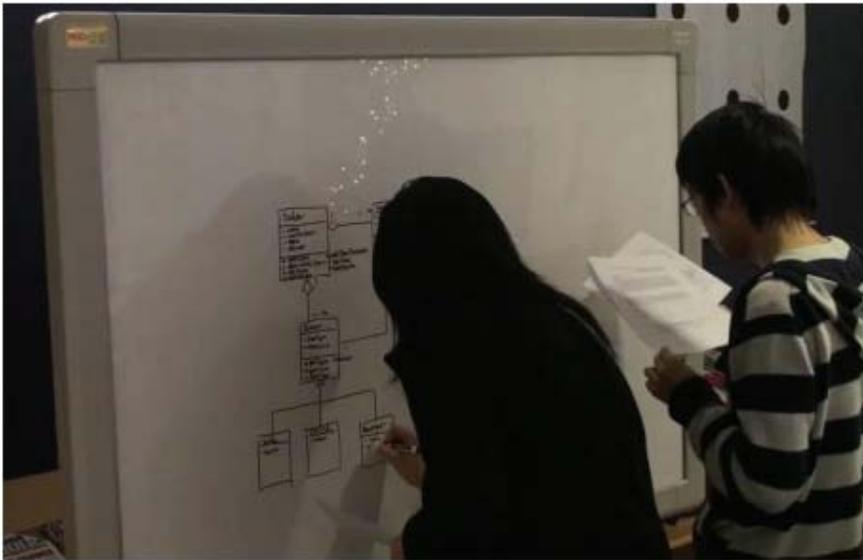
[Saldaña, J. (2013). *The coding manual for qualitative researchers*. Los Angeles: SAGE Publications.]

What is closed coding?

- Identifying and marking interesting items using a pre-established coding scheme
- Usually done on video, but could be transcripts, audio, etc.

When to use closed coding?

- Your research question involves counting or timing events
- You know a priori what the interesting events are
- You cannot mark the events by automatic logging



Closed vs. Open Coding

Closed Coding

- Suitable for timing or counting events
- Must know in advance what is relevant to your research question
- No richness
- Confirmatory research questions
- Largely quantitative

Open Coding

- Suitable for capturing and understanding meaning in your data
- You may have only some a priori idea about what is interesting
- Captures rich nuance
- Open-ended research questions
- Purely qualitative

Where do the (closed) codes come from?

- Previous research or theory
- May be obvious given your research questions / hypotheses
- Questions or topics from your interview guide
- Establish the coding scheme during the analysis:
 - If you already have a good idea:
 - View the video once or twice and agree on a set of codes
 - If you have no idea:
 - Step 1: Open coding (later)
 - Step 2: Closed coding: Revisit your data to apply the codes consistently throughout all your video.

[Mahyar et al., VAST 2014]

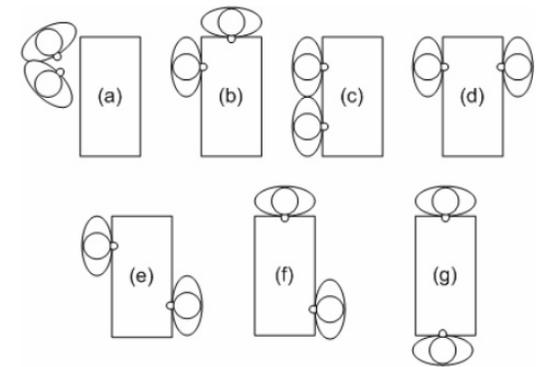


Code	Description	Example
DH	Having discussion or generating hypotheses	"US government is supplying the rebels with Lewisite."
RV	Referring to the visualization tool	"Link that with your apples. I will make a new node, linking Parazuelan."
CO	Coordinating the group	"Let's divide the work now, I will search for apples you look for flowers."
SA	Seeking awareness	"What do you guys got?"
VF	Verbalizing findings	"Former farm worker Francisco Dorado formed Shining Future in 1988."
QF	Questions about findings of another group member	"What did you find about apple bursting?"
RU	Relevant but otherwise uncategorized	"Oh okay found that article."

Table 1: Communication coding scheme.

What should you code?

- Depends on your research question
- Typical coded items may include:
 - Events
 - Behaviors
 - States of the people, system, or environment
 - Topics mentioned in an interview



Make sure you know exactly what you are coding and how you will use that information. Coding is very time consuming!

Code Types

Descriptive coding

- Codes are used to describe what is in the data

Analytic coding

- Codes are based on the analytical thinking by the researcher about why what is occurring in the data might be happening

Code Types

- Point events
 - Single point
- Duration events
 - Mark start and end – enables calculation of duration

Coding durations is much more time consuming. Only do it if you are certain you need duration information.

Code Types

Flat Coding

e.g. position of partners around a table:

- Same side
- Opposite sides
- Kitty corner

Hierarchical Coding

e.g. User goals:

- Share:
 - Share hypothesis
 - Share resources
 - Share idea
- Identify
 - Identify object
 - Identify person

Coding scheme

- Before applying closed coding, it is critical to establish a clear coding scheme
- List all the codes and their meanings
- Make sure all coders understand the scheme
 - Code a portion together, discuss disagreements, and modify the scheme.

Test your coding scheme by having another person use it. See if they understand it the same way you do.

Software Tools

- Pretty much necessary for closed coding of video
- Allow you to mark events via keystroke as the video plays – automatically logs the event + time
- Examples: Noldus Observer, MAXQDA



The screenshot displays the 'The Observer XT 11 - Parent-child interaction' software window. The main video area shows a father and a young girl sitting at a table, engaged in play with a pink toy and a book. The interface includes a 'Project Explorer' on the left, a 'Codes' panel with a tree view of behavior categories, and a 'Comment' panel with a text area and a table of recorded events.

Codes Panel:

- Verbal behavior
 - Talk to parent
 - Talk to doll
 - Talk to self
 - Play noise
 - Other verbal
 - No verbal
- Play behavior
 - Non-functional use
 - Functional use
 - Break up
 - Put together
 - Substitute
 - Agent play
 - Imaginary play
 - Game
 - No play behavior

Comment Panel:

Daddy, there are all letters here. It is difficult. A, B, C. Oooh, Daddy, what about all the letters? Do you know all the letters? A, B, C, D, E. You see, I know a lot of letters. This poster is in my school, in my classroom.

Event Log Table:

Time	Behavior	Modifier	Comment
47,66	No play behavior		
54,87	Functional use	Cash register Own intent	
89,94	Talk to parent		Daddy, there are all letters here.
102,42	No verbal		
108,47	No play behavior		
115,95	Talk to self		I'll put it in. Mmm, it does not d
125,07	No verbal		
129,06	Talk to parent		A Wins jigsaw that I have also g
133,33	No verbal		
136,19	Talk to parent		Yes
138,80	No verbal		
142,47	Talk to self		A baby born and a barbie. What
148,46	No verbal		

Multiple coders + inter-coder agreement

- Can your coding scheme be applied consistently by different researchers?
- Inter-coder agreement measures "the extent to which the different judges tend to assign exactly the same rating to each object" ([Tinsley & Weiss, 2000](#), p. 98);
- Often expected for papers that involve closed coding (but not open coding)
 - 2 coders is typical
 - Both must code some subset of the data with the same coding scheme

What about inter-coder agreement?

- or inter-coder reliability, or inter-rater reliability
- terms for the extent to which independent coders evaluate some data and reach the same conclusion
- measures "the extent to which the different judges tend to assign exactly the same rating to each object" ([Tinsley & Weiss, 2000](#), p. 98);
- in computational linguistics sometimes used to train an algorithm to do as well as a human

Inter-coder agreement measures

- **Cohen's kappa** – works for 2 coders, assumes nominal data
 - Gwet, K. L. (2014) "[Handbook of Inter-Rater Reliability \(4th Edition\)](#)"
- **Fleiss' kappa** – works for any fixed number of coders, assumes nominal data
 - Shrout, P. and Fleiss, J. L. (1979) "Intraclass correlation: uses in assessing rater reliability" in *Psychological Bulletin*. Vol. 86, No. 2, pp. 420–428
- **Krippendorff's alpha** – “generalizes several specialized agreement coefficients by accepting any number of observers, being applicable to nominal, ordinal, interval, and ratio levels of measurement, being able to handle missing data, and being corrected for small sample sizes”
 - (https://en.wikipedia.org/wiki/Inter-rater_reliability)
 - Krippendorff, K. (2013). *Content analysis: An introduction to its methodology, 3rd Edition*. Thousand Oaks, CA: Sage. pp. 221–250

Recap: Closed Coding

- Useful when you need to time or count events that can only be identified by a human observer
- Confirmatory research questions
- No richness, unlike open coding
- Advice:
 - Define and test your coding scheme in advance
 - Know your research question and code only what you need
 - Get help from a tool