

Qualitative Data Analysis: Common Phases, Strategic Differences

Ian Baptiste

Key words:
qualitative
research, research
methodology,
qualitative data
analysis

Abstract: This paper lays out an analytic framework to help rookie qualitative researchers recognize and appreciate common features of qualitative data analysis (QDA) while giving due consideration to strategic differences resulting from differences in expertise, context, and philosophy. The paper does not identify or illustrate specific QDA strategies. Rather, it raises questions the responsible analyst might consider at each phase of the process. I argue that all QDA (regardless of methodological or disciplinary orientation) comprise four interrelated phases: defining the analysis, classifying data, making connections between data, and conveying the message(s). This paper discusses the first three phases.

Key words:

[1. Introduction](#)

[2. QDA: Four Phases, Numerous Strategies](#)

[2.1 Phase 1: Defining the analysis](#)

[2.1.1 Ontology](#)

[2.1.2 A word about computer programs](#)

[2.1.3 Epistemology](#)

[2.1.4 Axiology](#)

[2.1.5 Causality](#)

[2.2 Phase 2: Classifying data](#)

[2.2.1 General questions guiding choice of analytic strategies](#)

[2.2.2 Questions to consider when tagging and labeling data](#)

[2.2.3 Questions to consider when developing and refining categories](#)

[2.3 Phase 3: Making connections—Constructing stories and theories](#)

[2.3.1 Questions to consider when constructing stories or theories](#)

[3. Conclusion](#)

[References](#)

[Author](#)

[Citation](#)

1. Introduction

Most qualitative research textbooks—whether they are general overviews such as DENZIN and LINCOLN (2000), MARSHALL and ROSSMAN (1995), MERRIAM (1998) or PATTON (1990); discipline/method-specific ones such as CARSPECKEN (1996), MOUSTAKAS (1994), STRAUSS and CORBIN (1998), STAKE (1995) or VAN MANEN (1990)—devote at least one chapter to qualitative data analysis (QDA). There are also texts devoted entirely to QDA such as COFFEY and ATKINSON (1996), DEY (1993), FELDMAN (1995), MILES and HUBERMAN (1994), SILVERMAN (2001) and WALCOTT (1994). The more general texts tend to focus

predominantly on what's common about QDA, muffling disciplinary and methodological differences. CRESWELL (1998) has written a general text that addresses strategic differences existing between and among five popular methods—biographical life histories, phenomenology, grounded theory, ethnography and case study. However, in his treatment of each method, CRESWELL shows a definite preference to a particular philosophical school over others. For instance, GLASER (1992)—a co-originator of grounded theory—differs philosophically with STRAUSS his co-constructor (see STRAUSS & CORBIN, 1990) on what is grounded theory and how data analysis should be conducted. Yet in his discussion of QDA in grounded theory, CRESWELL takes the STRAUSSIAN position, ignoring Glaser's objections (CRESWELL, 1998, chap. 8). Likewise, in his treatment of phenomenology CRESWELL leans more toward HUSSERL'S transcendentalism (HUSSERL, 1931; MOUSTAKAS, 1994) than to HEIDEGGER'S existentialism (HEIDEGGER, 1962, VAN MANEN, 1990). [1]

Understandably, the discipline/method-specific texts (CARSPECKEN, 1996; MOUSTAKAS, 1994; STRAUSS & CORBIN, 1998; STAKE, 1995; VAN MANEN, 1990) tend to focus on strategies specific to their particular discipline or method. Regarding those texts that are devoted entirely to QDA, some, such as MILES and HUBERMAN (1994) approach the task in much the same way as the general qualitative research texts—emphasizing what's common about QDA. Others (such as COFFEY & ATKINSON, 1996; or FELDMAN, 1995) highlight strategic differences resulting from disciplinary or methodological differences. However, constrained by space, these works are limited in the range of strategies they could adequately illustrate. In this paper I try to combine some of the virtues of both the general texts and the more discipline-specific ones. With the general texts, I try to emphasize the common features (phases) of QDA. But within that general framework, I suggest a broad range of strategic differences. I am able to do so, because I do not try to illustrate any particular strategy. Instead, within my discussion of each phase of QDA, I raise a series of questions to help my readers tease out strategic differences. It is in this sense (of posing questions) that this paper is analytic. It does not prescribe (or describe) specific analytic strategies. Instead, it provides a framework to help the rookie analyst make contextually-grounded, analytic choices. In developing this framework, I do not pretend to be neutral. I am decidedly constructivist in my philosophical orientation. So expect the questions I pose to be heavily influenced by that orientation (BERGER & LUCKMANN, 1967; LINCOLN & GUBA, 1985; VYGOTSKY, 1978). [2]

2. QDA: Four Phases, Numerous Strategies

In this paper, the term *phases* refer to major (distinguishable) intellectual tasks or goals. *Strategies*, on the other hand, refer to the way(s) we try to achieve our goals. For example, I regard classification as a major task or goal of any and all QDA. However how researchers go about classifying, in other words, the strategies each researcher uses may differ considerably. Phases do not occur as discrete, sequential, hierarchical steps; they are iterative, interactive and non-linear. Stages imply step-like processes—occurring one after the other. I also distinguish between strategies and tactics. Strategies are the overall plan or

approach we employ. In classifying data, for instance, one might use an intuitive or explicit strategy. *Tactics* are instances of a strategy; specific procedures employed within given strategies; what we do with what we have (ALINSKY, 1971, p.126). An explicit classification tactic might be to color code pieces of data; another might be tagging paragraphs using a QDA computer program. [3]

Regardless of the field of practice, disciplinary allegiance, research purposes or designs, I believe that there are four phases to all QDA: defining the analysis, classifying data, making connections between and among categories of data, and conveying the message/write-up. In this paper I discuss the first three. I have chosen not to discuss the fourth phase (conveying the message) for the following reasons: 1) my students and I are far less confused by what we read concerning this phase than we are about what we read concerning the other three phases; 2) there seems to be less equivocation in the literature concerning this fourth phase; 3) I frankly do not think that I have anything to say (at this time) that might improve our understanding of this phase. [4]

Every analyst, whether tacitly or overtly, takes into account three sets of considerations during any QDA process, they are: a) philosophical; b) contextual; and c) design. Philosophical considerations refer to the analyst's systems of values and beliefs concerning research in general. They also include the analyst's interests, ideological stance, and theoretical positions regarding the particular phenomena or issue(s) under investigation. For example, Africentric feminists are likely to interpret data quite differently from analysts who subscribe to neo-liberal or post-modern ideologies. [5]

I believe that for each analyst, these values, beliefs and interests are fairly stable over time and space. They determine what the particular analyst considers desirable; and they form the outer limits of what the particular analyst considers *theoretically* possible. For instance, if an analyst believes that is neither possible nor desirable to be objective, that belief will significantly govern how s/he conducts her/his analysis. That belief will also significantly influence how the analyst defines and establishes quality and credibility in her/his research. [6]

Design considerations refer to specific requirements imposed upon the analytic process by the very nature of the research questions themselves. These are considerations about research purposes and methods—including corresponding strategies and tactics. I concur with MERTON (1967), that at the broadest conceptual level, research is conducted either to generate hypotheses or to test them. And at a more intermediary conceptual level, the purposes of research might be divided into such overlapping categories as: descriptive, explanatory, exploratory, predictive, participatory, and so on. Within philosophical and contextual constraints (see below), each of these purposes dictates a different method or combination of methods. The analyst who is concerned with exploring and deeply understanding a phenomenon will employ a different combination of methods and corresponding strategies to the analyst who is primarily concerned with explanations or predictions. Popular examples of differing qualitative research methods include phenomenology, ethnography and grounded theory.

Objectivist and subjectivist orientations exist within every research method. For instance, in phenomenology there is the objectivist orientation evident in HUSSERL'S transcendental phenomenology, and the more subjectivist orientation evident HEIDEGGER'S hermeneutic or existential phenomenology (COHEN & OMERY, 1994; RAY, 1994, VAN MANEN, 1990). Similar objectivist and subjectivist orientations are found in grounded theory (CHARMAZ, 2000) and in ethnography (FETTERMAN, 1998; THOMAS, 1993). [7]

Contextual considerations refer to what's feasible given the analyst's knowledge and skills, resources, power and influence. Two analysts may hold similar philosophical and ideological positions. However, what each is able to do in a specific research context depends on their respective knowledge and skills, resources, power and influence. A poor, working-class doctoral student (without scholarship or other financial aid) is simply unable to do what her rich, well-resourced counterpart can do (all other things being equal). [8]

Let me illustrate how these three sets of considerations might play out in QDA decisions. Whether an analyst chooses, for example, to employ classical content analysis (BERELSON, 1952) or semiotic analysis (FELDMAN, 1995; MANNING, 1987) depends first, on philosophical considerations: epistemological assumptions the analyst makes about how meaning is extracted from texts; plus the analyst's ideological and theoretical stance regarding the particular phenomenon or issue under consideration. Second, the analyst's choice of strategy depends on design considerations: what kinds of purposes do the research questions evoke; do the purposes elicit the generation of hypotheses or do they evoke hypotheses testing; do the research questions demand description, explanation, prediction, understanding, or change in and of the phenomena under investigation; given the research questions and purposes, which of the two analytic strategy is likely to yield the best results? Third, the analyst's choice depends on contextual considerations: how skilled is the analyst in either of the two analytic strategies; given the analyst's resources (including time) and power how extensive and intensive can she employ the analytic strategy she has selected? The foregoing are the kinds of considerations every analyst makes in order to choose a particular analytic strategy. In many situations, however, these considerations are tacitly made. The suggestions and questions I present below will hopefully help analysts become more explicit (and consequently more trustworthy and reliable) in the analytic choices they make. [9]

The variety of QDA strategies that currently exist in the literature, springs, I believe, from the numerous permutations and combinations of positions that are possible, given these three sets of considerations—philosophical, design, and contextual. Any attempt, therefore, to provide a comprehensive list of possible QDA strategies would, it seems to me, be woefully inadequate and impractical. A more prudent and feasible way to help neophyte researchers is to raise questions that might draw out the important philosophical, design, and contextual considerations attending each phase. In the remainder of this paper I attempt to do just that—leaving it up to the analysts to decide which positions and corresponding strategies they deem most appropriate to take in each context. [10]

Because of the interactive, iterative, and non-linear character of the four phases of QDA, all three sets of considerations permeate each phase. Most of the questions I raise in phase 1, defining the analysis, point my readers (overtly) to philosophical or ideological considerations. However, that does not mean that philosophical considerations are quarantined in phase 1. In phase 1, for instance, I make comments and raise questions that urge my readers to consider whether and how to use computer programs in aiding their analysis. Although these comments and questions point overtly to the analyst's ontological and epistemological assumptions concerning meaning-making and textual interpretations, they also suggest design and contextual considerations. Let me provide a practical example of how contextual considerations might be implicit in what seems to be a purely philosophical matter. As a research methodologist, I have often advised students who insist on using frequency of occurrence of a particular word or phrase as the primary evidence of its meaningfulness or significance. As I get to know such students better (and as they come to trust me more) many of them would confide that their analytic decision was based, not on their own philosophical conviction, but rather on the demands of their more positivistic thesis chair. This is an example of how contextual considerations (covertly) drive analytic decisions. The point I wish to make is this: although the focus of my discussion and questions in phase 1 is overly philosophical, this emphasis should not preclude contextual or design considerations. [11]

By contrast, most of the questions I raise in phases 2 and 3 focus my readers' attention primarily toward design and contextual considerations. However, that does not mean that phases 2 and 3 are devoid of philosophical considerations. For example, in Phase 2 when I discuss tagging and labeling of data I ask: "In this particular study do I tag my data implicitly or explicitly?" Responses to this question depend (no doubt) on contextual and design considerations. However they point (if only tacitly) to philosophical considerations as well. [12]

This paper will become unnecessarily repetitious and laborious if I were to include all of the philosophical, design and contextual elements that attend each and every question I address in each of the three phases I explore below. My readers are therefore encouraged to address, not only those considerations that are made explicit by my questions, but also those that are implied. [13]

2.1 Phase 1: Defining the analysis

To conduct QDA analysts capture, record, interpret, and convey information. To do so they define the analysis, that is to say, they decide on the goals of the analysis; what counts as appropriate and sufficient information; and on how best to capture, record, interpret, and convey that information. This is a continuous process that begins with the initial conception of the study and proceeds through data gathering, reduction, and write-up. The decisions analysts make concerning defining the analysis may be tacit or overt. However, decisions (to define the analysis) are always made. Good analysts, therefore, strive to make these decisions as transparent and defensible as possible. [14]

The most crucial aspect to defining the analysis is attending to how our ideological and philosophical orientations structure what we do as analysts. Specific clusters of ideological and philosophical assumptions along with their accompanying practices are often referred to as research paradigms (KUHN, 1970; LINCOLN & GUBA, 1985, p.15). A paradigm is a worldview—a way of seeing, being, and acting in the world. Three popular paradigms discussed in the qualitative research literature are positivism, interpretivism, and constructivism (DENZIN & LINCOLN, 2000; LINCOLN & GUBA, 1985; MERRIAM, 1991; POLKINGHORNE, 1983). Research methodologists and philosophers of science do not agree on what constitutes a paradigm, on how many paradigms there are, on whether a person is limited to a single paradigm, or is free to choose whichever paradigm she/he desires (GREENE & CARACELLI, 1997; GUBA, 1990; KUHN, 1970; LAKATOS & MUSGRAVE, 1974; TASHAKKORI & TEDDLIE, 1998). However, most qualitative research methodologists agree that researchers are (inevitably) guided by a paradigm(s), and more and more of them are suggesting that attention to paradigmatic issue generally improves the quality of one's research. [15]

The qualitative research literature also refers broadly to this aspect of the research process as considerations of researcher identity or role (CROCKETT, 1973; LINCOLN & GUBA, 1985; QUIGLEY, 1997; STAKE, 1995). In short, I am suggesting that it is in defining our roles as analysts that the entire analysis, itself, is defined. To define ourselves, and consequently our analyses, I suggest that we examine four very interrelated domains: our ontology, epistemology, axiology, and notions of causality. [16]

2.1.1 *Ontology*

Research is always an attempt to investigate *something*. That thing is usually called "reality." Ontology deals with the question of what is real. Reality is a most difficult issue to grasp, so I will devote a quite a bit of space to elucidating it. Consider for a moment the question: *What do you mean when you say that something is real?* Is learning or electricity real? Are stones or trees real? Is poverty, homelessness, or co-dependency real? Are your thoughts, dreams and perceptions real? Are myths and fables real? Are flying horses or immortal women real? Are curriculums or "at risk students" real? Are phenomena such as andragogy, self-directed learners, and democratic societies real? [17]

What factors are likely to frame your answers to these questions? Is it the phenomenon's relative accessibility to others? Is it its relative separateness from, or independence of you, the knower—in a word, its objectivity? Is it its tangibility or relative permanence? Or might it be the phenomenon's clarity of definition, or its intelligibility, meaningfulness, or usability? If you privilege tangibility, objectivity, and accessibility, then your thoughts, dreams, and perceptions are not real. If you privilege usability, then learning and electricity might seem *more* real than stones and trees. On the contrary, if you privilege clarity of definition, then learning and electricity might seem less real than stones and trees. But what does it mean to say that something is less real? If a thing is less real does it still exist? Are there

things that exist that are *unreal*? Do our understandings of what it means to be real shape in any way what we might do as researchers? For instance, if a thing is "less real" does that mean that it is less worthy of being studied? Or does it simply mean that it needs to be studied in a different way? And, ultimately, does it matter at all (to the conduct of your analysis) whether and how things are real? Does the concept of reality make any sense at all? If we discard the concept, would we need to replace it? If so, with what? [18]

One of the most vexing questions surrounding the issue of reality is the distinction between the facticity and quality of a thing. Facticity asserts that a thing exists in some form. Quality tells us in what form the thing exists. Other terms for facticity are possibility or thatness. Other terms for quality are nature or whatness. An important question to consider is this: Is it possible to separate the *facticity* of a thing from its *quality*? Put differently, can I assert that a thing exists if I have absolutely no perception of what it is? In philosophical jargon: is it possible to separate my ontology of a thing from my epistemology of it? To illustrate: Can I assert that homelessness or democratic societies exist, if I have absolutely no idea of what homelessness or democratic societies are? Or, rather, is it my pre-conceived notions of what these things are which leads me to assert that they exist? And if I always possess pre-conceived notions of what things are before I could conduct investigations into them, then are my investigations "research" or "propaganda?" Are there really any differences between the two? If so, what are they? How do I minimize the risk of my research becoming propaganda? Positivists, for instance tackle this issue by paying attention to matters of replicability, reliability and objectivity (BABBIE, 1998, chap. 5). Constructivists, on the other hand, discount replicability, and deny objectivity, and instead concern themselves with notions such as dependability and confirmability (LINCOLN & GUBA, 1985, chap. 11). In summary, here are three overarching questions an analyst might ask regarding the ontological question:

1. In this research project what do I mean when I say that something is real?
2. How does my notion of reality shape the kinds of information I capture, record, interpret and convey?
3. How does my use of analytic tools (including computer programs) shape what I capture, record, interpret and convey? [19]

2.1.2 A word about computer programs

In the past decade or so, a number of computer programs has been developed to aid QDA—NUD*IST, NVivo, Ethnograph, Atlas/ti, to name a few (see WEITZMAN & MILES, 1995 for a more comprehensive list of programs and descriptions). Like any other analytic tool, computer programs may promote certain ways of construing and conducting QDA while precluding others. For instance because of the ease with which computers allow us to count things, it could be quite seductive for analysts to use "frequency of occurrence" as the sole measure of meaningfulness or significance. Clearly, the number of times a particular phrase appears in a text is not a sufficient measure of its importance or significance—

more semiotic and linguistic tools (I believe) are needed to make that assessment. Being one who uses computer programs, I certainly encourage their usage. However, I believe that it is important as analysts that we constantly think about how these programs are enhancing or constraining our analysis. Most qualitative research texts written after 1990 include a section discussing the role of computers in qualitative research. We are also beginning to see entire works devoted to the subject (see, for instance, BARRY, 1998; FIELDING & LEE, 1993; 1998; KELLE, 1997; KELLE & ASSOCIATES, 1995). [20]

2.1.3 Epistemology

Epistemology deals with the nature, sources and processes of knowledge and knowing. Regarding the nature of knowledge the analyst would address questions such as:

1. In this project, what terms best describe what I seek to find out or produce: knowledge; meaning; truth; information; data; etc.; etc?
2. Is my goal to discover the single, correct answer? Or is it to produce defensible perspectives? [21]

Regarding the sources of knowledge the analyst would address question such as:

1. In this project, what do I count as appropriate and inappropriate sources of knowledge? Do I focus only on behaviors that seem to be directly accessible to the five senses—touch, sight, sound, taste, and smell? Do I ignore those behaviors that are not directly accessible to the five senses—behaviors such as people's beliefs, perceptions, intentions, revelations, and so on?
2. In addition to the behavior of my research subjects, do I also include my own behavior and the behavior of my co-researchers as sources of data? Do I include as data, our (the researchers') modes and mechanisms of observation; and also our modes and mechanisms for capturing and storing data? Do I include our memory? [22]

Regarding the processes of knowing the analyst would address question such as:

1. In this project, do I desire to be a pure observer¹, and if so why?
2. How do I attain the status of a pure observer?
3. If I do not desire (or cannot attain) pure observation, what are my ethical alternatives? [23]

2.1.4 Axiology

Axiology is the domain of values and ethics. Regarding QDA, it comprises three issues: a) the place and role of the researcher's values in research; b) the role of research subjects; and c) the appropriate way(s) to use research products. [24]

¹ A pure observer is a researcher who is merely a fly on the wall, so to speak—neither affecting, or being affected by what she observes.

Regarding the role and place of the researcher's values in research the analyst might ask:

1. In this project, do I think it is *possible* to keep my values from influencing my analysis? If yes, how do I accomplish this? If no, when and how will I use my values in my analysis, and how will I defend the use of my values?
2. Do I think it is *desirable* to keep my values from influencing my analysis? If yes, why? If no, why not? [25]

Regarding the role of research subjects the analyst might ask:

1. In this project, how do I involve my research subjects? Do I use them as informants or participants; as objects to be studied; or as subjects—having an active decision-making role in all aspects of my study?
2. What factors govern the nature, level and extent of involvement I accord my research subjects?
3. Given that their participation may involve some level of personal risk (legal, mental or physical), how do I assure confidentiality and avert risk to my subjects? [26]

Regarding the use of research products the analyst might ask:

1. In this study, how do I use my research products? Do I generalize—that is to say do I sometimes apply my research "findings" to people who were not part of the study? If so, when and under what conditions do I generalize?
2. If you do not generalize, how else do I use my research products? Do I leave the application up to my subjects and readers? If so, why?
3. In this study, which do I consider to be the most important indicators of research quality: meaningfulness; usability (the ease with which the findings can be used); generalizability; objectivity, replicability, prestige; etc.; etc.? [27]

2.1.5 Causality

Causality is not a domain in the same sense as the other three discussed above. It is more integrally a part and parcel of all three. I pull it out for scrutiny because it is often overlooked in discussions of qualitative research (see LINCOLN & GUBA, 1985, chap. 6 as a notable exception). All research attempts to draw associations between ideas, people and/or events. These associations are called different things—cause, probable cause, mutual shaping, influence, determination, contribution, effects, and so on. Theorists differ as to the meaning and ontological status of these associations (BABBIE, 1998, pp.73-79; LINCOLN & GUBA, 1985, chap. 6). BABBIE, for instance assume that causes exist (out there) as an inherent constituent of phenomena (events, interactions, and so on). LINCOLN and GUBA on the other hand believe that causes are human impositions "upon a continuous stream of occurrence," and they argue that this imposition is based on the researcher's purposes and interests (1985, p.137).

Most graduate students I have come across have given little or no thought to this matter. For this reason, the questions I now pose are all the more important. The responsible analyst asks himself/herself questions such as:

1. What crucial associations or relationships do I observe in this study?
2. What terms do I use to describe those associations—cause, probable cause, mutual shaping, influence, contribution, etc., etc.?
3. When I say that "A" causes "B" what do I mean? In other words, what is my definition of causality, and how do I establish and demonstrate it?
4. If I do not believe in causality, what alternative concepts do I employ to discuss associations and relationship between related phenomena? [28]

2.2 Phase 2: Classifying data

As with phase 1, classifying begins at the inception of the research process and intensifies as the analysis proceeds. Ideological and philosophical considerations continue to permeate this, and subsequent phases. However, because I have focused extensively on philosophical considerations in phase 1, the questions I raise in the next two phases will focus more overtly on design and contextual issues. [29]

Of the design considerations affecting the strategic decisions of the analyst, one cannot be overstated: namely, the interpretive and linguistic possibilities (textual, contextual, syntactic, grammatical, and so on) embedded in the literary genre in which the data are captured and stored. For example, naturally occurring conversations (as a literary genre) lend themselves much better to discourse or narrative analysis than do structured interviews (RIESSMAN, 1993; STRAUSS, 1987; STUBB, 1983). As COFFEY and ATKINSON (1996) point out, the complementary application of different literary genre and linguistic devices often increases the richness and robustness of the analysis. [30]

Phase two includes two aspects: 1) tagging data, and 2) grouping tagged data. Tagging refers to the process of selecting from an amorphous body of material, bits and pieces that satisfy the researcher's curiosity, and help support the purpose of the study. Unless everything one "collects" and records is considered relevant and important (a rather rare, and frightening prospect), one must decide to count as important certain pieces of data and discount (if not completely reject) others. Consequently, though the strategies and tactics used for tagging data may vary widely in form, QDA is meaningless without some form of tagging. Labeling (i.e., assigning some distinguishing mark to selected data) is always an aspect of tagging. Labels may come from the data itself, or they may be imposed from outside. Labels may be meaningfully related to the data or mere heuristic devices. As such, labels may be images, numbers, symbols, words, phrases, themes, etc., etc. [31]

Once tagged and labeled, the analyst places data with similar characteristics into the same group or category. Other synonyms in the literature for categories are

constructs, concepts, variables, and themes. A category is an idea (a word or phrase) that stands for a set of objects or events with similar characteristics—a class (DEY, 1993, p.17f.). Categories may be mutually exclusive or overlapping; exhaustive or incomplete. To place data into categories implies the formation of definitions (whether tacit or explicit). It's advisable that analysts make their definitions as explicit as possible and refine them as the process proceeds. Based on his/her researcher identity (see discussion of phase 1) and theoretical sensitivity, the analyst constantly compares (GLASER, 1978; GLASER & STRAUSS, 1967; STRAUSS & CORBIN, 1998) the data—deciding whether and where particular tagged units fit; whether to discard, discount, narrow, expand, split, join, transfer, etc., previously tagged data; and whether to discard, discount, narrow, expand, split, join, refine, re-define, etc., previously constructed categories. To classify, then, is to tag, label, define, and refine groups of data. This is a messy process that goes back and forth between the four intellectual moments—tagging, labeling, defining, and refining. Below I raise questions the analyst might consider when classifying data. I have relied heavily on COFFEY and ATKINSON (1996), DEY (1993) and FELDMAN (1995) in formulating these questions. [32]

2.2.1 General questions guiding choice of analytic strategies

1. With what kinds of analytic strategies am I familiar?
2. Are my analytic knowledge and skills adequate for the task?
3. How much control do I have over the analytic decisions I make? Who or what else besides me controls the decisions I make?
4. Given my knowledge and skills, and my philosophical and contextual constraints what overall analytic strategies and tools would best serve my research purposes? [33]

2.2.2 Questions to consider when tagging and labeling data

1. In this particular study do I tag my data implicitly or explicitly?
2. If I choose the former, how do I demonstrate trustworthiness? [34]

If I choose the latter:

1. Do I familiarize myself with all the collected data (transcripts, fieldnotes, etc.,) by reading and re-reading them before I begin formal tagging, or do I jump straight into tagging?
2. Do I tag *serially*, i.e., proceeding through each transcript, one at a time? Or do I tag in *parallel*, i.e., by reading and comparing responses to the same question (across different transcripts) until all the questions are exhausted?
3. At the start of this process, do I tag and label everything that seems interesting or only those ideas that appear to be directly related to the research question(s)?

4. Are my initial labels *in vivo* codes—actual data taken from my transcripts or field notes; my paraphrasing of the actual data; or terms imposed from outside?
5. Do I tag before, during, or after I develop categories?
6. During my initial labeling what level of abstraction from the data do I use—low, moderate, high?²
7. Do I use explicit or implicit strategies³ to help me interpret the data? What are the advantages and disadvantages of using either?
8. What are the advantages and disadvantages of combining different interpretive strategies, such as semiotics with discourse or narrative analysis? [35]

2.2.3 Questions to consider when developing and refining categories

1. In this study, what criteria do I use to determine the number, types, and levels of categories I develop? Am I guided by the purpose of the study and research questions? Am I guided by parsimony or the need to obtain a holistic picture? Am I concerned that my categories exhibit conceptual equivalence, or that my categories are mutually exclusive? Am I guided by availability of time and other resources? Or are my standards arbitrary—am I just winging it as I proceed?
2. Do my categories exhaust the data I consider to be salient? In other words, have I classified all of the data I consider relevant? Or are there salient data I have not classified? If so, what other categories might I construct to adequately account for the unclassified data?
3. In what form(s) is it best to conceptualize my categories: nominal, ordinal, interval or ratio?⁴
4. Are my definitions (of categories) sufficiently developed to clearly distinguish between categories; to aid in the proper assignment of tagged units; and to assist in adequately fulfilling the purpose of the study?
5. Is each category adequately supported by tagged data?
6. Might a particular tagged unit be assigned to more than one category?
7. In what ways are the tagged units that I have grouped together similar? In what ways are they different? And how/why have I captured or concealed those differences?
8. Do I have memos or other data sources that might support, refute, shed light, provide context for, etc., particular tagged units and categories?
9. Do I have memos and other aids to help preserve meaning and context that might otherwise be lost due to fragmentation of data? [36]

2 Questions 6 through 8 also apply to the section titled "Questions to Consider When Developing and Refining Categories."

3 Examples of explicit interpretive strategies include structured approaches to narrative analysis, content analysis, deconstruction, discourse analysis, semiotics, dramaturgy, ethnomethodology, and so on (BERG, 1998; COFFEY & ATKINSON, 1996; DEY, 1993; FELDMAN, 1995; MANNING, 1987; AND RIESSMAN, 1993; AND STUBBS, 1983).

4 See DEY, 1993, chap. 2 for a very good discussion of types of variables (categories) one might employ in QDA.

2.3 Phase 3: Making connections—Constructing stories and theories

Some qualitative research studies that I have come across conclude with a mere glossary: a list of concepts (themes, variables, categories, etc.) and their definitions. Little or no attempt is made to relate these concepts to each other (or to existing theory) to build a cohesive whole. Glossaries are, no doubt, quite useful tools. However, they lack the richness, insightfulness, depth of understanding and explanatory potential of stories and theories. And in the final analysis, it seems to me that it is that richness, that depth of understanding that is the genius of qualitative research. The point of research is not to tell people what they already know. The point is to help our subjects and readers understand more broadly and deeply their experiences. This criterion of deeper/broader understanding demands that analysts develop stories or build theories. They do so by positing a parsimonious, integrated set of associations and relationships between and among the various concepts they have formulated—relationships that were previously undocumented, obscure, or unknown. [37]

This is scary terrain. During this creative process, neophyte qualitative analysts usually become aware that relationships are not *sui generis* (self-generating and self-evident). They come to realize that data do not speak for themselves; that it is they (the analysts) who must make their data say what they want them to say. How scary! How creative!! This realization is particularly scary to those of us who have been socialized to value only objective knowledge. This desire for objectivity is probably one of the reasons why many qualitative researchers stop at glossaries. Glossaries seem more objective—less subject to researcher bias. Telling stories or building theories seem to demand way more of the analyst's own subjectivities and values. [38]

As a social constructivist, I believe that there is really no way around human subjectivity. Creativity demands it. This constructive/creative phase is implicated in all QDA, I would argue. Constructivists revel in this human subjectivity while positivists shun or attempt to control it. In my experience as a teacher of research methodology, I have seen graduate students become frustrated and paralyzed by the realization that research products (findings) are merely human constructions and not objective Truth. On the contrary, this realization has been a source of liberation for other graduate students. It is liberating because such students begin to regard themselves as Subject (constructing knowledge and participating in transforming their world), rather than as Objects (merely consuming knowledge and adapting to pre-given, unchangeable realities). This recognition that they are constructing and not discovering knowledge also leads some students to accept that it is quite possible for researchers (using what seems to be the same⁵ data, and asking the same questions) to produce quite different (but equally trustworthy) results. The difference between the two groups of students usually lies in the value each places on objective Truth. The one group is entirely wedded

5 I say: "what seems to be" because the words, symbols, and images which make up data take on meaning only in the moment of interpretation. Before that moment they remain meaningless objects. At that moment of interpretation, the words, symbols, and images take on different meaning for different individuals, so that there is never a time really when two people are viewing the *same* data.

to it—believing that the attainment of objective Truth is both possible and desirable. The other group, believing that attainment of objective Truth is neither possible nor desirable, gladly embraces human subjectivity. [39]

As I did at the beginning, let me underscore that this phase (as all other phases) is not linear. In fact, although discussed last, it really begins with data collection or even before with the conceptualization of the study. It intensifies as the analysis proceeds. For instance, we start to develop hunches (rudimentary connections) as soon as we begin defining the analysis (phase 1). Our hunches become working hypotheses as we tag, label and categorize data (phase 2). And toward the end of the process, as we gain more knowledge and confidence about our data, our working hypotheses become full-fledged stories or theories. A word of caution: It is advisable that analysts do not get too wedded to their initial hunches and working hypotheses. Such premature commitment often leads the analyst to ignore important new insights and relationships that may greatly enrich her developing story or theory. [40]

2.3.1 Questions to consider when constructing stories or theories

1. In this particular analysis, what relationships have I posited between and among my various categories and themes?
2. How have I related them: by time, space, context, function, consequence, intensity, level, dimension, etc.?⁶
3. What relationships are subordinate and which ones are superordinate? Which relationships are associational, causal, consequential, contingent, etc., etc.?
4. What data (transcripts, observations, field notes, memos, memory, etc.) have I used to support the relationships I posit?⁷
5. What existing theories have I used to help me construct and support these relationships?
6. What parsimonious and cohesive story or theory am I telling?
7. Does my story or theory present a holistic picture of the phenomenon under investigation? If not how do I make it more holistic; or how do I defend its incompleteness?
8. How does my story or theory confirm (and/or) challenge existing theory? [41]

6 The issues raised in questions 2 & 3 address what methodologists refer to as theoretical sensitivity. For excellent discussions of these issues see GLASER and STRAUSS (1967), GLASER, (1978) and STRAUSS and CORBIN (1998).

7 Most qualitative research texts provide good discussions of the issues raised in questions 4-7. I find the treatment of these issues by COFFEY and ATKINSON (1996) and DEY (1993) to be particularly instructive.

3. Conclusion

Many of the graduate students I teach are often bewildered by the seeming mixed messages they get from qualitative research textbooks concerning QDA. At one extreme are discipline/method-specific texts that focus on discipline/method-specific strategies—a different set of strategies for each method or discipline (CARSPACKEN, 1996; MOUSTAKAS, 1994; STRAUSS & CORBIN, 1998; STAKE, 1995; VAN MANEN, 1990). At the other extreme are general texts that take a one-size-fit all approach (BERG, 1998; MILES & HUBERMAN, 1994; PATTON, 1990). Texts such as COFFREY and ATKINSON (1996) and FELDMAN (1995) are notable exceptions; while illustrating specific but different analytic strategies, they also try to highlight common features of all QDA. However, space limits the range of strategies they could reasonably illustrate. In this paper I tried to bridge the gap between the generalists and the specialists. I described 3 (of 4) phases that I believe are common to all QDA. However, within my discussion of each phase, I posed philosophical, design and contextual questions that I hope would elicit a broad range of strategic differences. I do not claim that my questions are exhaustive, but I believe that their consideration might illumine the QDA process and ease some of the frustrations I see in graduate students. [42]

References

- Alinsky, Saul (1971). *Rules for radicals: a practical primer for realistic radicals*. New York: Vintage Books.
- Babbie, Earl (1998). *The practice of social research* (8th edition). Belmont, CA: Wadsworth Pub.
- Barry, Christine (1998). Choosing qualitative data analysis software: Atlas/ti and Nudist compared. *Sociological Research Online*, 3(2), <http://www.socresonline.org.uk/3/3/4.html>
- Berelson, Bernard (1952). *Content analysis in communications research*. Glencoe, IL: Free Press.
- Berg, Bruce (1998). *Qualitative research methods for the social sciences* (3rd edition). Boston, MA: Ally and Bacon.
- Berger, Peter & Luckmann, Thomas (1967). *The social construction of reality: a treatise in the sociology of knowledge*. New York: Doubleday/Anchor Books.
- Carspecken, Phil (1996). *Critical ethnography in educational research: a theoretical and practical guide*. New York: Routledge.
- Charmaz, Kathy (2000). Grounded theory: objectivist and constructivist methods. In Norman Denzin & Yvonna Lincoln (Eds.), *Handbook of qualitative research* (2nd edition) (pp.509-535). Thousand Oaks, CA: Sage.
- Coffey, Amanda & Atkinson, Paul (1996). *Making sense of qualitative data: complementary research strategies*. Thousand Oaks, CA: Sage.
- Cohen, Marlene & Omery, Anna (1994). Schools of phenomenology: Implications for research. In Janice Morse (Ed.), *Critical issues in qualitative research methods* (pp.136-156). Thousand Oaks, CA: Sage.
- Creswell, John (1998). *Qualitative inquiry and research design: choosing among five traditions*. Thousand Oaks, CA: Sage.
- Crockett, Stanley (1973). The role of the researcher in educational settings: perspectives on research and evaluation. *Journal of Social Issues*, 29(1), 81-86.
- Denzin, Norman & Lincoln, Yvonna (2000) (Eds.). *Handbook of qualitative research* (2nd edition). Thousand Oaks, CA: Sage.

- Dey, Ian (1993). *Qualitative data analysis: a user-friendly guide for social scientists*. New York: Routledge.
- Feldman, Martha (1995). *Strategies for interpreting qualitative data*. Thousand Oaks, CA: Sage.
- Fetterman, David (1998). *Ethnography: step by step*. (2nd edition). Newbury Park, CA: Sage.
- [Fielding, Nigel](#) & Lee, Raymond (1998) (Eds.). *Computer analysis & qualitative research*. Thousand Oaks: Sage.
- Fielding, Nigel & Lee, Raymond (1993) (Eds.). *Using computers in qualitative research* (revised & updated). Thousand Oaks: Sage.
- [Glaser, Barney](#) (1992). *Basics of grounded theory analysis*. Mill Valley, CA: Sociology Press.
- Glaser, Barney (1978). *Theoretical sensitivity: advances in the methodology of grounded theory*. Mill Valley, CA: Sociology Press.
- Glaser, Barney & Strauss, Anselm (1967). *The discovery of grounded theory: strategies for qualitative research*. New York: Aldine de Gruyter.
- Greene, Jennifer & Caracelli, Valerie (1997). *Advances in mixed method evaluation: the challenges and benefits of integrating diverse paradigms*. San Francisco: Jossey-Bass.
- Guba, Egon (1990). *The paradigm dialog*. Newbury Park, CA.: Sage.
- Heidegger, Martin (1962). *Being and time* (J. MacQuarrie & E. Robinson, Trans). New York: Harper & Row.
- Husserl, Edmund (1931). *Ideas: general introduction to pure phenomenology* (D. Carr, Trans.). Evanston, IL: Northwestern University Press.
- [Kelle, Udo](#) (1997). Theory building in qualitative research and computer programs for the management of textual data. *Sociological Research Online*, 2(2), <http://www.socresonline.org.uk/2/2/1.html>.
- Kelle, Udo, with Prein, Gerald & Bird, Katherine (1995) (Eds.). *Computer-aided qualitative data analysis: theory, methods and practice*. Thousand Oaks, CA: Sage.
- Kuhn, Thomas (1970). *The structure of scientific revolutions*. Chicago: University of Chicago.
- Lincoln, Yvonna & Guba, Egon (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Manning, Peter (1987). *Semiotics and fieldwork*. Newbury Park: CA: Sage.
- Marshall, Catherine & Rossman, Gretchen (1995). *Designing qualitative research* (2nd edition). Newbury Park, CA.: Sage.
- Merriam, Sharon (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass.
- Merriam, Sharon (1991). How research produces knowledge. In John Peters & Peter Jarvis (Eds.), *Adult education: evolution and achievements in a developing field of study* (pp.42-65). San Francisco: Jossey-Bass.
- Merton, Robert (1967). *On theoretical sociology: five essays, old and new*. New York: Free Press.
- Miles, Matthew & Huberman, Michael (1994). *Qualitative data analysis: an expanded sourcebook* (2nd edition). Thousand Oaks, CA: Sage.
- Moustakas, Clark (1994). *Phenomenological research methods*. Thousand Oaks, CA: Sage.
- Patton, Michael (1990). *Qualitative evaluation and research methods* (2nd edition). Newbury Park, CA: Sage.
- Polkinghorne, Donald (1983). *Methodology for the human sciences: Systems of inquiry*. Albany, NY: SUNY.
- Quigley, Allan (1997). The role of research in the practice of adult education. *New Directions for Adult and Continuing Education*, 73, 3-22.
- Ray, Marilyn (1994). The richness of phenomenology: Philosophic, theoretic, and methodologic concerns. In Janice Morse (Ed.), *Critical issues in qualitative research methods* (pp.117-135). Thousand Oaks, CA: Sage.
- Riessman, Catherine (1993). *Narrative analysis*. Thousand Oaks, CA: Sage.
- Stake, Robert (1995). *The art of case study research*. Thousand Oaks, CA: Sage.

Strauss, Anselm (1987). *Qualitative analysis for social scientists*. New York: Cambridge University Press.

Strauss, Anselm & [Corbin, Juliet](#) (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd edition). Newbury Park, CA.: Sage.

Strauss, Anselm & Corbin, Juliet (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA.: Sage.

Stubbs, Michael (1983). *Discourse analysis: the sociolinguistic analysis of natural language*. Chicago: University of Chicago Press.

Tashakkori, Abbas & Teddlie, Charles (1998). *Mixed methodology: combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.

Van Manen, Max (1990). *Researching lived experience: human science for an action sensitive pedagogy*. Albany: SUNY.

Vygotsky, Lev (1978). *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wolcott, Harry (1994). *Transforming qualitative data, description, analysis, and interpretation*. Thousand Oaks, CA: Sage.

Weitzman, Eben & Miles, Michael (1995). *Computer programs for qualitative data analysis: a software sourcebook*. Thousand Oaks, CA: Sage.

Author

Dr. *Ian BAPTISTE* was born and raised in the Caribbean—Trinidad and Tobago, and Grenada. He did his undergraduate work in Agriculture at the University of the West Indies, Trinidad; his masters in Foundations of Education at Wheaton College, Illinois; and his doctorate in Adult and Continuing Education at Northern Illinois University. He taught at National-Louis University (Chicago) for three and a half years, and chaired the department of Adult Education for most of his tenure there. Dr. BAPTISTE joined the Penn State faculty in August 1997. He currently teaches graduate courses in qualitative research methods, adult teaching and learning, and learning in social action.

Dr. BAPTISTE's professional interests are in participatory approaches to community development; adult teaching and learning; and qualitative research methodology. He has been (and continues to be) a participatory research consultant to community-based organizations, NGOs and community colleges in Chicago, South Africa, and the Caribbean islands of Grenada and Trinidad. He currently chairs an interdisciplinary team of faculty and students at Penn State working to develop a Participatory Research Network of grassroots CBOs in central Pennsylvania. He is the Principal Investigator (PI) on a two-year research study (with three graduate students) exploring a phenomenology of adult learning; and, with one of his colleagues (Dr. CARR-CHELLMAN), he is Co-PI on an action research project which compares three Qualitative Data Analysis Programs—NUD*IST, NVivo & Atlas.ti. Dr. BAPTISTE is the author of several refereed publications. His most recent publication, "Educating lone wolves: pedagogical implications of human capital theory," appeared in the May 2001 of the *Adult Education Quarterly*.

Contact:

Ian Baptiste (Ed.D)

Assistant Professor of Education
The Pennsylvania State University
305B Keller Building
University Park,
PA 16802, USA

Phone: (814) 865-1958

Fax: (814) 865-0128

E-mail: ieb1@psu.edu

Citation

Baptiste, Ian (2001). Qualitative Data Analysis: Common Phases, Strategic Differences [42 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 2(3), Art. 22, <http://nbn-resolving.de/urn:nbn:de:0114-fqs0103226>.

Revised 3/2007