

- illustrative charts, matrices, diagrams (Miles & Huberman, 1994; Morgan, Fellows, & Guevara, 2008; Northcutt & McCoy, 2004; Paulston, 2000)
- interrelationship (Saldaña, 2003)
- quick ethnography (Handwerker, 2001)
- splitting, splicing, and linking data (Dey, 1993)
- survey research (Fowler, 2001; Wilkinson & Birmingham, 2003)
- thematic analysis (Auerbach & Silverstein, 2003; Boyatzis, 1998)
- within-case and cross-case displays (Gibbs, 2007; Miles & Huberman, 1994)

Notes

Structural Coding is an analytic cousin of Holistic Coding. The latter, however, is more exploratory and even tentative in nature, while the former is framed and driven by a specific research question and topic.

DESCRIPTIVE CODING

Sources

Miles & Huberman, 1994; Saldaña, 2003; Wolcott, 1994

Description

(This method is also called "Topic Coding" in some of the literature, but Descriptive Coding will be used in this manual to align with Wolcott's terminology.)

Descriptive Coding summarizes in a word or short phrase – most often as a noun – the basic topic of a passage of qualitative data. To clarify, Tesch (1990) differentiates that "it is important that these [codes] are identifications of the topic, not abbreviations of the content. The topic is what is talked or written about. The content is the substance of the message" (p. 119).

Applications

Descriptive Coding is appropriate for virtually all qualitative studies, but particularly for beginning qualitative researchers learning how to code data, ethnographies, and studies with a wide variety of data forms (e.g., interview transcripts, field notes, journals, documents, diaries, correspondence, artifacts, video).

Many ethnographic studies usually begin with such general questions as "What is going on here?" and such reflective questions as "What is this a study about?" Descriptive Coding is just one approach to analyzing the data's basic topics to assist with answering these types of questions. Turner (1994) calls this cycle the development of a "basic vocabulary" of data to form "bread and

butter" categories for further analytic work (p. 199). Description is the foundation for qualitative inquiry, and its primary goal is to assist the reader to see what you saw and to hear what you heard (Wolcott, 1994, pp. 55, 412). Descriptive Codes from data collected across various time periods and charted in matrices are also essential for assessing longitudinal participant change (Saldaña, 2003, 2008).

Example

An ethnographer walks through an inner city, lower income neighborhood in a large metropolitan area and takes field notes describing the setting. The field notes are written descriptively – i.e., factually and objectively as much as possible. The italicized "OC" sections interspersed throughout are Observer's Comments (Bogdan & Biklen, 2007, pp. 163–4), subjective impressions or memos embedded within the factual description, which also merit consideration for codes. Note how several Descriptive Codes repeat as the topics shift:

Driving west along the highway's access road and up Main St. to Wildpass Rd. there were ¹ abandoned warehouse buildings in disrepair, ² spray painted gang graffiti on walls of several occupied and unoccupied buildings. I passed a ³ Salvation Army Thrift Store, Napa Auto Parts store, a tire manufacturing plant, old houses in-between industrial sites, an auto glass store, Market/Liquors, Budget Tire, a check cashing service. ⁴ More spray paint was on the walls.

¹ BUILDINGS

² GRAFFITI

³ BUSINESSES

⁴ GRAFFITI

OC: ⁵ *There seems to be an abundance of car-oriented businesses in the neighborhood. Industrial looking atmosphere – no "showroom" qualities. Here is where "repair" is more important than sales.*

⁵ BUSINESSES

I parked on Turquoise Rd. and walked along the periphery of an elementary school lot. ⁶ The majority of the homes had dirt front yards; the only vegetation growing were weeds. Maybe one house per block had what would be called a lawn. The majority seem unattended, not cared for. The homes look like they were built in the 1930s, 1940s at the latest. ⁷ I saw spray paint on the "No Trespassing" sign of the elementary school and smaller gang symbols on it.

⁶ HOUSES

⁷ GRAFFITI

⁸ A beer bottle and beer can were against the fence of the school. ⁹ Surfaces on the houses vary – maybe one per block looks fairly well painted. The majority are peeling in paint, rotted wood, various materials (wood, stucco, tin) on the same house.

⁸ TRASH

⁹ HOUSES

OC: *Priorities, energies, financial resources do not go into the appearance of homes. There are more important things to worry about.*

¹⁰ I notice that the grand majority of homes have chain link fences in front of them. There are many dogs (mostly German shepherds) with signs on fences that say "Beware of the Dog."

OC: *There's an attempt to keep things at home safe. Protection from robbers, protection of property. Keep your distance – this is mine.*

Depending on the volume of data, specificity, or amount of detail you may need for data analysis and categorization, Descriptive Codes can also be assigned more detailed *subcodes* when needed (Miles & Huberman, 1994, p. 61). Gibbs (2007) explains that the most general code is called the "parent" while its subcodes are the "children"; subcodes that share the same parent are "siblings" in a hierarchy (p. 74):

¹ Some houses, because of their disrepair, look abandoned. But the things seen through the window and in the yard clue you that someone still lives there. ² One house has a picture portrait of Jesus and a cross on its front wall. I notice TV antennas on several roofs.

OC: ³ *No cable TV – a luxury; can't afford it; not a priority.*

⁴ Laundry hangs in the back yards of several homes on a clothesline. ⁵ There's a house with a small statue of the Virgin Mary in front, ⁶ "Beware of the Dog/*Pero*" signs. ⁷ Desk chairs, worn upholstered furniture are in the front yards of some homes.

OC: ⁸ *It's all we've got.*

⁹ "Pussy" is a spray painted word I decipher on an abandoned car.

Analysis

Descriptive Coding leads primarily to a categorized inventory, tabular account, summary, or index of the data's contents. It is essential groundwork for Second Cycle coding and further analysis and interpretation (Wolcott, 1994, p. 55). In the example above, all qualitative data passages coded with HOUSES, for example, would be extracted from the main body of field notes and reassembled together in a separate file for an organized and categorized narrative portrait of the environment for further analysis:

Priorities, energies, and financial resources of this neighborhood's residents do not go into the appearance of their homes (there are more important things to worry about).

¹⁰ SECURITY

¹ HOUSES-DISREPAIR

² HOUSES-DÉCOR

³ ECONOMICS-HOMES

⁴ HOUSES-YARDS

⁵ HOUSES-DÉCOR

⁶ SECURITY-SIGNS

⁷ HOUSES-YARDS

⁸ ECONOMICS-HOMES

⁹ GRAFFITI-OBSCENE

Houses appear to have been built in the 1930s through 1940s. I notice TV antennas (rather than cable) on several roofs. Surfaces on the houses vary. Maybe one per block looks fairly well painted. But the majority's exteriors have peeling paint, rotting wood, and an assembly of wood, stucco, and tin on the same house. Some, because of their disrepair, look unattended and abandoned. But look in the yards and inside the homes through their windows: people still live here.

Laundry hangs on clotheslines in the back yards of several homes. The majority have dirt front yards where the only vegetation growing is weeds. Maybe one house per block has what would be called a lawn; other front yards contain desk chairs and worn upholstered furniture. There's a house with a small statue of the Virgin Mary in front. Another house has a picture portrait of Jesus and a cross on its front wall.

Some recommended ways to further analyze Descriptive Codes are (see Appendix D):

- Hypothesis Coding, Domain and Taxonomic Coding, Pattern Coding, and Focused Coding
- content analysis (Krippendorff, 2003; Weber, 1990; Wilkinson & Birmingham, 2003)
- cross-cultural content analysis (Bernard, 2006)
- domain and taxonomic analysis (Schensul, Schensul, & LeCompte, 1999; Spradley, 1979, 1980)
- frequency counts (LeCompte & Schensul, 1999)
- graph-theoretic techniques for semantic network analysis (Namey, Guest, Thairu, & Johnson, 2008)
- grounded theory (Charmaz, 2006; Corbin & Strauss, 2008; Glaser & Strauss, 1967; Strauss & Corbin, 1998)
- mixed-methods research (Creswell, 2009; Creswell & Plano Clark, 2007; Tashakkori & Teddlie, 2003)
- qualitative evaluation research (Patton, 2002)
- quick ethnography (Handwerker, 2001)
- thematic analysis (Auerbach & Silverstein, 2003; Boyatzis, 1998)
- within-case and cross-case displays (Gibbs, 2007; Miles & Huberman, 1994)

Notes

Descriptive Coding is a straightforward method for novices to qualitative research, particularly for those first using CAQDAS programs. The method categorizes data at a basic level to provide the researcher an organizational grasp of the study.

IN VIVO CODING

Sources

Charmaz, 2006; Corbin & Strauss, 2008; Glaser, 1978; Glaser & Strauss, 1967; Strauss, 1987; Strauss & Corbin, 1998

Description

(In Vivo Coding has also been labeled "Literal Coding" and "Verbatim Coding" in selected methods literature. In this manual, In Vivo Coding will be used since it is the most well-known label.)

In vivo's root meaning is "in that which is alive," and as a code refers to a word or short phrase from the actual language found in the qualitative data record, "the terms used by [participants] themselves" (Strauss, 1987, p. 33).

Folk or *indigenous* terms are participant-generated words from members of a particular culture, subculture, or microculture. Folk terms indicate the existence of the group's cultural categories (McCurdy, Spradley, & Shandy, 2005, p. 26). For example, some homeless youths say that they "sp'ange" (ask passers-by for "spare change"). In Vivo Coding is one method of extracting these indigenous terms (and refer to Domain and Taxonomic Coding for more specific categorization guidelines).

Applications

In Vivo Coding is appropriate for virtually all qualitative studies, but particularly for beginning qualitative researchers learning how to code data, and studies that prioritize and honor the participant's voice. In Vivo Coding is one of the methods to employ during grounded theory's Initial Coding but can be used with several other coding methods in this manual.

In Vivo Coding is particularly useful in educational ethnographies with youth. The child and adolescent voices are often marginalized, and coding with their actual words enhances and deepens an adult's understanding of their cultures and world-views. In Vivo Coding is also quite applicable to action research (Stringer, 1999) since one of the genre's primary goals is to frame the facilitator's interpretations of terms "that participants use in their everyday lives, rather than in terms derived from the academic disciplines or professional practices" (p. 91).

Example

An adult female interviewer talks to Tiffany, a 16 year-old teenage girl, about her friendships at high school. Note how all In Vivo Codes are placed in quotation marks, and how virtually each line of data gets its own code:

I ¹ hated school last year. ² Freshman year, it was awful, I hated it. And ³ this year's a lot better actually. Um. I ⁴ don't know why. I guess, over the summer I kind of ⁵ stopped caring about what other people thought and cared more about, just, I don't know. It's ⁶ hard to explain. I ⁷ found stuff out about myself, and so I went back, and all of a sudden I found out that when I ⁸ wasn't trying so hard to ⁹ have people like me and to do ¹⁰ what other people wanted, people ¹¹ liked me more. It was ¹² kind of strange. Instead of ¹³ trying to please them all the time, they liked me more when I ¹⁴ wasn't trying as hard. And, I don't know, like every-, everybody might, um, people who are just, kind of, ¹⁵ friends got closer to me. And people who didn't really know me ¹⁶ tried to get to know me. ¹⁷ I don't know.

- ¹ "HATED SCHOOL"
- ² "FRESHMAN YEAR AWFUL"
- ³ "THIS YEAR'S BETTER"
- ⁴ "DON'T KNOW WHY"
- ⁵ "STOPPED CARING"
- ⁶ "HARD TO EXPLAIN"
- ⁷ "FOUND STUFF OUT"
- ⁸ "WASN'T TRYING SO HARD"
- ⁹ "HAVE PEOPLE LIKE ME"
- ¹⁰ "WHAT OTHER PEOPLE WANTED"
- ¹¹ "LIKED ME MORE"
- ¹² "KIND OF STRANGE"
- ¹³ "TRYING TO PLEASE THEM"
- ¹⁴ "WASN'T TRYING AS HARD"
- ¹⁵ "FRIENDS GOT CLOSER"
- ¹⁶ "TRIED TO KNOW ME"
- ¹⁷ "I DON'T KNOW"

Analysis

As you read interview transcripts or other documents that feature participant voices, attune yourself to words and phrases that seem to call for bolding, underlining, italicizing, highlighting, or vocal emphasis if spoken aloud. Their salience may be attributed to such features as impacting nouns, action-oriented verbs, evocative word choices, clever or ironic phrases, similes and metaphors, etc. If the same words, phrases, or variations thereof are used often by the participant (such as "I don't know" in the example above), and seem to merit an In Vivo Code, apply it. In Vivo Codes "can provide a crucial check on whether you have grasped what is significant" to the participant, and may help "crystallize and condense meanings" (Charmaz, 2006, p. 57). Thus, keep track of codes that are participant-inspired rather than researcher-generated by always putting In Vivo Codes in quotation marks: "HATED SCHOOL".

Key writers of grounded theory advocate meticulous work and that an In Vivo (or other) Code should appear next to every line of data. Depending on your goals, In Vivo Codes can be applied with less frequency, such as one word or phrase for every three to five sentences. In the interview excerpt above, rather than 17 In Vivo Codes I could have limited the number to four: "FRESHMAN YEAR AWFUL", "FOUND STUFF OUT", "WASN'T TRYING SO HARD", "FRIENDS GOT CLOSER". But don't infer that there is a fixed rule or formula for an average number of codes per page or a recommended ratio of codes to text. Trust your instincts with In Vivo Coding.

When something in the data appears to stand out, apply it as a code. Researcher reflection through analytic memo writing, coupled with Second Cycle coding, will condense the number of In Vivo Codes and provide a reanalysis of your initial work.

In Vivo Codes capture "behaviors or processes which will explain to the analyst how the basic problem of the actors is resolved or processed" (Strauss, 1987, p. 33) and "help us to preserve participants' meanings of their views and actions in the coding itself" (Charmaz, 2006, p. 55). In Vivo Codes can provide imagery, symbols, and metaphors for rich category, theme, and concept development. Strauss (1987, p. 160) also recommends that researchers examine In Vivo Codes not just as themes but as possible *dimensions* of categories — i.e., the continuum or range of a property.

In Vivo Codes could be used as the sole coding method for the First Cycle of data analysis, and the sole method of choice for small-scale studies, but that may limit the researcher's perspective on the data, a perspective that can contribute to more conceptual and theoretical views about the phenomenon or process. Sometimes the participant says it best; sometimes the researcher does. Be prepared and willing to mix and match coding methods as you proceed with data analysis.

CAQDAS programs such as NVivo, MAXQDA, and ATLAS.ti make In Vivo Coding easy by permitting the analyst to select a word or small phrase from the data, clicking a dedicated icon, and assigning the selected text as an In Vivo Code. But be aware that some CAQDAS functions will retrieve multiple text units only if they share the *exact* same code you have applied to them. In Vivo coded data most often are so unique that they will require careful review and self-categorization into an NVivo node, for example.

Some recommended ways to further analyze In Vivo Codes are (see Appendix D):

- Second Cycle coding methods
- action research (Altrichter, Posch, & Somekh, 1993; Stringer, 1999)
- case studies (Merriam, 1998; Stake, 1995)
- domain and taxonomic analysis (Schensul, Schensul, & LeCompte, 1999; Spradley, 1979, 1980)
- frequency counts (LeCompte & Schensul, 1999)
- grounded theory (Charmaz, 2006; Corbin & Strauss, 2008; Glaser & Strauss, 1967; Strauss & Corbin, 1998)
- interactive qualitative analysis (Northcutt & McCoy, 2004)
- memo writing about the codes (Charmaz, 2006; Corbin & Strauss, 2008; Glaser, 1978; Glaser & Strauss, 1967; Strauss, 1987)
- metaphoric analysis (Coffey & Atkinson, 1996; Todd & Harrison, 2008)
- narrative inquiry and analysis (Clandinin & Connelly, 2000; Coffey & Atkinson, 1996; Cortazzi, 1993; Daiute & Lightfoot, 2004; Riessman, 2008)

- phenomenology (van Manen, 1990)
- poetic and dramatic constructions (Denzin, 1997; Glesne, 2006; Knowles & Cole, 2008; Saldaña, 2005a)
- polyvocal analysis (Hatch, 2002)
- portraiture (Lawrence-Lightfoot & Davis, 1997)
- qualitative evaluation research (Patton, 2002)
- thematic analysis (Auerbach & Silverstein, 2003; Boyatzis, 1998)

Notes

Researchers new to coding qualitative data often find In Vivo Coding a safe and secure method with which to begin. But be wary of overdependence on the strategy because it can limit your ability to transcend to more conceptual and theoretical levels of analysis and insight.

PROCESS CODING

Sources

Bogdan & Biklen, 2007; Charmaz, 2002; Corbin & Strauss, 2008; Strauss & Corbin, 1998

Description

Process Coding uses gerunds ("ing" words) exclusively to connote action in the data (Charmaz, 2002). Simple observable activity (e.g., reading, playing, watching TV, drinking coffee) and more general conceptual action (e.g., struggling, negotiating, surviving, adapting) can be coded as such through a Process Code. The processes of human action can be "strategic, routine, random, novel, automatic, and/or thoughtful" (Corbin & Strauss, 2008, p. 247).

Applications

Process Coding is appropriate for virtually all qualitative studies, but particularly for those that search for "ongoing action/interaction/emotion taken in response to situations, or problems, often with the purpose of reaching a goal or handling a problem" (Corbin & Strauss, 2008, pp. 96–7). For grounded theory, Process Coding happens simultaneously with Initial Coding and Axial Coding, and a search for *consequences* of action/interaction is also part of the process. Subprocesses are usually "the individual tactics, strategies, and routine actions that make up the larger act" (Strauss & Corbin, 1998, p. 169).

Like In Vivo Coding, Process Coding is not necessarily a specific method that should be used as the sole coding approach to data, though it can be with small-scale projects.

Examples

A first example from field notes taken while observing at an elementary school uses Process Coding to document basic activity. Note how the codes are all gerund-based:

¹ Mrs. Jackson rises from her desk and announces, "OK, you guys, let's get lined up for lunch. Row One." Five children seated in the first row of desks rise and walk to the classroom door. Some of the seated children talk to each other. ² Mrs. Jackson looks at them and says, "No talking, save it for the cafeteria. ³ Row Two." Five children seated in the second row of desks rise and walk to the children already standing in line. ⁴ A girl asks Mrs. Jackson, "Does it smell like pizza?" Mrs. Jackson smiles and replies, "Yeah, smells good, doesn't it?"

¹ LINING UP FOR LUNCH

² MANAGING BEHAVIOR

³ LINING UP FOR LUNCH

⁴ BANTERING

As a second example, an adult female interviewer talks to a teenage girl about rumors. Note how the codes are all gerund-based (and note that the interviewer's questions and responses are not coded – just the participant's responses):

TIFFANY: Well, ¹ that's one problem, that [my school is] pretty small, so ² if you say one thing to one person, ³ and then they decide to tell two people, ⁴ then those two people tell two people, and ⁵ in one period everybody else knows. ⁶ Everybody in the entire school knows that you said whatever it was. So ...

I: Have you ever had rumors spread about you?

TIFFANY: Yeah, ⁷ it's just stupid stuff, completely outlandish things, too. ⁸ I, I don't really want to repeat them.

I: That's OK, you don't have to.

TIFFANY: ⁹ They were really, they were ridiculous. ¹⁰ And the worst thing about rumors, ¹¹ I don't really care if people think that, because obviously they're pretty stupid to think that in the first place. But ¹² the thing I care about is, like, last year, especially freshman year, was a really horrible year school-wise. And, ¹³ I guess it was good in a way that you find out who your real friends are, because ¹⁴ some of them turned on me and ¹⁵ then started to say that those things

¹ PROBLEMIZING SCHOOL SIZE

² SAYING ONE THING

³ TELLING OTHERS

⁴ TELLING OTHERS

⁵ EVERYBODY KNOWING

⁶ KNOWING WHAT YOU SAID

⁷ REJECTING RUMORS

⁸ NOT REPEATING WHAT WAS SAID

⁹ REJECTING RIDICULOUSNESS

¹⁰ CRITICIZING RUMORS

¹¹ NOT CARING WHAT PEOPLE THINK

¹² REMEMBERING A HORRIBLE YEAR

¹³ FINDING OUT WHO YOUR REAL FRIENDS ARE

¹⁴ TURNING ON YOU

¹⁵ SAYING THINGS ARE TRUE

were true and, like, ¹⁶ then people thought, "Well that person's her friend, so they must know." And so, ¹⁷ it just made the entire thing worse. And ¹⁸ you really learn a lot about people and, uh, and ¹⁹ who your real friends are. LuAnn's ²⁰ probably the only person who's really stuck by me this entire time, and ²¹ just laughed at whatever they said.

¹⁶ ASSUMING BY OTHERS

¹⁷ MAKING THINGS WORSE

¹⁸ LEARNING A LOT ABOUT PEOPLE

¹⁹ LEARNING WHO YOUR FRIENDS ARE

²⁰ STICKING BY FRIENDS

²¹ LAUGHING AT WHAT OTHERS SAY

Analysis

The conventions of story-line are used in analytic memo writing when reviewing data for process – e.g., the first step, the turning point, the second step, the third step. Participant language with transitional indicators such as "if," "when," "because," "then," "and so," etc. clue the researcher to a sequence or process in action. These sequences or processes can be ordered as a numeric series of actions, listed as a bullet-pointed set of outcomes, or graphically represented with first-draft illustrations as a flow diagram. Simple examples based on the interview above include:

Narrative – Spreading Rumors:

1. [!]if you say one thing to one person,
2. and then they decide to tell two people,
3. then those two people tell two people,
4. and in one period everybody else knows.

Coded Process – Spreading Rumors:

1. SAYING ONE THING
2. TELLING OTHERS
3. TELLING OTHERS
4. EVERYBODY KNOWING

See Figure 3.3 for an illustrated process of spreading rumors.

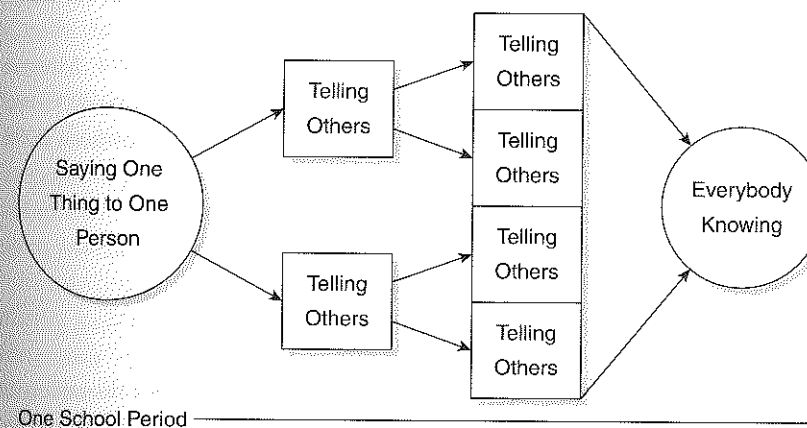


FIGURE 3.3 An illustrated process for spreading rumors

Consequences of Rumors [the Process Codes that led to the outcomes are listed for reference]:

- False accusations (SAYING THINGS ARE TRUE, ASSUMING BY OTHERS)
- Confrontation (TURNING ON YOU)
- Hurt feelings (MAKING THINGS WORSE, CRITICIZING RUMORS)
- Bad memories (NOT REPEATING WHAT WAS SAID, REMEMBERING A HORRIBLE YEAR)
- Reinforcement of loyalties (FINDING OUT WHO YOUR REAL FRIENDS ARE, LEARNING WHO YOUR FRIENDS ARE, STICKING BY FRIENDS)
- Social awareness (LEARNING A LOT ABOUT PEOPLE)
- "High road" personal growth (REJECTING RUMORS, REJECTING RIDICULOUSNESS, CRITICIZING RUMORS, NOT CARING WHAT PEOPLE THINK, LAUGHING AT WHAT OTHERS SAY)

Researcher reflection through analytic memo writing, coupled with Second Cycle coding, will condense the number of Process Codes and provide a reanalysis of your initial work. Dey (1993) encourages consideration of the complex interplay of factors that compose a process and how we can "obtain a sense of how events originate and evolve, and their shifting significance for those involved. Process refers to movement and change over time. In place of a static description, we can develop a more dynamic account of events" (p. 38). CAQDAS program linking functions enable you to mark the trail of participant process throughout the data corpus.

Some recommended ways to further analyze Process Codes are (see Appendix D):

- Second Cycle coding methods
- action research (Altrichter, Posch, & Somekh, 1993; Stringer, 1999)
- case studies (Merriam, 1998; Stake, 1995)
- cognitive mapping (Miles & Huberman, 1994; Northcutt & McCoy, 2004)
- decision modeling (Bernard, 2006)
- grounded theory (Charmaz, 2006; Corbin & Strauss, 2008; Glaser & Strauss, 1967; Strauss & Corbin, 1998)
- illustrative charts, matrices, diagrams (Miles & Huberman, 1994; Morgan, Fellows, & Guevara, 2008; Northcutt & McCoy, 2004; Paulston, 2000)
- memo writing about the codes (Charmaz, 2006; Corbin & Strauss, 2008; Glaser, 1978; Glaser & Strauss, 1967; Strauss, 1987)
- splitting, splicing, and linking data (Dey, 1993)
- thematic analysis (Auerbach & Silverstein, 2003; Boyatzis, 1998)
- vignette writing (Erickson, 1986; Graue & Walsh, 1998)

Notes

To appreciate the breadth & depth of Corbin & Strauss's (2008) discussion of Process Coding, readers are referred to their book, *Basics of Qualitative Research*, for

a full explanation and thorough examples of memo writing that capture process, and how micro- and macro-levels of analysis can be projected onto the data.

Also refer to Dramaturgical Coding for a comparable approach to analyzing a participant's tactics and strategies.

INITIAL CODING

Sources

Charmaz, 2006; Corbin & Strauss, 2008; Glaser, 1978; Glaser & Strauss, 1967; Strauss, 1987; Strauss & Corbin, 1998

Description

(Earlier publications in grounded theory refer to Initial Coding as "Open Coding." Charmaz's [2006] term will be used in this manual since it implies an initiating procedural step in harmony with First Cycle coding processes.)

Initial Coding is breaking down qualitative data into discrete parts, closely examining them, and comparing them for similarities and differences (Strauss & Corbin, 1998, p. 102). The goal of Initial Coding, particularly for grounded theory studies, is "to remain open to all possible theoretical directions indicated by your readings of the data" (Charmaz, 2006, p. 46). It is an opportunity for you as a researcher to reflect deeply on the contents and nuances of your data and to begin taking ownership of them. Initial Coding is not necessarily a specific formulaic method. It is a First Cycle, open-ended approach to coding the data with some recommended general guidelines.

Initial Coding can employ In Vivo Coding or Process Coding, for example, or other selected methods profiled in this manual. At times you may notice that elements of a possible or developing category are contained within the data. If so, code them during the initial cycle.

Applications

Initial Coding is appropriate for virtually all qualitative studies, but particularly for beginning qualitative researchers learning how to code data, ethnographies, and studies with a wide variety of data forms (e.g., interview transcripts, field notes, journals, documents, diaries, correspondence, artifacts, video).

Initial Coding is intended as a starting point to provide the researcher with analytic leads for further exploration and "to see the direction in which to take [this] study" (Glaser, 1978, p. 56). But Clarke (2005) recommends a period of "digesting and reflecting" on the data before beginning Initial Coding ventures (p. 84). All proposed codes during this cycle are tentative and provisional. Some

codes may be reworded as analysis progresses. The task can also alert the researcher that more data are needed to support and build an emerging theory.

Charmaz (2006) advises that detailed, line-by-line Initial Coding (as it is outlined below) is perhaps more suitable for interview transcripts than for researcher-generated field notes. But Clarke (2005), in her postmodern approach to grounded theory, stresses the need to examine the nonhuman material elements of our social world that *will* be found in field notes and artifacts. She also prescribes that for initial visual discourse/materials analysis, coding is not important but interpreting and analytic memo writing are critical. It is thus the memos themselves – the researcher's narratives – that are coded for further analysis.

Example

An adult female interviewer asks a 16 year-old girl about her social friendships (note that the interviewer's questions, responses, and comments are not coded – just the participant's responses). It is not required during Initial Coding, but since one of the eventual goals of grounded theory is to formulate categories from codes, I occasionally include not only a Process Code (e.g., LABELING), but also subcode it with specific referents (e.g., LABELING: "GEEKY PEOPLE", LABELING: "JOCKS"). These referents may or may not later evolve into categories, dimensions, or properties of the data as analysis continues:

[I: Last week you were talking about the snobby girls at lunchroom. And then you just were talking about that you didn't like some people because they were cliquish. So what kind of, who are your friends? Like, what kind of people do you hang out with?]

TIFFANY: ¹ I hang out with everyone. Really, ² I choose. Because I've been [living] here for so long, you get, ³ I can look back to kindergarten, and at some point I was ⁴ best friends with everybody who's been here, ⁵ practically. [I: You mean in choir?]

TIFFANY: ⁶ Almost everybody in my grade. No, in school. And so there are ⁷ certain people that I've just been ⁸ friends with since forever. And then it's ⁹ not fair of me to stereotype either, like, say, "Oh well, like, the ¹⁰ really super popular pretty girls are all mean and ¹¹ they're all snobby and they all talk about each other," ¹² 'cause they don't. Some of them, some of them don't. And ¹³ those are the ones I'm friends, friends with. And then there are

- ¹ "HANGING OUT WITH EVERYONE"
- ² "CHOOSING" WHO YOU HANG OUT WITH
- ³ RECALLING FRIENDSHIPS
- ⁴ "BEST FRIENDS WITH EVERYBODY"
- ⁵ QUALIFYING: "PRACTICALLY"
- ⁶ QUALIFYING: "ALMOST"
- ⁷ FRIENDS WITH "CERTAIN PEOPLE"
- ⁸ FRIENDS WITH "SINCE FOREVER"
- ⁹ "NOT FAIR TO STEREOTYPE"
- ¹⁰ LABELING: "REALLY SUPER POPULAR PRETTY GIRLS"
- ¹¹ IDENTIFYING STEREOTYPES
- ¹² DISPELLING STEREOTYPES
- ¹³ CHOOSING FRIENDS: "SUPER POPULAR PRETTY GIRLS"

the, there are the ¹⁴ geeky people. ¹⁵ Some of them though, the geeks at our school, aren't like harmless geeks. They're like ¹⁶ strange-psycho-killer-geek-people-who-draw-swastikas-on-their-backpacks ¹⁷ kind of geeks. [I: So are they like Colorado ... ?] TIFFANY: Yeah, ¹⁸ some of them are kind of, it's really scary. ¹⁹ But then again ²⁰ there's not the complete stereotype. Some of the, ²¹ not all of them are completely, like, wanna kill all the popular people. So, ²² I'm friends with those people. And then the ²³ jocks, ²⁴ not all of the guys are idiots, so ²⁵ I'm friends with the ones who can carry on a conversation.

[I: (laughs) So, so you wouldn't put yourself in any of the ...]

TIFFANY: ²⁶ I'm friends with someone because of who they are, ²⁷ not because of what group they, they hang out in basically. 'Cause I think ²⁸ that's really stupid to be, like, ²⁹ "What would people think if they saw me walking with this person?" or something.

[I: So you wouldn't define yourself with any specific group?]

TIFFANY: ³⁰ No.

[I: Do you think anyone else would define you as ...]

TIFFANY: I think people, ³¹ people define me as popular. Mainly because I would rather hang out with someone who's ³² good hearted but a little slow, compared to someone ³³ very smart but very evil (chuckle).

Analysis

Though you're advised to code quickly and spontaneously, pay meticulous attention to the rich dynamics of data through line-by-line coding – a "microanalysis" of the corpus (Strauss & Corbin, 1998, p. 57) or what is

- ¹⁴ LABELING: "GEEKY PEOPLE"
- ¹⁵ QUALIFYING: "SOME OF THEM"
- ¹⁶ LABELING: "STRANGE-PSYCHO-KILLER-GEEK"
- ¹⁷ QUALIFYING: "KIND OF"
- ¹⁸ QUALIFYING: "SOME OF THEM"
- ¹⁹ QUALIFYING: "BUT THEN..."
- ²⁰ DISPELLING STEREOTYPES
- ²¹ QUALIFYING: "NOT ALL OF THEM"
- ²² CHOOSING FRIENDS: "GEEKS"
- ²³ LABELING: "JOCKS"
- ²⁴ DISPELLING STEREOTYPES
- ²⁵ CHOOSING FRIENDS: JOCKS
"WHO CAN CARRY ON A CONVERSATION"
- ²⁶ CRITERIA FOR FRIENDSHIP: "WHO THEY ARE"
- ²⁷ CRITERIA FOR FRIENDSHIP: NOT GROUP MEMBERSHIP
- ²⁸ ETHICS OF FRIENDSHIP
- ²⁹ NOT CONCERNED WITH WHAT OTHERS THINK
- ³⁰ MAINTAINING INDIVIDUALITY
- ³¹ DEFINING SELF THROUGH OTHERS: "POPULAR"
- ³² CRITERIA FOR FRIENDSHIP: "GOOD HEARTED BUT SLOW"
- ³³ CRITERIA FOR FRIENDSHIP: NOT THOSE "VERY SMART BUT VERY EVIL"

colloquially referred to as data “splitting.” Selected writers of grounded theory acknowledge that such detailed coding is not always necessary, so sentence-by-sentence or even paragraph-by-paragraph coding is permissible depending on your research goals and analytic work ethic. Codes in parentheses or accompanied with question marks can be part of the process for analytic follow-up or memo writing.

One major facet of Initial Coding to grounded theory methodologists is the search for *processes* – participant actions that have antecedents, causes, consequences, and a sense of temporality. CHOOSING FRIENDS and DISPELLING STEREOTYPES (the “-ing” codes) are two such active processes culled from Initial Coding in the example above that may or may not be developed further during Second Cycle coding.

Also during this or later cycles of grounded theory coding, there will be a search for the *properties* and *dimensions* of categories – conceptual ideas that bring together similarly-coded and related passages of data. In the example above, the process of CHOOSING FRIENDS has the categories GEEKS and JOCKS, each with selected properties or CRITERIA FOR FRIENDSHIP – e.g., JOCKS “WHO CAN CARRY ON A CONVERSATION”. These categories and their properties may or may not be developed further during Second Cycle coding.

A personal debriefing or “reality check” by the researcher is critical during and after the initial coding of qualitative data, thus an analytic memo is written to reflect on the process thus far. Notice that *memo writing also serves as a code- and category-generating method*. An analytic memo excerpt based on the coding example above reads:

30 May 2007

CODE: DISCRIMINATING

Adolescence seems to be a time of contradictions, especially when it comes to choosing your friends. Tiffany is aware of the cliques, their attributed labels, and their accompanying stereotypes at school. Though she herself labels a group in broad terms (“super popular pretty girls,” “geeks,” “jocks”), she acknowledges that there are subcategories within the group and exceptions to the stereotypes. Though she states she chooses friends for “who they are” and does not care what others think of her for the friends she “hangs out” with, she has explicit criteria for who from the stereotyped group she’s friends with. She thinks others define her as “popular” and, as such, is aware of the stereotypical attributes from others that come with being part of that group. Her language is peppered with qualifiers: “some of them,” “not all of them,” “kind of,” “practically,” “almost,” etc. Developmentally, as an adolescent she seems right on target for *discriminating* – DISCRIMINATING cognitively the exceptions to stereotypes, while being SOCIALLY DISCRIMINATE about her own friendships.

Strauss & Corbin (1998) advise that at least ten interviews or observations with detailed coding are necessary for building a grounded theory (p. 281). Though

the First Cycle process is labeled “initial” or “open,” suggesting a wide variance of possibilities, Glaser (1978) recommends that the codes developed at this stage somehow relate to each other (p. 57). Strauss & Corbin (1998) also recommend that a separate list of all emerging codes be kept as analysis continues to help the researcher visualize the work in progress and to avoid duplication. CAQDAS code lists will help immensely with this.

Some recommended ways to further analyze Initial Codes are (see Appendix D):

- Second Cycle coding methods, particularly Focused Coding and Axial Coding
- grounded theory (Charmaz, 2006; Corbin & Strauss, 2008; Glaser & Strauss, 1967; Strauss & Corbin, 1998)
- memo writing about the codes (Charmaz, 2006; Corbin & Strauss, 2008; Glaser, 1978; Glaser & Strauss, 1967; Strauss, 1987)
- situational analysis (Clarke, 2005)
- thematic analysis (Auerbach & Silverstein, 2003; Boyatzis, 1998)

Notes

Initial Coding can range from the descriptive to the conceptual to the theoretical, depending on what you observe in and infer from the data, and depending on your personal knowledge and experiences you bring to your reading of the phenomena – “experiential data” that should be integrated into your analytic memo writing, yet not the central focus of it.