

# Chapter 93

## Qualitative Research Methods for Science Education

Frederick Erickson

This chapter on qualitative research methods in science education is divided into four major sections devoted to (1) the purposes of qualitative research, (2) data collection, (3) data analysis and (4) preparing reports.

### Purposes of Qualitative Research

The essential purposes of qualitative research are to document in detail the conduct of everyday events and to identify the meanings that those events have for those who participate in them and for those who witness them. The emphasis is on discovering *kinds* of things that make a difference in social life; hence, an emphasis is placed on *qualitas* rather than on *quantitas*. This priority of emphasis does not mean that information about frequency is irrelevant to qualitative inquiry, for good qualitative research reports the range and frequency of actions and meaning perspectives that are observed, as well as their occurrence, narratively. The crucial problem for the qualitative researcher, however, is determining the “qualities” of social action and meaning.

Qualitative research in education is especially appropriate when we want:

- Detailed information about implementation
- To identify the nuances of subjective understanding that motivate various participants in a setting
- To identify and understand change over time

Human social action and opinion are locally distinct and situationally contingent. What at first glance can seem to be the same sort of setting, event, or point of view

---

F. Erickson (✉)

Inaugural George F. Kneller Professor of Anthropology of Education, Emeritus,  
Los Angeles, CA 90095-1521, USA  
e-mail: ferickson@gseis.ucla.edu

can be subtly different in kind despite surface similarity. When we are not certain about the details of local implementation of educational practices, then documentation through qualitatively sensitive narrative description is necessary. We need to be able to answer the question “What was the treatment, specifically?” before we try to answer the question “What were the effects of the treatment?”

At the most fundamental level, we need to determine whether or not the intended program was implemented in its most ordinary and material aspects. Were there classrooms and teachers available? Were the classrooms equipped adequately (e.g., with laboratory tables)? If there were tables, did their water taps and sinks work properly? Were the teachers prepared for the new teaching methods and materials? Did the books get published in time and did they actually arrive at the classrooms when the school term began?

At a more subtle level, we can study implementation by observing and documenting classroom discourse and pedagogy (e.g., Roth and Roychoudhury 1993). For example, in the newer “constructivist” approaches to teaching science and mathematics, the emphasis is on the students’ construction of knowledge. To encourage such an active stance toward learning, we in the USA assume that a teacher leading a class discussion would avoid entirely, or at least use quite infrequently, traditional “teacher questions” (i.e., known information questions in which the teacher knows the answer and the students know that the teacher knows). However, because of the power of customary cultural expectations of both teachers and students, it is difficult to change these conversational patterns in classroom discourse. If classroom discussion involves the teacher and students continually sliding back into their old habits of known information questions and emphasis on procedural correctness for getting the right answer, we could say that the new “constructivist” curriculum was not actually implemented, even though everyone went through the motions of implementation.

Two examples of questions that are too general in focus to be useful to inform educational practice are: “Did the teachers like it?” and “Do the students understand it?” Which teachers liked or disliked what aspects, in which situations? Some kinds of dislike can come with unfamiliarity, while other kinds of dislike stem from a sense that the teacher’s identity is being violated in following recommended practices. This can involve the teacher’s identity as a professional, as a man or woman, as a member of a certain social class or ethnic/religious identification group, or some combination of these (e.g., Glasson and Lalik 1993).

Which students understood what, in the doing of it and after the fact of doing? Much qualitative research in science education has been motivated by the desire to gain more specific understanding of the cognitive processes by which students understand and misunderstand science content and its discourse (e.g., Roth 1994). Here, too, as with the issue of likes and dislikes, identifying what science content “means” to varying students involves probing subtle differences, especially the distinction between literal, referential meaning and more metaphoric social meaning. A student can “understand” the periodic table cognitively in a literal way while also “understanding” that such knowledge feels alien – that knowledge has become a metaphor for *not me*. The entire understanding of the student involves a combination of both referential and social meaning. This is to say that, in teaching and learning

science, there are always issues of hidden curriculum combined with manifest curriculum, for students and for teachers as well.

The understandings of members in the setting, while the central focus of qualitative inquiry, are not considered uncritically by the researcher. There could be contradictions between intentions and actions and there can be systematic blind spots in the awareness of both teachers and students. Part of the responsibility of the qualitative researcher is to go beyond what the local actors understand explicitly, identifying the meanings that are outside the awareness of the local actors, and revealing the hidden curriculum so that it can be faced critically by teachers and students (see especially Lemke 1990). There is also the issue of curriculum integration – intended and unintended. For example, science can be a rich environment for the acquisition of literacy skills (Florio-Ruane 1982), but many teachers might not realize this.

Teaching and learning in science education are discursive activities (Lemke 1990). By *discourse*, I mean both its small letter “d” sense and its large letter “D” sense. In the small letter “d” sense, discourse refers to the conduct of immediate social interaction by verbal and nonverbal means. Science is talked and written in words – its ideas are not only expressed in numbers. Learning science is learning a new dialect and, as with the acquisition of other aspects of language, learning the dialect of science occurs in face-to-face conversation with others. Which kinds of classroom conversations appear to offer especially rich opportunities for understanding science? Which conversational roles appear to be most productive for students and for teachers? Sociolinguistically informed microanalysis of classroom discourse, in the small letter “d” sense, offers much potential for study of the acquisition of scientific understanding.

Discourse also has a broader meaning, in a large letter “D” sense. Learning is not only a matter of participation in an immediate conversation, but it also involves joining in a larger Conversation whose interlocutors, language, topics, and political and economic interests range far in social space and time. To do science and to know it is to engage as an interlocutor in that larger conversation – with Newton, Einstein, and Heisenberg, for example, and with their financial patrons – King Charles II of seventeenth-century England and the German and American governments and business interests of the twentieth century, including munitions manufacturers and what was to become the aerospace industry. To engage in the Discourse of science is to adopt not only a dialect, but a voice – a stance toward the phenomenal world and to society. In this larger sense, the Discourse of science can be thought of as the totality of knowledge and social situation that it takes to adopt successfully the roles of doing science, as a student, as a teacher, as a researcher, or as one who seeks and receives scholarly and financial sponsorship (Gee 1990).

Not everyone wants to buy into such roles – sometimes the risks might seem too great. All learning involves risk. Yet, to take the leap of risk as a learner, I think that there must not only be a safe and predictable learning environment, but also the learner must have a sense of entitlement, an audacity.

In societies throughout the world, the sense of entitlement is unequally distributed. Those from upper-class and middle-class backgrounds have more entitlement than do the very poor. For the already advantaged, the life project is progress – perhaps even

advancement. For the very poor, the life project is survival – do not lose what you already have, and do not risk much because the stakes of trying and failing are so high. This might explain, in part, why constructivist approaches are resisted by some students, by some parents, and by some teachers. The ambiguity of not knowing a right answer or a right procedure is scary business for a learner.

I do not believe that it is impossible that those who have grown up in circumstances in which they and their parents have little power and little respect could dare to try at difficult school tasks. But I do think that they might need special encouragement and special safety in the classroom – the safety to be imperfect and in process (Erickson and Shultz 1991).

Scientific knowledge is power, as is all other knowledge, according to Foucault (1979). If we wish to change the distribution of scientific knowledge and prestige in society through a new kind of science education, then the study and practice of science education needs to address issues of the political economy and the semiotics of scientific knowledge and of its acquisition. The “meanings” of scientific knowledge and skill are deeply embedded in issues of power, risk, trust, legitimacy, and in-group/out-group distinction and ranking. Studies of student “misconceptions,” which do not address issues of power/knowledge, seem narrow and shallow in comparison. They fail to mine the richness of meaning that is inherent in the study and practice of science.

## Data Collection

### *Research as Searching*

To do research is to pay unusually close attention and to reflect deliberately on what we have seen and heard. “Re-searching” is to seek and seek again, recursively. The basic issues in designing strategies for data collection are to think where we would need to be searching, with whom and in what relationships. Addressing such issues is necessary in order to gather evidence to warrant the assertions that one would like to be able to make in answer to the main research questions that have been posed in the study.

These issues – where to be as researcher, with whom, and how – have both intellectual and ethical dimensions. Because the literature on qualitative research has emphasized issues of data collection and research ethics, and because of the limitations of space, I do not discuss data collection in detail here. Rather, I state briefly a number of points, which I think are especially important. These points and more detailed discussion also can be found in other publications I have authored on qualitative research methods (e.g., Erickson 1986). In addition, I have found especially helpful the writings of Hammersley and Atkinson (1983), Miles and Huberman (1984), Clifford and Marcus (1986), Strauss (1987), Bogdan and Biklen (1992), LeCompte et al. (1992), Denzin and Lincoln (1994), Wolcott (1994), Denzin (1996), and Lareau and Shultz (1996).

### *Framing Questions*

The research report will consist of answers to the questions, which one has framed; thus, good questions are at the heart of the inquiry. Of course, because settings are locally distinct, one cannot anticipate fully in advance the circumstances that will be encountered when the study has begun. Research questions, data collection operations, and research role relationships necessarily change during the course of a qualitative study. In spite of this it is useful to frame questions in advance and think of the kinds of evidence that we would want to have accumulated in order to answer those questions, as well as anticipating issues of ethics.

### *Variety in Sources and Kinds of Evidence*

The participant observer uses two primary means of data collection: looking and asking. What people's doings mean to them might be apparent from looking, but often determining this also necessitates asking them by means of informal and formal interviewing. We also might need to ask because we cannot be everywhere in the present and because we cannot observe what has happened in the past. Yet asking is often more intrusive than watching, even when the asking is done very informally. The ideal process, in my view, is a recursive process of observation and interview in which, at each step along the way, insights gained by one method (either by looking or by asking) are followed up using the other method.

Looking and asking in a setting can produce differing sources and kinds of data, each with a distinct epistemological status as evidence: field notes written by an observer; interview comments; machine recordings; and site documents, including demographic and historical material. An effective data collection design includes as many of these different sources as possible, and always includes observation, interviewing, and collection of site documents and often including machine recording as well. As data analysis proceeds, when hunches about patterns that were developed on the basis of field notes are cross-checked and confirmed by reference to interview data or site documents, one has a stronger evidentiary claim than if evidence came from only one information source (the formal term for this is "triangulation"). Indeed, if we think of the evidence collected in a qualitative study that warrants a particular concluding assertion as consisting of information bits, an assertion warranted by 500 bits from field notes, 500 from interviews, 250 from site documents, and 250 from videotape analysis is more credible than an assertion warranted by 4,000 bits from interview comments or from field notes alone.

Thus, in designing data collection strategies one needs to anticipate the variety in kinds, sources, and amounts of evidence that will be necessary in order to draw credible conclusions and present them in a report. Data collection strategies can be planned in general at the outset. For example, in a study of the changing student conceptions of dynamics in a high school physics class, one can anticipate needing

to observe class sessions firsthand (for a certain number of days over a number of weeks, or for a complete unit or topic of subject matter, or for an entire semester or school year) and possibly also needing to videotape on specified days – placing a wireless microphone alternately on various focal individuals in the class. One would want to interview students and the teacher outside of class. One also would want to collect student work (notebooks, scribbles on worksheets, and journals) for focal individuals and perhaps for the whole class. Perhaps, on a daily basis, this material might be photocopied and then handed back to the students (this could be done by arranging to use the photocopying machine in the school office immediately after the class meeting). One might also want access to school records for demographic and family information as well as for prior scores on achievement tests, comments by prior teachers, or attendance and credit accumulation information. In addition, one might want demographic and historical information on the neighborhoods of students, including census tract and block data. One also might wish to interview the parents of focal students.

### ***Ethics and the Negotiation of Entry***

Researchers are obliged ethically to anticipate what will be done in data collection, analysis, and reporting, and to explain to those studied why it will be done that way rather than some other. In order to negotiate entry and deal responsibly with the concerns of those who will be studied it is necessary to tell them how we plan to conduct the study so that they can consider and give us advice about what that will mean to them in convenience and in safety. Without such knowledge their consent will not be genuinely informed. Written agreements are helpful in specifying the conditions of research.

### ***Risk***

The primary ethical obligation of the researcher, as it is of the physician, is to do no harm. Since qualitative research does not involve biochemical intervention of the sort found in medical research, the risks of physical harm are minimal. Usually this is true also for risks of social harm. As most qualitative research topics in education are framed, ordinarily the maximum risk to school students, teachers, or administrators is that of slight psychological harm due to embarrassment or to anxiety concerning the possibility of embarrassment. Admittedly, sometimes more than embarrassment could be at risk (e.g., if student performance were to be revealed as extremely poor or as involving academic dishonesty or if teaching or administrative performance were to be revealed as gravely incompetent). In such cases, administrative or legal punishment might result from exposure through research.

The risks of embarrassment or of administrative sanction for those studied are greatest when research information is shared in the local setting itself. For example, if a videotape or narrative vignette portraying a teacher is presented at a national meeting of researchers, few if any consequences to the teacher “back home” are likely to follow. However, if the same tape or vignette were shown to that teacher’s principal when the principal disagreed with the teacher’s approach to teaching, the risk of harm to the teacher would be much greater. Explicit agreements with teachers *and* with administrators about the circumstances under which information will be made available from the research, locally and nationally, can reduce anxiety about being videotaped.

### ***Informed Consent***

Consent that is genuinely informed and without coercion reduces the risk of social harm because it affirms the dignity and respects the agency of those who will be involved in the study. My experience has been that those studied become most anxious when they do not know the real purposes, potential audiences, and substantive foci of the research, as well as the boundaries around their participation that can be expected. Qualitative research requires not merely grudging and passive assent, but active participation in and commitment to the research by those who are studied. The best way to achieve trust with participants in the research relationship is by being trustable as a researcher – forthright and specific about what will be involved in participation in the study and respectful of the character and rights of those who agree to participate.

Issues of access and consent can be especially complicated when the classroom teacher is the researcher. Roles with colleagues and supervisors need to be renegotiated and at least oral assent granted by them. For example, if a teacher or principal is studying her own practice, and she takes notes in a staff meeting that will be used later as a resource for evidence (perhaps becoming the basis for a narrative vignette of a portion of that meeting that would appear in the written research report), assent to that teacher’s presence in that meeting *as a researcher* rather than as an ordinary colleague is ethically necessary. In practitioner research, just as in research conducted by outsiders to the school, it is necessary not only to gain general and collective consent for research that might involve other persons as non-focal research subjects (e.g., by a collective vote of staff in a meeting or, in the case of primary school children, by the school principal acting *in loco parentis*), but also to gain specific consent from those who will be studied as focal individuals – from parents (in the case of children of primary school age) and from the individuals themselves (in the case of older children and adults).

Conditions need to be negotiated for those colleagues being observed so that they are able to declare certain material off the record or on the record, or to declare certain material out of bounds entirely, and to know clearly when the practitioner-researcher’s “research light” is on or off. For researchers who visit a school as outsiders, the

“research light” issue is less complicated logistically and ethically, unless the outsiders are in the role of advocates and/or collaborators with those in the school. In that case, the same conditions for consent and for “research light” notification obtain for outsiders as they do for insiders who are conducting practitioner research.

## Data Analysis

### *Finding the Data*

In qualitative research, analysis is a boot-strapping operation in which, reflexively, assertions and questions are generated on the basis of evidence, and evidence is defined in relation to assertions and questions. Data analysis, informal and formal, begins as one is negotiating entry to the research site. It often continues in restudy after supposedly “final” reports are written. In a fundamental sense, data reanalysis never stops, and this is why it is sometimes so difficult for qualitative researchers to bring their work to closure.

Bodies of information are collected in fieldwork and are held in documentary sources in various media such as field notes, interview tapes, videotapes, and site documents. These are not yet *data* as they appear in raw form; they are more appropriately regarded as *resources for potential data*. The documentary sources contain many thousands of information bits, not all of which are relevant to the inquiry that is being conducted. Analysis consists in recursive review of information sources with a question or assertion in mind, deciding progressively which information bits to attend to further and, perhaps even more importantly, which not to attend to. This reminds me of an aphorism from the graphic arts that states that “to draw is to leave things out.”

In experimental research, the decisions about what will constitute data are made in advance of data collection and analysis. In participant observational research, data analysis and data definition are largely a matter of post hoc decision making. Such decision making is not capricious. As in historical research, it follows certain principled lines.

The fundamental issue is determining the extent of generalization, not as one’s assertions apply to settings beyond the one that was studied (i.e., to external generalization), but as the findings concerning patterns in the setting are supported by evidence from within the setting (i.e., to internal generalization which involves generalization within the case rather than beyond it).

### *Finding Assertions*

One can start with a tentative, working assertion about a pattern whose generalization within the setting could be checked later. For example, in a study of student conceptions of physics, one might want to assert that students hold an implicitly



Aristotelian conception of dynamics at the beginning of teaching a physics unit. One can also start analysis by drafting a narrative vignette, or by presenting an interview quote that illustrates students' physics conceptions.

Taking the former course is to begin analysis narratively by telling a story. This is incipient analysis because any coherent narrative account contains within itself an implicit theory of the organization of the events that the narrative describes. Beginning qualitative researchers and even more experienced ones often find that stating assertions is intimidating; it is premature to state a conclusion, one thinks. In that circumstance, breaking into analysis through narrative is an appropriate strategy.

### *Searching Data Sources for Evidence*

Whether one begins analysis by framing a working assertion, or by telling a story in first draft, the next steps are crucial. They involve testing the evidentiary warrant for the assertion that is explicitly stated or is implicit in the narrative account. Such testing requires searching the entire corpus of information sources for any information that might bear on the working assertion. (A working assertion can be thought of as a tentative answer to a particular research question.)

To return to our hypothetical example of a study of student conceptions of dynamics, field notes of observations would be searched for any evidence that might confirm or disconfirm assertions about student conceptions. Interviews with students also would be reviewed with the same issue in mind, as would site documents, videotapes of classroom interaction, and any other possible sources of evidence that might bear on the issue of student conceptions of dynamics. If one knew that certain sources of evidence (e.g., site documents and a certain round of interviews) did not contain evidence about student conceptions of dynamics, these sources could be ignored in the search. However, any source with potential for data that bear on the assertion should be reviewed at least once.

The initial search for evidence needs to be exhaustive in order to ensure that crucial disconfirming evidence was not systematically ignored. Because there are many connected assertions in a final qualitative report, linked hierarchically across differing levels of generality and involving differing levels of inference, the corpus of research materials is searched repeatedly, considering each single assertion and each set of assertions in turn. In these searches, the researcher begins to employ verbal coding categories or some other means (such as colored markers highlighting portions of the field notes), thus indicating where relevant information is in the research corpus and what the content of those data are. As some assertions are disconfirmed in the search, they are revised, and the search is undertaken again with coding categories adjusted accordingly.

For example, for some students, a hard binary distinction between Aristotelian and Newtonian conceptions does not seem warranted in the data; those students seem neither Aristotelian nor Newtonian. Perhaps they are confused – one is not sure at first how to characterize these conceptions that do not fit easily into the dichotomy that had appeared at first glance. Having discovered during the search

a three-way typology of student conceptions (Aristotelian, Newtonian, and “other/possibly confused”), the researcher goes back to the sources previously reviewed when the binary typology was in mind. The researcher then re-sorts the data to see if the three-way typology can contain all the instances that were identified.

### ***Analytic Induction***

This recursive process of reviewing evidence with an assertion in mind, revising the assertion in the light of the evidence, and then reviewing the evidence again has been called the “constant comparative” method of data identification and analysis (Glaser and Strauss 1967). I find that term misleading. The process of comparison is indeed recursive and progressive but not *constant*. The point is that one continues reviewing evidence until all relevant data have been identified and compared. One then goes on to another assertion or chain of assertions. I prefer the classic term *analytic induction* (Lindesmith 1947).

Gradually, through such a process of progressive problem solving, one finds that certain kinds of phenomena – actions, opinions, and kinds of social actors in the setting – covary in regular ways. One discovers post hoc various comparison groups, or sets of persons, actions, and opinions that are progressively regrouped as comparative analysis proceeds.

To return to our hypothetical example, one discovers that more of the students who retain Aristotelian conceptions of dynamics sat in the back of the room and that they also got average grades in English and Social Studies, in contrast to those students who sat in the front of the room, among whom was to be found the largest proportion (in the class as a whole) of students with Newtonian conceptions. More of the students who seemed confused asked for help from the teacher than did the students who held Aristotelian conceptions. When help was asked for, it tended to be done politely.

A number of the students who held Aristotelian conceptions were boys, and they appeared to be less polite in class overall than were the possibly confused students among whom, as a set, girls were overrepresented. As a set, these impolite students were also of lower socioeconomic status (and racial minority status was overrepresented in that set) in contrast to those who seemed confused or those who held Newtonian conceptions. Although most of the boys who held Aristotelian conceptions were impolite in class (and interviews with the Social Studies teachers revealed that these students, for the most part, were impolite in their classrooms as well), there were a few boys who were somewhat more polite to the physics teacher than others in their set who held Aristotelian conceptions. Those polite boys – some of whom were of white working-class background and some of whom were African-American – did not seek help from the teacher, but they also did not appear as impolite as the others in their set. Reanalysis of interviews with those students revealed that their conceptions were changing somewhat in a Newtonian direction, and that their responses appeared a bit like those in the “other” group, yet still distinct from them.

Looking now at all the students in the set of those whose conceptions were neither clearly Aristotelian nor Newtonian, it appears that some were increasingly more aware of the contradictions in the Aristotelian position and that what, at first analysis might have been seen as “confusion” in their conceptions, was better construed as a movement in the direction of Newtonian conceptions. This was true in this set of students more for those who sought help from the teacher, but it also was true for those students who did not seek help, yet were relatively polite in class. Were the polite students somehow more willing to take seriously what was being taught than were the impolite students? Were the polite students trying harder to learn? Were they less alienated from the Discourse of science than were the impolite students?

From such lines of questioning and reasoning, working recursively back and forth between hunches and data, one progressively arrives at new insights. The data show patterns of covariation across partially ordered sets of persons, actions, and opinions, considered together comparatively. (Usually the sets are *partially ordered* in that all members of a set are not identical and some features or properties of members of a set might be shared with members of another set, with proportions of different types of set members varying across sets.)

In our hypothetical example, “confusion” begins to be seen as a process of shifting conceptions and of changing identification with the Discourse, and the persistence of Aristotelian conceptions seems partly to be a matter of attitude on the part of students, especially male students of working-class and racial minority background. This appears to be not only a matter of willingness to seek help but, more fundamentally, a matter of student stance toward the teaching and the course content – toward the School Discourse – because some students who were polite but did not seek help tended to be moving toward the Newtonian conception (even though some students who were both polite and sought help seemed further along in a Newtonian direction than those who were polite and did not seek help). More of the polite students, however, considered together as a set, held conceptions that were moving more in a Newtonian direction than did most of the students who were impolite. However, there were some male students of upper-middle-class background, African-American and white, who held Aristotelian conceptions and were impolite, just as there were some white and African-American working-class students who held Newtonian conceptions. Thus the patterns of covariation between social background and academic performance were not simple, as the various sets in comparison groups were partially ordered.

The kind of reasoning sketched in our hypothetical example is not primarily a search for cause, as in the physical and biological sciences. It is a search for understanding. Which kinds of actions make sense, for which social actors, in which social situations? When one is alienated from a Discourse, how does it make sense to *work at* not learning what is taught? Goodenough (1981, pp. 54–57) observed that, because social life is so contingent, the kind of prediction that is possible in the hard sciences is not attempted in interpretive qualitative research. One does not attempt to predict that certain events will happen. One does want to be able to predict how people will react if a certain event happens – what sense they will make of it. Such understanding is the main aim of qualitative inquiry.

## ***Frequency Counts and Discrepant Instances***

Even though the analysis is “qualitative,” it is apparent that the researcher must pay careful attention to frequency of occurrence, especially to relative frequency, in comparing different kinds of phenomena across differing comparison groups. It is necessary to count things and to make decisions carefully about what things to count and in which sets.

Usually in analysis that proceeds by analytic induction, the researcher identifies ways in which actions, opinions, or types of persons usually occur. These are the typical phenomena. One is also interested, however, in the atypical or those few discrepant cases whose closer analysis often can lead to new insights. Discrepant instances are not leftovers in analysis (e.g., working-class students who are not impolite or alienated in the physics classroom). Such instances and the circumstances in which they occur are scrutinized carefully. This is another reason why counting is important in qualitative research. It is essential, in exhaustive analysis of all instances of a field of phenomena, to identify the frequency of occurrence of all the types and subtypes, if one is to be able to distinguish between the typical and the atypical.

In a qualitative analysis, one wants to discover, through analytic induction, a few general assertions – pattern statements with a wide enough reach that they connect by threads, as it were, to sub-assertions, which ultimately are connected by threads to data bits across multiple sources in the total corpus of information sources. The most satisfactory analysis is one in which, by pulling a few threads at the top node of a *set of sets* of connected threads, more discrete bits of data are tugged on within the whole corpus of information sources than would have been by any other top-level nodes of threads (i.e., by alternate lines of interpretation and analysis).

## ***Changing the Questions***

What if, during the analytic process of progressive problem solving, it seems to the researcher that the questions posed at the outset of the study need to be revised? In an experimental study this would spell disaster. In a qualitative study it simply means that the post hoc analysis is working properly – discovering subtleties and contingencies that could not have been foreseen when the study was undertaken. This is why we do participant observational fieldwork: to discover what could not have been anticipated by the deductive reasoning of armchair theorizing. Often, even during fieldwork, as the result of partial or incipient analysis, it begins to seem that the research questions need to change. That represents progress. It is not a problem but an opportunity.

I am very suspicious of sustained participant observational research in which the initial research questions are not revised (in subtle shades of meaning or more fundamentally) by the time the study is concluded. In such studies, I suspect that the observer concentrated too hard on collecting information that confirmed initial

assumptions, and then in data analysis overlooked all the contradictions and discrepant cases that might have been apparent had the researcher taken a more tentative and judicious stance with regard to evidence and conclusions. Analytic induction, when successful, teaches us fresh insights – something we could not have known before we started our inquiry.

## **Preparing Reports**

### *Getting Started*

Because qualitative data analysis never stops, experientially the researcher never feels ready to complete a report or often even to begin it. At such junctures, our intellectual integrity and sense of humility in contemplating the unknown can be liabilities if we let them immobilize us. It might help to remember, then, that qualitative reporting is inherently tentative. A qualitative research report can be thought of as a rendering or a construction. It is not the reality it attempts to represent.

As a text, the report consists of answers to the primary research questions of a study. It is an exercise in rhetoric. It makes an argument that, to be successful, must be both clear and persuasive. Clarity and coherence in reporting do not come on the first try; it is necessary to sketch and rewrite, drafting sections in nonlinear order and addressing diverse themes. Usually, one does not start writing the beginning of the report but rather writes drafts of the middle section first, which is the main descriptive account. Only then does one draft a concluding section and then, finally, one turns to writing the introductory section.

Writing a report involves making strategic decisions about what material to include, how to sequence it, and how to handle the inevitable tension between presenting evidence and overview. By presenting descriptive detail as evidence one convinces the reader but also risks confusing the reader with information overload. Conversely, presenting an overview maximizes clarity but risks failing to persuade the reader because of lack of evidence and lack of subtlety in reporting and analysis. Thus, there is a danger that one's report will be either thick and murky or thin and trivial. How to achieve a report that satisfies both the need for evidentiary warrant and for clarity is a difficult task.

### *Not Trying to Say Everything*

Participant observational fieldwork amasses huge amounts of potential data. Only some of these become data through analysis, yet still more data are identified than could be included in any report. Thus, the process of sketching a series of first drafts most essentially involves deciding what *not* to include in them. Which of the many

pertinent vignettes will not be told as illustrations? Which interview comments will not be presented? The same aphorism applies for reporting as for data collection and analysis: to draw is to leave things out.

### ***Showing the Range and Frequency of Variation***

In presenting the argument of a report, it is desirable to show not only the most typical or obvious patterns, but also the full range of variation and relative frequency of occurrence of data. The atypical, discrepant instances, as well as the more typical ones, need to be reported if the report is not to be one-dimensional and superficial.

In the most effective qualitative research reports, information about relative frequency is not simply presented to the reader as a matter of faith in the author's integrity and judgment, using fuzzy cover statements in narrative such as "usually" or "sometimes" or "most people." It is both possible and desirable in qualitative reporting to be specific about frequency through the display and interpretation of simple frequency tables.

There are three main types of text in a qualitative research report: particular description; general description; and orienting commentary. Each of these types, which are discussed below, can be thought of as containing subtypes.

#### ***Particular Description***

This consists of narrative reporting of detailed evidence concerning the actions and beliefs of sets of persons in the setting that is being studied. Particular description can take the form of narrative vignettes that portray the actions of particular persons in specific events, or of quotes of what particular persons said in various interviews, or of quotes from particular site documents, or of a bit of demographic or historical information that applies specifically to a certain setting, such as a single classroom, household, or school building.

#### ***General Description***

This consists of synoptic reporting that displays evidence for the existence of certain distinct patterns in the overall ecology of action and belief in the setting being studied (i.e., its overall social organization and culture). Particular description, through vignettes and quotes, presents pieces of the overall social ecology. General description shows patterns of *generalization within the case*. It provides an evidentiary warrant for the relative typicality or atypicality of the specific vignettes and quotes that appear in the report and it portrays synoptically the setting

as a sociocultural whole. General description also could include historical, economic, or demographic information that situates the local activities that were observed firsthand within the wider ecology of broader sociohistorical processes.

General description can take the form of frequency tables (the simpler the better), of analytic charts and typologies that identify key patterns of contrast in the setting, or of time lines and flowcharts that identify sequential patterns. General description is also done in words: “The teacher’s approach stressed reasoning over memorization of facts.” “Most of the students held an Aristotelian conception of dynamics, while some held a Newtonian conception and this did not change during the course of instruction.”

This latter kind of general narrative reporting easily becomes hazy and it can represent unwarranted assertions about generalization within the case. Writing that “two thirds of the 24 students held an Aristotelian conception of dynamics while one third held a Newtonian conception” is a way to make “most” or “some” more specific and less hazy in narrative reporting. It is preferable, in saying that “the teacher’s approach stressed reasoning over memorization of facts,” to follow the generalization with an illustration of the kind of teaching that is meant. This illustration could be performed by presenting a narrative vignette of such teaching and then showing the relative frequency of that kind of teaching in a frequency table which shows the number of times that this kind of teaching was observed *and* which also displays the frequency of occurrence of all contrasting kinds of teaching that were observed. Combinations of general and particular description are much clearer substantively and are also more persuasive to the reader than presenting only general description, through which one sees patterns in a forest but learns nothing about the trees, or presenting only particular description, in which a tree might be exquisitely described but the reader has no sense of the forest.

### ***Orienting Commentary***

The third type of text in a narrative report is that of orienting comments. One subtype is the *interpretive or theoretical comment*, which might or might not invoke research literature: “That the Aristotelian conceptions of the students did not change is understandable because...” Another subtype of orienting comment is a summary of what has been said in a previous major section: “And so, we have seen that...” Yet another subtype of orienting comment is that of foreshadowing what is to come next in the text and of after-shadowing that which has just been presented. I think of these as “road signs.” They let the reader know where the text is going and where it has just been.

Road sign commentary is necessary at each of the junctures in the text of the report: at the beginning of a new section consisting of multiple chapters; at the beginning and ending of each chapter; and at the beginning and ending of each new section within a chapter. We can think of this as *general commentary*. Even at the beginning and ending of each unit of particular description within a section of a

chapter, some orienting comment is helpful. This can be thought of as *particular commentary*. Before and after each narrative vignette or interview quote, it is necessary to present specific orienting comments which (1) identify the substantive point to be illustrated by the particular example and (2) identify special details to which the reader should attend in the example.

### ***Writing a Whole Section in a Report***

This is analogous to stringing beads of varying sizes and shapes together into a necklace. The section would begin with general foreshadowing commentary that identifies an assertion – a substantive point, which will be illustrated in the section to come by means of units of specific description. In addition, the general commentary might outline the content to come (e.g., a sequence consisting of certain vignettes and interview quotes, then a discrepant instance, then information from a site document, and finally a frequency table which shows the typicality and atypicality of the various events and comments that were illustrated more specifically in the section).

After the general foreshadowing commentary, units of particular description follow. There might be two vignettes illustrating a typical kind of event, followed by two or three interview quotes that identify the points of view of actors in the events that were reported narratively (as noted above, before and after each of these units of particular description, brief specific commentary would be placed in order to keep the reader oriented). A vignette of a discrepant event might then be presented. Perhaps a few interview comments pertinent to the discrepant instance might follow or quotations from site documents might follow the discrepant instance. A frequency table or analytic chart, which showed how the various units of particular description fit into a more general pattern of evidence might then follow. Each successive unit of new information in the section would be preceded and followed by interpretive commentary. The section would be concluded with general commentary that reviewed the evidence and the issues that had just been presented.

Whether in a classic book-length monograph or in a journal article-length presentation, alternation of particular and general description and of particular and general orienting commentary is found in the best examples of qualitative research reporting. This feels assertive and it is. It is a new experience for beginning researchers, who might wish to try to let the story tell itself. Yet, unless the writer takes on a voice of executive commentary, actively leading the reader's attention through the text, the details of the report will not speak coherently to the reader.

### ***Short Reports***

In preparing an article-length report or an oral presentation that is limited to 15 or 20 min, there is a temptation to skip the particular description and try to tell the



whole story of the study by means of general description. In my judgment, that is a mistake. It is better to narrow the range of coverage and state a single main point in an opening few paragraphs. Then I recommend selecting a few pertinent and vivid narrative vignettes and interview quotes to present, showing the typicality or atypicality of those instances by means of general description in a frequency table, framing each of the preceding units of descriptive reporting with orienting commentary, and concluding the article with a summary substantive discussion. Such a brief report, which sacrifices breadth for depth, will show clearly a few things and be much more effective than an attempt to “tell it all” in a voice of hazy, general description.

## **Toward Better Qualitative Research**

### *Criticism of Qualitative Research from Within the Field*

Currently, qualitative research faces serious criticism, not only from “hard science” advocates external to such research, but among qualitative researchers themselves. From insiders, there has been serious criticism of an overly authoritative voice in some forms of qualitative research, particularly ethnography (e.g., Clifford and Marcus 1986; Denzin 1996). To some extent, taking care to show clear evidence for assertions mitigates these criticisms. Other critics question the entire rationalist project of research. Critical social theory shows how ideological interests that are taken for granted and thus are invisible, or are deliberately obscured drive social research. Postmodern theorists challenge the possibility of a distinction between observer and observed, subject and object.

One consequence of this criticism has been a certain loss of nerve among qualitative researchers. A more positive consequence can be found in various attempts to bring the voices and perspectives of those studied into a more prominent place in research reports. Focus on meaning from the point of view of the social actor is a hallmark of ethnography. A way to improve the quality of ethnography involves taking more care that perspectives are not misunderstood because of faulty analysis or because of the re-voicing of opinions through editorial paraphrase. In some cases, however, the attempt to highlight the “voices” of those studied has led, in my judgment, to an overreliance on interview alone as a research approach. What makes this problematic, especially in a report, is that it can mask the editorial hand of the author. An interview quote is selected by an author and placed carefully in the report. It does not have the same epistemological status, in its written form, as a comment made directly by a speaker to a hearer in an actual speech situation. Simply relying on interview data, in other words, does not resolve the power/knowledge issues raised by the critics of naive realism in qualitative research reporting. The author still maintains tremendous executive power in the construction of a qualitative research report. This needs to be clear both to the author and to the audience.

In a sense, the author always will have more power than those who are portrayed by the author. With authorial authority comes professional responsibility and a sense of this has been heightened by those critics of qualitative research who have arisen from within that work's own ranks.

Another response to the criticism that traditional qualitative research invites abuses of authority by researchers (including those of self-deception in data identification and analysis) has been for researchers to try to redress the imbalance of power by sharing it more fully with those who are studied. Both participatory action research and practitioner research are attempts to address the power/ knowledge issues involved in social research (e.g., Anderson et al. 1994).

### ***Possibilities***

Despite their limits, qualitative methods can make important contributions to science education research. Qualitative research most essentially addresses issues of the literal and metaphoric meaning of actions to social actors, while it also documents those actions in the concrete details of their routine enactment. It is the most fundamentally constructivist research method available to us. It enables us to see and understand how, in the conduct of daily life, all persons are busy, active, and making sense.

Education as a social institution is heavily invested in the notion that only some are fully making sense and are “on task,” and that others make less sense and are less active or less “motivated” (McDermott and Varenne 1995). That deep cultural belief, embedded in the workings of history which reproduces inequality in society, is manifested and reinforced so ubiquitously in the habitual conduct of teaching and learning in schools that it leads us, as educational researchers and as educational practitioners, to overlook the full diversity of ways of making sense and the full diversity of tasks – as defined by social actors – in which students and teachers are engaged. The sensitivity of qualitative research to nuances of activity and meaning in learning environments lends richness and depth to the study of the teaching and learning of science, and it is from that substantive perspective – perhaps more than the methods of research themselves – that future research in science education can benefit.

### **References**

- Anderson, G., Herr, K., & Nihlen, A. S. (1994). *Studying your own school: An educator's guide to qualitative research*. Thousand Oaks, CA: Sage.
- Bogdan, R., & Biklen, S. (1992). *Qualitative research for education*. Boston: Allyn and Bacon.
- Clifford, J., & Marcus, G. E. (Eds.). (1986). *Writing culture: The poetics and politics of ethnography*. Berkeley, CA: University of California Press.
- Denzin, N. (1996). *Interpretive ethnography: Ethnographic practices for the 21st century*. Thousand Oaks, CA: Sage.

- Denzin, N., & Lincoln, Y. (1994). *Handbook of qualitative research*. Thousand Oaks, CA: Sage.
- Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 119–161). New York: Macmillan.
- Erickson, F., & Shultz, J. (1991). Students' experience of the curriculum. In P. Jackson (Ed.), *Handbook of research* (pp. 465–485). New York: Macmillan.
- Florio-Ruane, S. (1982). The problem of dead letters. In W. Doyle & T. Good (Eds.), *Focus on teaching: Readings from the Elementary School Journal* (pp. 55–61). Chicago, IL: University of Chicago Press.
- Foucault, M. (1979). *Discipline and punish: The birth of the prison*. New York: Random House.
- Gee, J. (1990). *Sociolinguistics and literacies: Ideology in discourses*. London: Falmer Press.
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: strategies for qualitative research*. Chicago, IL: Aldine.
- Glasson, G. E., & Lalik, R. V. (1993). Reinterpreting the learning cycle from a social constructivist perspective: A qualitative study of teachers' beliefs and practices. *Journal of Research in Science Teaching*, 30, 187–207.
- Goodenough, W. (1981). *Culture, language and society*. Menlo Park, CA: Benjamin/Cummings.
- Hammersley, M., & Atkinson, P. (1983). *Ethnography: Principles in practices*. London: Tavistock.
- Lareau, A., & Shultz, J. (1996). *Journeys through ethnography: Realistic accounts of field work*. New York: Westview.
- LeCompte, M., Preissle, J., & Milroy, W. (Eds.). (1992). *The handbook of qualitative research in education*. New York: Academic Press.
- Lemke, J. (1990). *Talking science: Language, learning and values*. Norwood, NJ: Ablex.
- Lindesmith, A. R. (1947). *Addiction and opiates*. Chicago, IL: Aldine.
- McDermott, R., & Varenne, R. (1995). Culture as disability. *Anthropology and Education Quarterly*, 26, 324–348.
- Miles, M. B., & Huberman, A. M. (1984). *Qualitative data analysis: A sourcebook of new methods*. Beverly Hills, CA: Sage.
- Roth, W.-M. (1994). Experimenting in a constructivist high school physics laboratory. *Journal of Research in Science Teaching*, 31, 197–223.
- Roth, W.-M., & Roychoudhury, A. (1993). The development of science process skills in authentic contexts. *Journal of Research in Science Teaching*, 30, 127–152.
- Strauss, A. (1987). *Qualitative analysis for social scientists*. Cambridge, UK: Cambridge University Press.
- Wolcott, H. F. (1994). *Transforming qualitative data: Description, analysis and interpretation*. Thousand Oaks, CA: Sage.