3 Coding in Grounded Theory Practice

The first analytic turn in our grounded theory journey brings us to coding. Grounded theory coding requires us to stop and ask analytic questions of the data we have gathered. These questions not only further our understanding of studied life but also help us direct subsequent data-gathering toward the analytic issues we are defining. Grounded theory coding consists of at least two phases: initial and focused coding. During initial coding we study fragments of data—words, lines, segments, and incidents—closely for their analytic import. From time to time, we may adopt our participants' telling terms as *in vivo* codes. While engaging in focused coding, we select what seem to be the most useful initial codes and test them against extensive data. Throughout the process, we compare data with data and then data with codes. We may follow special procedures to elaborate our codes or move to extant theoretical codes but only if indicated by our emerging analysis. Signposts and guides make our sojourn with coding accessible and ease our way around obstacles.

Consider the following interview excerpt from Bonnie Presley, who had long known she had systemic lupus erythematosus and had recently learned that she also had discoid lupus erythematosus. At the time of this interview, Bonnie was 48 years old and divorced from her second husband. After leaving a partner with whom she had lived for several years, she lived alone with her three cats. During the past year, she had had several immobilizing episodes of illness; the first one had been life-threatening. Currently, she was attempting to regain her strength after being ill for almost three months. Bonnie's good friend and neighbor, Linda, was keeping a watchful eye on her. Linda was bringing Bonnie food and made her tea since Bonnie felt too weak to care for herself.

Although Bonnie's adult daughter, Amy, now lived in the area, their calls and Amy's visits remained sporadic. Years before, Amy could not understand how the mother she had known as a fitness buff could have become so sedentary. Bonnie's youthful physical appearance belied her health status because her symptoms remained invisible to an untrained eye. In the early years of her illness, Bonnie had found it difficult to tell Amy about her illness and its seriousness.[1] Amy had moved away before Bonnie first became ill and Bonnie had either understated what was happening or avoided telling Amy. Bonnie recounted her realization about how she had told Amy the news about her recent crisis:

> She found out from Linda that I was, had been in bed for days and she called me up, 'You never tell me, and I have to find out from Linda,' and 'Why don't you tell me who you are and what's going on and ...' Well, I don't know how long after that, but that Saturday the pain started right here and it, throughout the day it got worse and worse and worse. And she–I kept thinking that, well, I can deal with this, so I took some kind of a pain pill and nothing helped. And that was about one in the afternoon. Well, it got worse and worse so that every time I took a breath the pain was horrible, so by seven, eight o'clock that night, I was scared because I knew that if it got any worse I wasn't going to be able to breathe. So I called her and then I told her what was going on, that I was going to be driven to the doctor because they were going to try giving me shots of zylocain or something to try to locate a point to where maybe it would go in there and numb the pain for me so that I could breathe. Well, I called her and told her this. And I have a car phone. She says, 'Well, Mom I'll call you later or you call me.' Well, I didn't call her; she didn't call me. That was Saturday night. She didn't call me until—she called me about noon on Monday, and I finally said, 'Well look, this is why I don't tell you, because when I told you Saturday night, you never called, you didn't care or anything and it really hurt my feelings. So that's why I don't tell you when I have this going on.' And she said to me, 'Well, Mom, you sounded perfectly fine.' And I said, 'Well, what do you expect me to do, become an emotional wreck or something?' I said, 'I have to keep everything still and quiet in me in order to control, because if I went into emotional frenzy, I would have not been able to breathe,' you know. So she started really trying to understand that just because I was scared to death, I was in horrible pain, but when I called her, I guess I was just a normal mom.

What sense might we make of stories like Bonnie's? How do we synthesize hundreds of pages of interviews, fieldnotes, documents, and other texts to develop a grounded theory? Whether we have collected stories, scenes, or written statements, we study and define these materials to analyze what happened and what they might mean.

Coding means categorizing segments of data with a short name that simultaneously summarizes and accounts for each piece of data. Your codes show how you select, separate, and sort data to begin an analytic accounting of them.

Qualitative coding, the process of defining what the data are about, is our first analytic step. Coding means naming segments of data with a label that simultaneously categorizes, summarizes, and accounts for each piece of data.[2] Coding is the first step in moving beyond concrete statements in the data to making analytic interpretations. We aim to make an interpretative rendering that begins with coding and illuminates studied life.

BOX 3.1 GROUNDED THEORY CODING EXAMPLE

Receiving second-hand news Being left out; Accusing mother of repeated not telling; (questioning ethical stance?) Being confronted Facing self and identity questions; Demanding self-disclosure and information Experiencing escalating pain Expecting to manage pain Inability to control pain Rapid worsening of pain Having excruciating pain Becoming frightened; Foreseeing breathing crisis Breaking the news; Informing daughter of plan Explaining projected treatment Having access for making contact Leaving follow-up contact open-ended No follow-up Ascertaining the time between contacts Explaining lack of disclosure Accusing daughter of not caring Expressing hurt; Assuming lack of caring; Making negative inferences (of a moral lapse?) Accounting for not telling Sounding fine Questioning daughter's expectations Explaining need for emotional control Seeing life-threatening risk of losing control Teaching that mode of telling does not reflect state of being Sounding like a 'normal' mom	She found out from Linda that I was, had been in bed for days and she called me up, 'You never tell me, and I have to find out from Linda,' and 'Why don't you tell me who you are and what's going on and ...' Well, I don't know how long after that, but that Saturday the pain started right here and it, throughout the day it got worse and worse and worse. And she–I kept thinking that, well, I can deal with this, so I took some kind of a pain pill and nothing helped. And that was about one in the afternoon. Well, it got worse and worse so that every time I took a breath the pain was horrible, so by seven, eight o'clock that night, I was scared because I knew that if it got any worse I wasn't going to be able to breathe. So I called her and then I told her what was going on, that I was going to be driven to the doctor because they were going to try giving me shots of zylocain or something to try to locate a point to where maybe it would go in there and numb the pain for me so that I could breathe. Well, I called her and told her this. And I have a car phone. She says, 'Well, Mom I'll call you later or you call me.' Well, I didn't call her; she didn't call me. That was Saturday night. She didn't call me until—she called me about noon on Monday, and I finally said, 'Well look, this is why I don't tell you, because when I told you Saturday night, you never called, you didn't care or anything and it really hurt my feelings. So that's why I don't tell you when I have this going on.' And she said to me, 'Well, Mom, you sounded perfectly fine.' And I said, 'Well, what do you expect me to do, become an emotional wreck or something?' I said, 'I have to keep everything still and quiet in me in order to control, because if I went into emotional frenzy, I would have not been able to breathe,' you know. So she started really trying to understand that just because I was scared to death, I was in horrible pain, but when I called her, I guess I was just a normal mom.

Our codes show how we select, separate, and sort data to begin an analytic accounting of them. Qualitative codes take segments of data apart, name them in concise terms, and propose an analytic handle to develop abstract ideas for interpreting each segment of data. As we code, we ask: which theoretical categories might these statements indicate?

You might have wondered what qualitative codes look like and how researchers construct them. A quick look at my codes of Bonnie Presley's story will give you an idea (see Box 3.1).

The codes in Box 3.1 attempt to portray meanings and actions in Bonnie's story. We gain a sense of both Bonnie's and Amy's concerns, as Bonnie presents them. Her story shows how telling news can be fraught with problems. Misunderstandings and dilemmas arise. Hesitancies occur. Accusations ensue. Explanations follow. Telling the news can open the self to view, risk emotional costs, and force questions about relationships. Not telling or delayed telling can also rent or rupture bonds. Familial failures, ethical slights, and moral claims accrue, from one or another person's view. Rhetorical styles may be meant–or misunderstood–as delivering fundamental judgments. For both Bonnie and Amy, disclosing illness became a contested area in which charged questions ignited about whom each was to the other. Events may force disclosure, as Bonnie's story indicates. What people tell, when they tell it, and how they tell it all matter. How Bonnie told her daughter affected how her daughter understood and acted on the news. Bonnie had concentrated on not risking loss of emotional control but later realized that her straightforward way of informing Amy may have understated the seriousness of the episode and fueled misunderstandings. By maintaining emotional control when informing her daughter, Bonnie's daughter thought she 'sounded perfectly fine,' like 'just a normal mom.'

Note that the codes stick closely to the data, show actions, and indicate how dilemmas surrounding disclosure arise. Certain codes, such as 'being left out,' 'facing self and identity questions,' 'demanding self-disclosure and information,' are central to analyzing Bonnie's story, as are those about accounting, explaining, and providing reasons. Other codes preserve events, suggest contexts, and portray viewpoints, such as 'receiving second-hand news,' 'expecting to manage pain,' and 'sounding like a "normal" mom.' Many of the codes are short. They also imply crucial relationships between telling and self, as defined by both self and other. Hence, the codes suggest building categories concerned with telling, disclosing, self, and identity. I placed two codes in parentheses because they are less firmly apparent here than others and represent ideas to look for in further data. Consistent with a grounded theory emphasis on emergence, questions about these codes arise from my reading of the data rather than emanating from an earlier frame applied to them.

Grounded Theory Coding

Grounded theory coding generates the bones of your analysis. Theoretical integration will assemble these bones into a working skeleton. Thus, coding is more than a beginning; it shapes an analytic frame from which you build the analysis.

I lay out coding strategies for developing the frame. Try them. See how they work for you. Grounded theory coding fosters studying action and processes, as you can see in the codes of Bonnie Presley's story.

> Coding is the pivotal link between collecting data and developing an emergent theory to explain these data. Through coding, you *define* what is happening in the data and begin to grapple with what it means.

Coding is the pivotal link between collecting data and developing an emergent theory to explain these data. Through coding, you *define* what is happening in the data and begin to grapple with what it means. The codes take form together as elements of a nascent theory that explains these data and directs further data-gathering. By careful attending to coding, you begin weaving two major threads in the fabric of grounded theory: generalizable theoretical statements that transcend specific times and places and contextual analyses of actions and events.

Grounded theory coding consists of at least two main phases: 1) an initial phase involving naming each word, line, or segment of data followed by 2) a focused, selective phase that uses the most significant or frequent initial codes to sort, synthesize, integrate, and organize large amounts of data. While engaged in initial coding, you mine early data for analytic ideas to pursue in further data collection and analysis. Initial coding entails a close reading of the data as indicated by my codes of Bonnie Presley's story. During initial coding, the goal is to remain open to all possible theoretical directions indicated by your readings of the data. Later, you use focused coding to pinpoint and develop the most salient categories in large batches of data. Theoretical integration begins with focused coding and proceeds through all your subsequent analytic steps.

The actual research you conduct through analyzing your data likely differs–at least somewhat–from what you may have planned earlier in a research or grant proposal. We learn through studying our data. Qualitative coding guides our learning. Through it, we begin to make sense of our data. How we make sense of it shapes the ensuing analysis. Careful attention to coding furthers our attempts to understand acts and accounts, scenes and sentiments, stories and silences from our research participants' view. We want to know what is happening in the setting, in people's lives, and in lines of our recorded data. Hence, we try to understand our participants' standpoints and situations, as well as their actions within the setting.

The logic of grounded theory coding differs from quantitative logic that applies *preconceived* categories or codes to the data. As the example above illustrates, we *create* our codes by defining what we see in the data. Codes emerge as you scrutinize your data and define meanings within it. Through this active coding, you interact with your data again and again and ask many different questions of them. As a result, coding may take you into unforeseen areas and new research questions.

Language plays a crucial role in how and what we code. Most fundamentally, the empirical world does not appear to us in some natural state apart from human experience. Rather we know the empirical world through language and the actions we take toward it. In this sense, no researcher is neutral because

language confers form and meaning on observed realities. Specific use of language reflects views and values. We share one language with colleagues and perhaps another with friends; we attribute meanings to specific terms and hold perspectives. Our codes arise from the languages, meanings, and perspectives through which we learn about the empirical world, including those of our participants as well as our own. Coding impels us to make our participants' language problematic to render an analysis of it. Coding should inspire us to examine hidden assumptions in our own use of language as well as that of our participants.

We *construct* our codes because we are actively naming data–even when we believe our codes form a perfect fit with actions and events in the studied world. We may think our codes capture the empirical reality. Yet it is *our* view: we choose the words that constitute our codes. Thus we define what we see as significant in the data and describe what we think is happening. Coding consists of this initial, shorthand defining and labeling; it results from a grounded theorist's actions and understandings. Nonetheless, the process is interactive. We interact with our participants and subsequently interact with them again many times over through studying their statements and observed actions and re-envisioning the scenes in which we know them. As we define our codes and perhaps later refine them, we try to understand participants' views and actions from their perspectives. These perspectives usually assume much more than what is immediately apparent. We must dig into our data to interpret participants' tacit meanings. Close attention to coding helps us to do that.

Close attention to coding follows the first grounded theory mandate: *Study your emerging data* (Glaser, 1978).

From the beginning, you may sense that the process of coding produces certain tensions–between analytic insights and described events, whether spoken accounts or written observations, between static topics and dynamic processes, and between participants' worlds and professionals' meanings.

Initial Coding

The Logic of Initial Coding

When grounded theorists conduct initial coding, we remain open to exploring whatever theoretical possibilities we can discern in the data. This initial step in coding moves us toward later decisions about defining our core conceptual categories. Through comparing data with data, we learn what our research participants view as problematic and begin to treat it analytically. During initial coding, we ask:

- 'What is this data a study of?' (Glaser, 1978: 57; Glaser & Strauss, 1967)
- What does the data suggest? Pronounce?
- From whose point of view?
- What theoretical category does this specific datum indicate? (Glaser, 1978)

Initial coding should stick closely to the data. Try to see actions in each segment of data rather than applying preexisting categories to the data. Attempt to code

with words that reflect action. At first, invoking a language of action rather than of topics may feel strange. Look closely at actions and, to the degree possible, code data *as* actions. This method of coding curbs our tendencies to make conceptual leaps and to adopt extant theories *before* we have done the necessary analytic work.

Students often believe that they must rely on earlier concepts and invoke them before they begin coding to make their qualitative research legitimate. They make statements like, 'I'm going to use Max Weber's concept of routinization,' or 'My advisor wants me to use Anselm Strauss's concept of "negotiations".' Such approaches preclude ideas from emerging as you code events. The openness of initial coding should spark your thinking and allow new ideas to emerge. Earlier grounded theory rules prescribed conducting initial coding without having preconceived concepts in mind (Glaser, 1978, 1992). I agree with Glaser's approach of keeping initial coding open-ended yet acknowledge that researchers hold prior ideas and skills. As Dey (1999: 251) states, 'There is a difference between an open mind and an empty head.' Try to remain open to seeing what you can learn while coding and where it can take you. In team research, several individuals may code data separately and then compare and combine their different codings.

Initial codes are provisional, comparative, and grounded in the data. They are provisional because you aim to remain open to other analytic possibilities and create codes that best fit the data you have. You progressively follow up on codes that indicate that they fit the data. Then you gather data to explore and fill out these codes.

Initial grounded theory coding can prompt you to see areas in which you lack needed data. Realizing that your data have gaps–or holes–is part of the analytic process. It is inevitable when you adopt an emergent method of conducting research.[3] After all, making 'discoveries' about the worlds you study and pursuing these discoveries to construct an analysis is what grounded theory is about. Such discoveries reflect what you learn and how you conceptualize it. The advantage of grounded theory strategies is that you may learn about gaps and holes in your data from the earliest stages of research. Then you can locate sources of needed data and gather them. Hence, simultaneous data collection and analysis can help you go further and deeper into the research problem as well as engage in developing categories.

Codes are also provisional in the sense that you may reword them to improve the fit. Part of the fit is the degree to which they capture and condense meanings and actions. Compelling codes capture the phenomenon and grab the reader.

Initial Coding Practices

Speed and spontaneity help in initial coding. Working quickly can spark your thinking and spawn a fresh view of the data. Some codes fit the data and grab the reader immediately. You can revise others to improve the fit. My original code of the first line of Bonnie Presley's story above was 'receiving news indirectly.' It condensed the statement but the neutral wording drained the incident

of its intensity and importance. Changing the code to 'receiving second-hand news' suggested the reduced value of the news, implied the receiver's diminished status, and alluded to her angry response.

Comparing incidents of the same order between data spurs you to think analytically about them. Bonnie Presley revealed a reluctance to tell her daughter, delayed in telling her, and imparted difficult news in a matter-of-fact manner. Yet, from time to time, she and Amy talked about their problems in giving and getting news about Bonnie's illness. Because Bonnie no longer had much contact with her own mother, dilemmas of disclosure did not arise with her. No disclosures occurred. Bonnie's grandmother, of whom she was very fond, had partly raised her. Bonnie protected her grandmother from worry by treating her situation lightly and by minimizing the implications of her symptoms. My data included other cases of inter-generational tensions. Several other single women I studied who had no children and few close family ties had conflicted relationships with their aging mothers. As geographical and emotional distance increased, these women correspondingly curtailed sharing their news. From the data and brief descriptions above, avoiding disclosure, delaying disclosure, and controlling information all emerged as salient codes.

Glaser (1978) shows how coding with gerunds helps you detect processes and stick to the data. Think of the difference in imagery between the following gerunds and their noun forms: describing versus description, stating versus statement, and leading versus leader. We gain a strong sense of action and sequence with gerunds. The nouns turn these actions into topics. Staying close to the data and, when possible, starting from the words and actions of your respondents, preserves the fluidity of their experience and gives you new ways of looking at it. These steps encourage you to begin analysis from their perspective. That is the point. If you ignore, gloss over, or leap beyond participants' meanings and actions, your grounded theory will likely reflect an outsider's, rather than an insider's view. Outsiders often import an alien professional language to describe the phenomenon. If your data are thin and if you don't push hard in coding, you may mistake routine rationales for analytic insights. Thus, accepting participants' orchestrated impressions at face value can lead to outsider analyses.

Picking up general terms from an interview such as 'experience' or 'event' and calling them codes tells you little about the participant's meaning or action. If general terms seem significant, qualify them. Make your codes fit the data you have rather than forcing the data to fit them.

A code for coding:

- Remain open
- Stay close to the data
- Keep your codes simple and precise
- Construct short codes
- Preserve actions
- Compare data with data
- Move quickly through the data.

In short, remain open to what the material suggests and stay close to it. Keep your codes short, simple, active and analytic. The first two guidelines above reflect your stance toward coding. The remaining guidelines suggest how to do coding.

Word-by-Word Coding

The size of the unit of data to code matters. Some grounded theorists conduct nuanced coding and move through their data word by word. This approach may be particularly helpful when working with documents or certain types of ephemera, such as Internet data. Word-by-word analysis forces you to attend to images and meanings. You may attend to the structure and flow of words, and how both affect the sense you make of them, as well as their specific content.

Line-by-Line Coding

For many grounded theorists, line-by-line coding is the first step in coding (see Box 3.2). Line-by-line coding means naming each line of your written data (Glaser, 1978). Coding every line may seem like an arbitrary exercise because not every line contains a complete sentence and not every sentence may appear to be important.[4] Nevertheless, it can be an enormously useful tool. Ideas will occur to you that had escaped your attention when reading data for a general thematic analysis.[5]

Line-by-line coding works particularly well with detailed data about fundamental empirical problems or processes whether these data consist of interviews, observations, documents, or ethnographies and autobiographies. For example, if you plan to study how older women who have been full-time homemakers handle divorce, you have identified an area to explore about which you may hear stories in interviews, support groups, and job training programs that take on vivid meanings when studied line by line.

Detailed observations of people, actions, and settings that reveal visibly *telling* and *consequential* scenes and actions lend themselves to line-by-line coding. Generalized observations such as 'the meeting droned on' give you little substance to code.

Fresh data and line-by-line coding prompt you to remain open to the data and to see nuances in it. When you code early in-depth interview data, you gain a close look at what participants say and, likely, struggle with. This type of coding can help you to identify implicit concerns as well as explicit statements. Engaging in line-by-line coding helps you to refocus later interviews. The following flexible strategies help you code:

- Breaking the data up into their component parts or properties
- Defining the actions on which they rest
- Looking for tacit assumptions
- Explicating implicit actions and meanings
- Crystallizing the significance of the points
- Comparing data with data
- Identifying gaps in the data.

By using these strategies flexibly and following leads in your data, coding leads to developing theoretical categories, some of which you may define in your initial codes. Stick with what you define in your data. Build your analysis step-by-step from the ground up without taking off on theoretical flights of fancy. Having a credible amount of data that speaks to your research topic further strengthens the foundation of your study.

Your research participants' actions and statements teach you about their worlds, albeit sometimes in ways they may not anticipate. Studying your data through line-by-line coding sparks new ideas for you to pursue. Hence, the grounded theory method itself contains correctives that reduce the likelihood that researchers merely superimpose their preconceived notions on the data. Line-by-line coding provides an early corrective of this type.

In the examples of line-by-line coding in Box 3.2, my interest in time and self-concept comes through in the first two codes in Excerpt 1. Note how I kept the codes active and close to the data. Initial codes often range widely across a variety of topics. Because even a short statement or excerpt may address several points, it could illustrate several different categories. I could use the excerpt in Box 3.2 to show how avoiding disclosure serves to control identity. I could also use it either to show how a respondent learns that other people see his or her illness as inexplicable or how each day is unpredictable. Having multiple interviews of the same individuals allows me to see how social and emotional isolation begins and progresses.

The logic of 'discovery' becomes evident as you begin to code data. Line-by-line coding forces you to look at the data anew. Compare what you see when you read a set of fieldnotes or an interview as an entire narrative with what you gain when you do word-by-word, line-by-line, or incident-by-incident coding on the same document. Entire narratives may net several major themes. Word-by-word, line-by-line, segment-by-segment, and incident-by-incident coding may generate a range of ideas and information. Therefore, you 'discover' ideas on which you can build.

Initial codes help you to separate data into categories and to see processes. Line-by-line coding frees you from becoming so immersed in your respondents' worldviews that you accept them without question. Then you fail to look at your data critically and analytically. Being critical about your data does not necessarily mean being critical of your research participants. Instead, being critical forces asking *yourself* questions about your data. These questions help you to see actions and to identify significant processes. Such questions include:

- What process(es) is at issue here? How can I define it?
- How does this process develop?
- How does the research participant(s) act while involved in this process?
- What does the research participant(s) profess to think and feel while involved in this process? What might his or her observed behavior indicate?
- When, why, and how does the process change?
- What are the consequences of the process?

Through coding each line of data, you gain insights about what kinds of data to collect next. Thus, you distill data and direct further inquiry early in the

BOX 3.2 INITIAL CODING: LINE-BY-LINE CODING

	Excerpt 1 *Christine Danforth, age 37, lupus erythematosus, Sjögren's syndrome, back injuries*
	Lupus erythematosus is a systemic, inflammatory autoimmune disease of the connective tissue that affects vital organs as well as joints, muscles, and nerves. Sjögren's syndrome is a related autoimmune inflammatory disease characterized by dry mucous membranes of the eyes and mouth.
Shifting symptoms, having inconsistent days Interpreting images of self given by others Avoiding disclosure Predicting rejection	If you have lupus, I mean one day it's my liver; one day it's my joints; one day it's my head, and it's like people really think you're a hypochondriac if you keep complaining about different ailments ... It's like you don't want to say anything because people are going to start thinking, you know, 'God, don't go near her, all she is–is complaining about this.' And I think that's why I never say anything because I feel like everything I have is related one way or another to the lupus but most of the people don't know I have lupus, and even those that do are not going to believe that ten different ailments are the same thing. And I don't want anybody saying, you know, [that] they don't want to come around me because I complain.
Keeping others unaware Seeing symptoms as connected Having others unaware Anticipating disbelief Controlling others' views Avoiding stigma Assessing potential losses and risks of disclosing	
	Excerpt 2 *Joyce Marshall, age 60, minor heart condition, recent small CVA (stroke)*
	In her case, the stroke left her with weakness, fatigue, and slowed responses when tired.
Meaning of the CVA	I have to see it [her CVA] as a warning. I can't let myself get so anxious. I have to live one day at a time.
Feeling forced to live one day at a time	
Having a worried past	I've been so worried about John [her husband who had had life-threatening heart attacks and lost his job three years before retirement] and preparing to get a job [her first in 38 years] ... It's just so hard with all this stress ... to concentrate on what I can do today. I always used to look to the future. I can't now; it upsets me too much. I have to live one day at a time now or else there may not be any me.
Earlier losses Difficulty in living one day at a time; concentrating on today Giving up future orientation Managing emotions through living one day at a time Reducing life-threatening risk	

data collection. Line-by-line coding gives you leads to pursue. If, for example, you identify an important process in your fifteenth interview, then you can return to earlier respondents and see if that process explains events and experiences in their lives. If not, you can seek new respondents who can illuminate this process. Hence, your data collection becomes more focused, as does your coding.

Coding Incident to Incident

Whether or not you conduct line-by-line coding depends on the type of data you have collected, their level of abstraction, the stage of the research process, and your purpose for collecting these data. Grounded theorists often conduct a close cousin of line-by-line coding through a comparative study of incidents. Here you compare incident with incident, then as your ideas take hold, compare incidents to your conceptualization of incidents coded earlier. That way you can identify properties of your emerging concept.

A similar logic applies to observational data. Making comparisons between incidents likely works better than word-by-word or line-by-line coding, in part because the fieldnotes already consist of your own words (see, for example, Charmaz & Mitchell, 2001). Compare incident to incident. Concrete, behavioristic descriptions of people's mundane actions may not be amenable to line-by-line coding, particularly when you observed a scene but do not have a sense of its context, its participants, and did not interact with them. Students often think observing behavior in public places is the easiest type of qualitative research to conduct. Not so. Both the researcher's data and analytic approach make a difference. Few novices have the eye and ear to record nuances of action and interaction. More likely, they record concrete behaviors in a general way and gradually learn to make more acute observations.

Still, detailed observations alone do not guarantee creating an insightful theoretical analysis although they may generate excellent description. The mode of analysis matters. Comparative methods help you to see and make sense of observations in new, analytic ways. Conducting a line-by-line coding of one observation after another of people's actions in a public place may not spark fresh ideas. Instead, making comparisons between observations gives you clues to follow if not immediate ideas. If the people you study bring you into their world, for example, you may record all kinds of incidents in anecdotes, conversations, and observations in your fieldnotes that abound with meaning. You may see first-hand how your participants manage daily life without them telling you–and you may learn much more.

The more unproblematic–that is, routine, familiar, and ordinary–observed events seem to you, the more problematic creating an original conceptual analysis of them will be. Breaking through the ordinariness of routine events takes effort. To gain analytic insights from observations of routine actions in ordinary settings, first compare and code similar events. Then you may define subtle patterns and significant processes. Later, comparing *dissimilar* events may give you further insights.

Using Comparative Methods

Whatever unit of data you begin coding in grounded theory, you use *'constant comparative methods'* (Glaser & Strauss, 1967) to establish analytic distinctions–and thus make comparisons at each level of analytic work. At first, you compare data with data to find similarities and differences. For example, compare interview statements and incidents within the same interview and compare statements and incidents in different interviews. Making sequential comparisons helps. Compare data in earlier and later interviews of the same individual(s) or compare observations of events at different times and places. When you conduct observations of a routine activity, compare what happens on one day with the same activity on subsequent days.

If your codes define another view of a process, action or belief than your respondent(s) hold, note that. Your observations and ideas do matter. Do not dismiss your own ideas if they do not mirror the data. Your ideas may rest on covert meanings and actions that have not entirely surfaced yet. Such intuitions form another set of ideas to check. Our task is to make analytic sense of the material, which may challenge taken-for-granted understandings.

What you see in your data relies in part upon your prior perspectives. Rather than seeing your perspectives as truth, try to see them as representing one view among many. That way, you may gain more awareness of the concepts that you employ and might impose on your data. To illustrate, you might already possess a repertoire of psychological concepts that you ordinarily invoke to understand behavior. Invoking these concepts in your codes can lead you to prejudge what is happening. Try to avoid assuming that respondents, for example, repress or deny significant 'facts' about their lives. Instead, look for how they understand their situations before you judge their attitudes and actions through your own assumptions. Seeing the world through their eyes and understanding the logic of their experience brings you fresh insights. Afterwards, if you still enlist disciplinary terms as codes, you will use them more consciously rather than automatically. Thus, you can elect to use only those terms that fit your data.

Advantages of Initial Coding

From the start, careful word-by-word, line-by-line, incident-by-incident coding moves you toward fulfilling two criteria for completing a grounded theory analysis: fit and relevance. Your study fits the empirical world when you have constructed codes and developed them into categories that crystallize participants' experience. It has relevance when you offer an incisive analytic framework that interprets what is happening and makes relationships between implicit processes and structures visible.

Careful coding also helps you to refrain from imputing your motives, fears, or unresolved personal issues to your respondents and to your collected data. Some years ago, a young man in my seminar conducted research on adaptation to disability. He had become paraplegic himself when he was hit by a car while bicycling. Stories of courage, hope, and innovation filled his ten in-depth interviews. Narratives of grief, anger, and loss permeated his analysis of them. After I noted that his analysis did not reflect his collected material, he realized how

his feelings had colored his perceptions of other people's disabilities. His was an important realization. However, he might have arrived at it before he handed in his paper had he done more assiduous coding. Line-by-line coding might have changed his ideas about his date early in the analysis.

Coding forces you to think about the material in new ways that may differ from your research participants' interpretations. Your analytic eye and disciplinary background lead you to look at their statements and actions in ways that may not have occurred to them. By studying the data, you may make fundamental processes explicit, render hidden assumptions visible, and give participants new insights. Thomas (1993) says that a researcher must take the familiar, routine, and mundane and make it unfamiliar and new. Think of seeing a once-familiar landscape with a fresh eye after a long absence. You see familiar landmarks with acuity unlike days past when they blurred together. Word-by-word and line-by-line coding help you to see the familiar in new light. Incident coding aids you in discovering patterns and contrasts. You may gain surprising insights about how people's actions fit together or come into conflict. You also gain distance from your preconceptions and your participants' taken-for-granted assumptions about the material so that you *can* see it in new light.

In Vivo Codes

Grounded theorists generally refer to codes of participants' special terms as *in vivo* codes. Their specialized terms provide a useful analytic point of departure. *In vivo* codes help us to preserve participants' meanings of their views and actions in the coding itself. *Pay attention to language while you are coding. In vivo* codes serve as symbolic markers of participants' speech and meanings. Whether or not they provide useful codes in the later more integrated analysis depends on how you treat them analytically. Like any other code, they need to be subjected to comparative and analytic treatment. Although the terms may be catchy, *in vivo* codes do not stand on their own in a robust grounded theory; these codes need to be integrated into the theory. When you scrutinize them carefully, three kinds of *in vivo* codes prove to be useful:

- Those general terms everyone 'knows' that flag condensed but significant meanings
- A participant's innovative term that captures meanings or experience
- Insider shorthand terms specific to a particular group that reflect their perspective.

In vivo codes that condense meanings consist of widely used terms that participants assume everyone shares. In contrast, take participants' usage as problematic rather than reproducing it. Hence, we look for their implicit meanings and attend to how they construct and act upon these meanings. In doing so, we can ask, what analytic category(ies) does this code suggest? Unpacking such terms not only gives you a great opportunity to understand implicit meanings and actions but also to make comparisons between data and with your emerging categories.

Today, everyone knows what the general term 'battered woman' means; however, certain groups assume specific meanings when they use the term. Donileen Loseke (1992) discovered that claims-makers' use of the term depicted a particular set of characteristics that did not fit all women who suffered physical abuse. For claims-makers, a battered woman meant an economically and emotionally dependent mother who suffered repeated, escalating physical abuse, had low self-esteem and poor coping skills, could not rely on informal help or formal services, and had no place to go. These claims-makers then acted on their meanings when deciding who would receive services and what these services should include. An older, affluent woman without children would not fit their definition, despite having been beaten.

Some *in vivo* codes simultaneously reflect condensed meanings of a general term and reveal an individual's fresh perspective. After suffering a sudden onset of a serious chronic condition, one man said he intended to pursue 'making a comeback' (Charmaz, 1973). By borrowing a term from once-successful celebrities, he defined his stance toward dealing with chronic illness. Other participants' actions and statements indicated that they shared this stance, although they did not invoke this vivid term.

In vivo codes are characteristic of social worlds and organizational settings. For example, Calvin Morrill's (1995: 263–268) glossary of executives' terms in one corporation included both general terms and specific labels that no doubt furthered his understanding of how they dealt with conflict. Executives imbued some terms, such as 'bozo,' 'roadblock,' or 'jumping ship,' with meanings that echoed ordinary parlance, although many terms assumed specific meanings within the organization and evoked metaphors of combat, violence, and violation. Morrill includes among them:

BLACK KNIGHT	An executive who often engages in covert action against opponents, does not support his intra-departmental colleagues in disputes ... ; (in take-over imagery, *black knight* refers to an unfriendly acquirer from the perspective of an acquired firm). (p. 263)
FLYING LOW	Not confronting an offender with longstanding grievances against their behavior. (p. 265)
RAPE	An executive's allowing himself or herself to be publicly criticized by another without *calling out* the challenger. (p. 266)
SMALL BURSTS OF FIRE	Short public criticisms of a colleague delivered in rapid succession. (p. 267)
VAPORIZING	Terminating an executive from the company or creating the conditions under which an executive resigns from the corporation. (p. 267)

At organizational or collective levels of analysis, *in vivo* codes reflect assumptions, actions, and imperatives that frame action. Studying these codes and exploring leads in them allows you to develop a deeper understanding of what is happening and what it means. Such codes anchor your analysis in your research participants' worlds. They offer clues about the relative congruence between your interpretation of participants' meanings and actions and their overt statements and actions. *In vivo* codes can provide a crucial check on whether you have grasped what is significant. Elijah Anderson (2003) speaks to this point in his ethnographic memoir of his study (1976) of African-American men who hung out on a Chicago street corner. Anderson discovered three groups: 'respectables,' 'non-respectables,' and 'near-respectables.' He related these categories to his teacher, Howard Becker, who asked him what the men called themselves. Anderson reviewed his data and realized that the men's terms of 'regular,' 'hoodlum,' and 'winehead,' stood out. After invoking their terms and clarifying what the men meant by them, Anderson stated that he experienced a dramatic increase in his understanding of their worlds. From a grounded theory standpoint, for example, it would be fascinating to explicate the process of how men become defined as belonging to one category or another, who designates and enforces the categories, and how these categories make actions predictable.

In each study you conduct, participants will word or write things in ways that crystallize and condense meanings. Hearing and seeing their words anew allows you to explore their meanings and to understand their actions through coding and subsequent data collection. Pursue telling terms. One young doctor with severe diabetes explained himself as being 'supernormal' (Charmaz, 1973, 1987). As our conversation unfolded, his meaning of supernormal became clear. Not only did he intend to manage being a physician without his condition deterring him, he also aimed to excel beyond his peers. His hopes and plans symbolized identity goals in social life that transcended psychological predilections. Once I grasped the idea of pursuing supernormal identity goals, I saw this process reflected in other participants' actions and stated intentions. Similarly, other *in vivo* codes emerged as I heard many people advocate 'taking one day at a time,' and listened to their stories of having 'good days' and 'bad days.' Subsequently, I sought the condensed meanings and actions that these terms covered and coded for them.

Focused Coding

Focused coding is the second major phase in coding. These codes are more directed, selective, and conceptual than word-by-word, line-by-line, and incident-by-incident coding (Glaser, 1978). After you have established some strong analytic directions through your initial line-by-line coding, you can begin focused coding to synthesize and explain larger segments of data. Focused coding means using the most significant and/or frequent earlier codes to sift through large amounts of data. One goal is to determine the adequacy of those codes. Focused coding requires decisions

> Focused coding means using the most significant and/or frequent earlier codes to sift through large amounts of data. Focused coding requires decisions about which initial codes make the most analytic sense to categorize your data incisively and completely.

about which initial codes make the most analytic sense to categorize your data incisively and completely.

But moving to focused coding is not entirely a linear process. Some respondents or events will make explicit what was implicit in earlier statements or events. An 'Aha! Now I understand,' experience may prompt you to study your earlier data afresh. Then you may return to earlier respondents and explore topics that had been glossed over, or that may have been too implicit to discern initially or unstated.

BOX 3.3 FOCUSED CODING

Excerpt 1 *Christine Danforth, age 37, lupus erythematosus, Sjögren's syndrome, back injuries*

Avoiding disclosure / Assessing potential losses and risks of disclosing

If you have lupus, I mean one day it's my liver; one day it's my joints; one day it's my head, and it's like people really think you're a hypochondriac if you keep complaining about different ailments ... It's like you don't want to say anything because people are going to start thinking, you know, 'God, don't go near her, all she is–is complaining about this.' And I think that's why I never say anything because I feel like everything I have is related one way or another to the lupus but most of the people don't know I have lupus, and even those that do are not going to believe that ten different ailments are the same thing. And I don't want anybody saying, you know, [that] they don't want to come around me because I complain.

Excerpt 2 *Joyce Marshall, age 60, minor heart condition, recent small CVA (stroke)*

Feeling forced to live one day at a time / Concentrating on today / Giving up future orientation / Managing emotions / Reducing life-threatening risk

I have to see it [her CVA] as a warning. I can't let myself get so anxious. I have to live one day at a time.

I've been so worried about John [her husband who had had life-threatening heart attacks and lost his job three years before retirement] and preparing to get a job [her first in 38 years] ... It's just so hard with all this stress ... to concentrate on what I can do today. I always used to look to the future. I can't now; it upsets me too much. I have to live one day at a time now or else there may not be any me.

The strength of grounded theory coding derives from this concentrated, active involvement in the process. You act upon your data rather than passively read them. Through your actions, new threads for analysis become apparent. Events, interactions, and perspectives come into analytic purview that you had not thought of before. Focused coding checks your preconceptions about the topic.

In the first excerpt in Box 3.3, I selected the codes 'avoiding disclosure' and 'assessing potential losses and risks of disclosing' to capture, synthesize, and understand the main themes in the statement. In the second, the following codes were most useful: 'feeling forced to live one day at a time,' 'concentrating on today,' 'giving up future orientation,' 'managing emotions,' and 'reducing life-threatening risk.' Again, I tried to keep the codes active and close to the data. Through focused coding, you can move across interviews and observations and compare people's experiences, actions, and interpretations. Note how the codes condense data and provide a handle on them.

Consistent with the logic of grounded theory, coding is an emergent process. Unexpected ideas emerge. They can keep emerging. After you code a body of data, compare your codes and data with each other. A telling code that you constructed to fit one incident or statement might illuminate another. An earlier incident may alert you to see a subsequent one with incisiveness. I had witnessed several tense moments with couples during which spouses declared that their partner's disabilities robbed him or her of former competencies.

Consider the following fieldnotes from an early interview with Andrei, a retired college professor, and his wife, Natasha, both of whom had chronic illnesses:

> I asked [Andrei], 'Did you keep up with professional work after you retired?' He said: 'I used to teach extension courses but due to the budget and that governor, there isn't any money for extension courses.' She [Natasha] cut in [to me], 'Andrei used to be an extremely successful speaker; partly his enthusiasm, partly his articulateness, but with the speech problems, he can't do it ...' [He said, slowly and painfully] 'The schools don't have any money ... I can't speak very well.'
>
> I felt desperately sorry for him at this point. Whether or not both factors were at play at the point when they stopped calling him for extension teaching, this was a terrible moment for him when she said it. Regardless of the real reason, at this precise moment knowing *what* she thought of his deteriorating competence was critical to him. Participating in this short sequence was like watching someone who was observing his own identity crumbling away – it was painful both for him and for me, although I got the impression that she was so caught up in her perceptions of accuracy that she didn't actually see how it defaced him. ... Acknowledging that he can't speak very well was said like an admission of guilt or inferiority that was previously hidden from view. (Charmaz, 1983: 119–120)

From such early observations, I developed the code of 'identifying moment.' In each case, the judgment imparted a shocking image of whom the ill person had become. Such disquieting views proclaimed negative changes and underscored their permanence. The code 'identifying moment' alerted me to other brief interactions in which someone conferred a significant identity on a person with chronic illness. One example occurred some years later when I entered a

care home that primarily served impoverished elders. The assistant at the desk said that her supervisor had not informed her that I was coming to talk with residents (as had been arranged). Six elders in wheelchairs were lined up against the wall and one middle-aged woman was walking toward the desk. The people in wheelchairs perked up and regarded me with interest, as is common in institutions where few residents have visitors. Without looking up, the assistant nodded toward the middle-aged woman and said, 'You can talk to Mary there; she's one of the smart ones and there aren't many of them.' At this pronouncement, six heads in wheelchairs drooped in unison. Mary looked proud to be chosen. I realized that I had just witnessed another identifying moment–a positive one for Mary but a negative one for the other residents.

Through comparing data to data, we develop the focused code. Then we compare data to these codes, which helps to refine them. In the first example, I compared situations in which participants had freely discussed the disability in question before, with those in which they had not. Before the incident when Andrei acknowledged his speech difficulties, his physician had told me that Andrei's impaired speech was never openly discussed. I also compared these incidents for their intensity and impact. At first, the code only represented negative identifying moments. As I obtained more data, I found and defined positive identifying moments. 'Identifying moments' began as a code, which I developed as a category (Charmaz, 1991a). Because the notion of identifying moments resonates with many experiences, Will van den Hoonaard (1997) treats it as a sensitizing concept for other researchers to use as a starting point.

Axial Coding

Strauss and Corbin (1990, 1998; Strauss, 1987) present a third type of coding, axial coding, to relate categories to subcategories. Axial coding specifies the properties and dimensions of a category.

Strauss (1987: 64) views axial coding as building 'a dense texture of relationships around the "axis" of a category.' Thus, axial coding follows the development of a major category, although it may be in an early stage of development. The purposes of axial coding are to sort, synthesize, and organize large amounts of data and reassemble them in new ways after open coding (Creswell, 1998).

> Axial coding relates categories to subcategories, specifies the properties and dimensions of a category, and reassembles the data you have fractured during initial coding to give coherence to the emerging analysis.

Initial coding fractures data into separate prices and distinct codes. Axial coding is Strauss and Corbin's (1998) strategy for bringing data back together again in a coherent whole. According to Strauss and Corbin (p. 125), axial coding answers questions such as 'when, where, why, who, how, and with what consequences.' With these questions, a researcher can describe the studied experience more fully, although Corbin and Strauss contend that linking

relationships between categories occurs on a conceptual rather than descriptive level. For them, analyzing data means converting text into concepts, which seems to be the intent of Strauss's and Corbin's use of axial coding. These concepts specify the dimensions of a larger category. Axial coding aims to link categories with subcategories, and asks how they are related. Clarke views axial coding as elaborating a category and uses diagramming to integrate relevant categories.[6] For her, an integrative diagram aims to link categories with categories to form a substantive theory of action.

While engaged in axial coding, Strauss and Corbin apply a set of scientific terms to make links between categories visible. They group participants' statements into components of an organizing scheme to answer their questions above. In one such organizing scheme, Strauss and Corbin include: 1) *conditions*, the circumstances or situations that form the structure of the studied phenomena; 2) *actions/interactions*, participants' routine or strategic responses to issues, events, or problems; and 3) *consequences*, outcomes of actions/interactions. Strauss and Corbin use conditions to answer the why, where, how come, and when questions (p. 128). Actions/interactions answer by whom and how questions. Consequences answer questions of 'what happens' because of these actions/interactions.

Axial coding provides a frame for researchers to apply. The frame may extend or limit your vision, depending on your subject matter and ability to tolerate ambiguity. Students who prefer to work with a preset structure will welcome having a frame. Those who prefer simple, flexible guidelines–and can tolerate ambiguity–do not need to do axial coding. They can follow the leads that they define in their empirical materials.

Although I have not used axial coding according to Strauss and Corbin's formal procedures, I have developed subcategories of a category and showed the links between them as I learned about the experiences the categories represent. The subsequent categories, subcategories, and links reflect how I made sense of the data.

The earlier coding examples of Bonnie Presley and Christine Danforth's interviews indicate that telling other people about having a chronic illness poses emotional and interactional dilemmas. Such dilemmas arose in many interviews; I had not planned to study them. Not surprisingly, the first two categories that I saw in early interviews were disclosing illness and avoiding disclosure. I outlined their respective properties through comparing data with data of the same kind of experience or event. The apparent pain in participants' stories led me to view 'disclosing' as revealing and often risky. Bonnie Presley's risks included exacerbating a medical crisis. Many other people risked making themselves emotionally vulnerable and having uncontrollable feelings. Disclosing was not a neutral form of talking.

Next, I reexamined the data I had coded during initial coding. Participants dealt with information about themselves both by avoiding disclosure of illness and by telling people about it; however, some forms of telling lacked control and sometimes not telling at all occurred when participants felt overwhelmed. When participants' lacked control in telling, they exposed themselves by blurting out their concerns instead of managing and metering self-revelations.

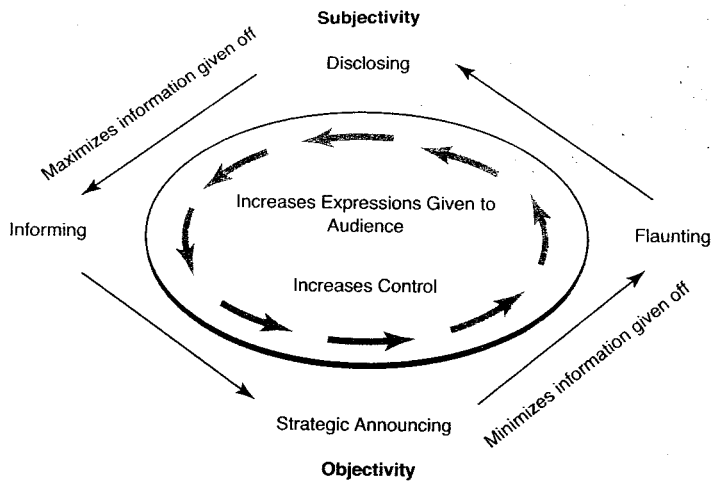

FIGURE 3.1 Forms of telling

Subsequently, I coded for the range between spontaneous statements and staged pronouncements. I linked forms of telling explicitly to the relative absence or presence of participants' control in relaying information and the extent to which they invoked explicit strategies. After discovering that people invoked different forms of telling, I then looked more closely at the following:

- Biographical and interactional contexts of their telling
- Social and experiential conditions affecting whom various participants told
- Participants' stated intentions for telling
- What participants told these individuals
- How participants told them.

I coded for if, when, how, and why participants changed their earlier forms of telling. These strategies may lead to charting causes and conditions of the observed phenomenon. In my analysis of forms of telling (see Figure 3.1), studying these data led to seeing that the participant's subjective stake in telling exceeded what a researcher could plot along a simple continuum. Rather, subjectivity and objectivity meet when participants flaunt illness. It became apparent that some individuals flaunt illness when it has caused them unresolved problems of self-acceptance and acceptance by others.

No explicit frame guided my analytic constructions of participants' accounts and experiences or elicited the emphasis. Although axial coding may help researchers to explore their data, it encourages them to *apply* an analytic frame to the data. In that sense, relying on axial coding may limit what and how researchers learn about their studied worlds and, thus, restricts the codes they construct.

Whether axial coding helps or hinders remains a question (see Kelle, 2005). Whether and to what extent it offers a more effective technique than careful

comparisons remains debatable. At best, axial coding helps to clarify and to extend the analytic power of your emerging ideas. At worst, it casts a technological overlay on the data–and perhaps on your final analysis. Although intended to obtain a more complete grasp of the studied phenomena, axial coding can make grounded theory cumbersome (Robrecht, 1995).

Theoretical Coding

Theoretical coding is a sophisticated level of coding that follows the codes you have selected during focused coding. Glaser (1978: 72) introduced theoretical codes as conceptualizing 'how the substantive codes may relate to each other as hypotheses to be integrated into a theory.' In short, theoretical codes specify possible relationships between categories you have developed in your focused coding. Glaser (1992) argues that these codes preclude a need for axial coding because they 'weave the fractured story back together' (Glaser, 1978: 72). Theoretical codes are integrative; they lend form to the focused codes you have collected. These codes may help you tell an analytic story that has coherence. Hence, these codes not only conceptualize how your substantive codes are related, but also move your analytic story in a theoretical direction.

Glaser (1978) presents a series of 18 theoretical coding families that include analytic categories such as his 'Six Cs: Causes, Contexts, Contingencies, Consequences, Covariances, and Conditions' (p. 74), 'degree,' 'dimension,' 'interactive,' 'theoretical,' and 'type' coding families as well as ones that derive from major concepts such as 'identity–self,' 'means–goals,' 'cultural,' and 'consensus' families. Several of Glaser's coding families indicate a specific analytic category but merge conceptual distinctions. For example, the 'unit' family includes the following structural units: group, family organizational, aggregate, territorial, societal, status and role units. Glaser also includes situations, social worlds and social contexts, which certainly may serve as units of analysis but connote emergent, rather than structural properties. In *Doing Grounded Theory* (1998), Glaser enlarges on several earlier coding families and extends the list to include more coding families such as: the 'paired opposite,' 'representation,' 'scale,' 'random walk,' 'structural-functional,' and 'unit identity' families.

If you use them skillfully, theoretical codes may hone your work with a sharp analytic edge. They can add precision and clarity–as long as they fit your data and substantive analysis. They can aid in making your analysis coherent and comprehensible. Depending on the data you have and on what you learn about them, you may find that your analysis takes into account several coding families. For example, you may clarify the general context and specific conditions in which a particular phenomenon is evident. You may be able to specify the conditions under which it changes and to outline its consequences. You might learn its temporal and structural orderings and discover participants' strategies for dealing with them. If you understand the temporal ordering, you likely include an analysis of process. Thus, despite not delving into substance, this short example alone brings in the following analytic coding families: the 'Six Cs,' 'temporal ordering,' 'ordering' (Glaser includes structural ordering here, see