Research Methodologies in Science Education: Qualitative Data

Julie C. Libarkin
Science Education Department, Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, MS-71, Cambridge MA 02138, jlibarki@cfa.harvard.edu
Josepha P. Kurdziel
Teaching and Teacher Education, University of Arizona, Education Room 703, Tucson AZ 85721, kurdziel@u.arizona.edu

INTRODUCTION

Tony, a geology graduate student, is going to be observing students in an introductory classroom in an attempt to determine how the course is affecting student learning. He is interested in determining not only the change in students' understanding of geological concepts, but also the effect of the course on student attitudes, opinions, and interactions with their peers. He has also decided to collect examples of student work and interview students at the end of the semester.

Qualitative data are both inherently different and remarkably similar to quantitative data. By definition, quantitative data consist of numbers, while qualitative data usually consist of words, either created directly by those under study, or passively by the researcher. How then can these two distinct types of data be considered similar? Fundamentally, all quantitative data are derived from qualitative decisions and qualitative data can be transformed into quantitative data. Social science (and nursing) researchers have created a number of methodologies for quantifying the qualitative.

Qualitative analysis can be used in student evaluation, course/program assessment, or basic research. Student evaluation typically includes only rudimentary qualitative analysis (such as the reading and grading of term papers) and contains some inherent assumptions. Student evaluation is designed to quantify the extent of student learning and content knowledge. Students are judged relative to each other or to a predetermined "perfect" score. Assessment of a program or course, on the other hand, is designed to determine the effect on participants, usually independent of expected outcomes. Similarly, basic research is typically focused on an open-ended question free of value judgments and, like much research in the "hard" sciences, is designed to increase the researcher's understanding of a situation or culture. The types of data gathered and the methodologies used to interpret these data will depend upon the inherent goals of the research.

WHAT ARE QUALITATIVE DATA?

Qualitative data are typically sub-divided into three components (Patton, 1990; Miles and Huberman 1994): 1) Interviews, including focus groups; 2) direct observation; and 3) written documents, including both unsolicited (classroom assignments) and solicited (surveys) documents (Fig. 1). In the following pages we will review these types of qualitative data and provide a plausible example of how these data can be gathered in your classrooms. For a closer look at qualitative research in practice, visit The Qualitative Report, an online journal that discusses research, techniques, and personal experiences. Journal contents can be accessed through http://www.nova.edu/ssss/QR/index.html.

Observing classrooms - Tony is a stratigrapher, and he has spent a number of field seasons collecting and analyzing stratigraphic data. He knows that before he enters the classroom to begin gathering data he must decide what question he is trying to answer, which types of data will help answer that question, and how he will collect data, including the tools, such as a tape recorder or video camera, he will need to bring with him. After a few weeks of observing and videotaping, Tony feels comfortable with the students in the class and is enjoying observing and occasionally helping to teach the lesson.

Observations can be made either passively or actively. Passive observation, often called direct observation, requires scrutiny without interaction (Bogdan and Bicklen, 1992). A direct observer is detached from the situation being observed, and should have no influence on the participants. Direct observation in classrooms is often carried out through the use of videotapes, although non-intrusive observation can be carried out in person if the observer takes care not to interact with students. Researchers can also immerse themselves in the context being observed by becoming a participant (Bogdan and Bicklen, 1992). Participant observation is most often conducted by anthropologists studying other societies, but immersion in a classroom as a student/observer is possible.

Words to the wise about classroom observations - take careful notes! A map of the room and a seating chart can be invaluable, and it is important to record student characteristics, (e.g., gender, age, ethnicity) that may prove useful later. Additionally, it is always a good idea to type out handwritten notes or transcribe audio or

videotapes as soon after the observation as possible. This provides an opportunity for the researcher to clarify hastily scribbled notes, or to recall incidents outside the range of mechanical recorders. Finally, students often view observers as an additional resource, especially in laboratory settings. In Tony's case, the course instructor thought of Tony as a resource, rather than a passive observer. It is important to explain the role of the observer to instructors, and perhaps students, prior to entering the classroom. When encountering questions from students, we as observers have found that a shrug of the shoulders and a suggestion to "ask your professor" results in minimal observer-participant interaction.

Interviewing - During final exams Tony arranges to interview five students. He has three questions he wants to ask:

1) Did you find the course stimulating and would you take another course in this subject area?

2) What would you change about the course content?

3) What would you change about the instructor's teaching style or other course features?

There are two types of interviews that are generally used by the education community: focus groups and one-on-one interviews. Many faculty use focus groups in formative evaluation, such as mid-semester focus groups designed to solicit student opinion about new teaching innovations. Researchers may use focus groups as either a primary mechanism for gathering data or as a starting point from which a larger research agenda will be designed and clarified. Focus groups are useful because they typically take a short amount of time to prepare and execute and allow a significant number of participants to voice their opinions and ideas (Ponsford and Masters, 1998). Interviews, on the other hand, are much more time consuming, often requiring several hours for a single interview. Additionally, the data entry associated with interviews is much more time-intensive than for focus groups. As a result, focus groups are more frequently utilized in formative evaluation, although interviews are a very common research tool.

Several key "do's and don'ts" for conducting focus groups and interviews are useful to discuss. First and foremost, as with questionnaires, researchers must be careful to ask only *singular* questions (Patton, 1990; Seidman 1998). A singular question prompts the interviewee for only one response. That is, asking a student, "*Did you find the course stimulating and would you take another course in this subject area?*" may result in confusion and missed data. On the other hand, asking a student, "*Did you find the course stimulating?*" and "*Would you take another course in this subject area?*" will result in much clearer responses, especially if you are trying to uncover cause and effect. Secondly, focus groups, by their very nature, can yield valuable information about group dynamics and peer interactions. As such, it is important to allow wiggle room; be prepared, know the basic questions that you wish to cover, but allow freedom for new ideas to evolve. Similarly, interviews can be unstructured, with the same benefits as focus groups. However, it is important to have common ground in all one-on-one interviews to ensure that interviews will be useful in a synthesizing analysis. We have found it valuable to have students complete a short questionnaire at the start of each focus group or interview. This questionnaire can then be used to guide the discussion.

Written Documents - To augment his observations and interviews, Tony has asked the instructor to make copies of some of the term papers written by the students. He collects eight papers, five from the beginning of the semester and three from the end, all written by different students. The names of the students were removed from the copies to protect their privacy.

When considering the analysis of written documents, most classroom researchers immediately think of student-generated work. Indeed, the analysis of exams, term papers, homework, notes, and other materials can provide a rich and complex view of a course and its effects. It is also important to consider materials created by the instructor. The analysis of syllabi, assignments, and exams is especially important since these materials initiate student work. Any questionnaires or surveys solicited from students are also valuable sources of information. The questionnaires themselves should be reviewed carefully; again, student responses will always be constrained by the questions you ask! Finally, if written documents are going to be used in conjunction with observations or interviews, it is important to have a method for connecting the different types of data. In Tony's case, he would have collected the most useful information if he had asked each of his interviewees to provide him with papers written over the course of the semester.

What is a Case Study? - Social science researchers use case studies to document specific details about an individual, group, or institution (Patton, 1990; Miles and Huberman 1994). These details can be generalized to a larger group, and can provide new insights that are quite different from those generated by broader studies. Case reports are valuable, as they tend to engage the reader in the unique perspective of a single individual or group. These studies are composed of as much data as possible, and generally include myriad interviews, observations, and other data. Additionally, cases studies are often longitudinal in nature; that is, they are conducted over an

extended period of time, such as over the course of a semester or academic career. Finally, a case study of a program or group may be composed of individual-scale case studies. This layering adds rich context to a qualitative study which is absent in more generalized qualitative methods. Although not necessarily involved in all case studies, many researchers expand the impact of the case study process by comparing findings between cases using the analysis techniques described below.

Analyzing Qualitative Data - To find order in the myriad layers of qualitative data generated by the methods described above, a field of research known generally as *content analysis* was developed. Content analysis focuses on the search for patterns within a text, be it field notes of a classroom observation or a student term paper. Three types of analyses are common in qualitative research: *thematic content analysis*, where themes are extracted from the text, *indexing*, where specific words are viewed in context, and *quantitative descriptive analysis*, or word counting.

Thematic Content Analysis - Thematic content analysis is the most subjective of the qualitative analytical techniques, although with practice biases can be eliminated. In conducting a thematic analysis, themes are usually identified inductively, such that themes emerge naturally from the data (Patton, 1990; Denzin and Lincoln, 1998). Inductive analysis limits bias, as the researcher is not imposing external themes on the data; however, with practice, predetermined thematic analysis can be a quick method for analyzing data. Once dominant themes have been identified in the data through open coding, the researcher links and reorganizes themes in an attempt to develop a dominant structure. This structure will eventually evolve into a conceptual framework of the system under study. Quotes and anecdotes are useful as examples of the types of data that led to the extraction of themes and connections, and should be used to bolster arguments (Golden-Biddle and Locke, 1997).

Indexing and Quantitative Descriptive Analysis - Indexing and quantitative descriptive analysis (QDA) are additional methods that can be used to quantify qualitative data (Trochim, 2001). Both techniques involve documenting word occurrences, although indexing is primarily concerned with the context in which words exist. Context is usually defined as the words that immediately precede and follow the targeted term. QDA, on the other hand, reports the frequency with which words are used throughout a text, independent of a specific context. Computer programs have been developed for conducting these types of analyses, although coding by hand is certainly possible. Qualitative data can always be transformed into quantitative data, although it is not necessarily desirable to do so. Converting the qualitative to the quantitative strips the data of the context in which they occur, and any data transformations should be made cautiously.

Interpreting Qualitative Data - Unlike quantitative research, the interpretation of qualitative data is essentially unbounded. Qualitative researchers are faced with an infinite number of variables, and must be able to ferret out dominant themes and effects. It is ultimately up to the investigator to define cause and effect, and up to other researchers to evaluate study conclusions. Eisner (1991) suggests three criteria that are useful in deciding if a qualitative interpretation is relevant. Researchers should ask themselves: 1) Does the "story" laid out by the original data, analyses, and interpretations make sense both within the context of the study and from an outsider's point of view?; 2) Do multiple data sources tell the same story, and if not, can this be explained? (The use of multiple data sources is called triangulation.); and 3) Does the study explain a situation in a useful way; that is, is the study a valuable addition to the existing body of knowledge? As a reader, a qualitative study should be judged carefully, always with a critical eye turned towards validity and reliability.

Validity and Reliability - As with any type of research, the validity and reliability of a qualitative study must be established if the results of the research are going to be meaningful to others. While a strict set of statistical guidelines can be used to judge the appropriateness of quantitative research, qualitative research is usually not testable by standard statistical means. Some qualitative researchers would argue that validity and reliability only have meaning within the narrow frame of reference of the subject under study, and therefore these criteria can never be satisfied in qualitative research. Taking a less strict approach to the issue of validity and reliability, Lincoln and Guba (1985) established a set of four measures that can be used: credibility, transferability, dependability, and confirmability. Although just one point of view, we find these criteria to be very useful in bolstering the strength of our qualitative studies, analyses, and conclusions.

Credibility - Because qualitative research is focused on studying an individual or entity (such as a program), the subject's perspective on the project results can be used to establish validity. The participants themselves can verify interpretations made about their attitudes and opinions, as well as inferred causal relationships.

Transferability - One of the ultimate goals of all research is the establishment of the value of the research outside of the immediate context under study. The ability to apply conclusions to other settings depends strongly upon the investigator's ability to document the setting of the original study, giving subsequent researchers an oppor-

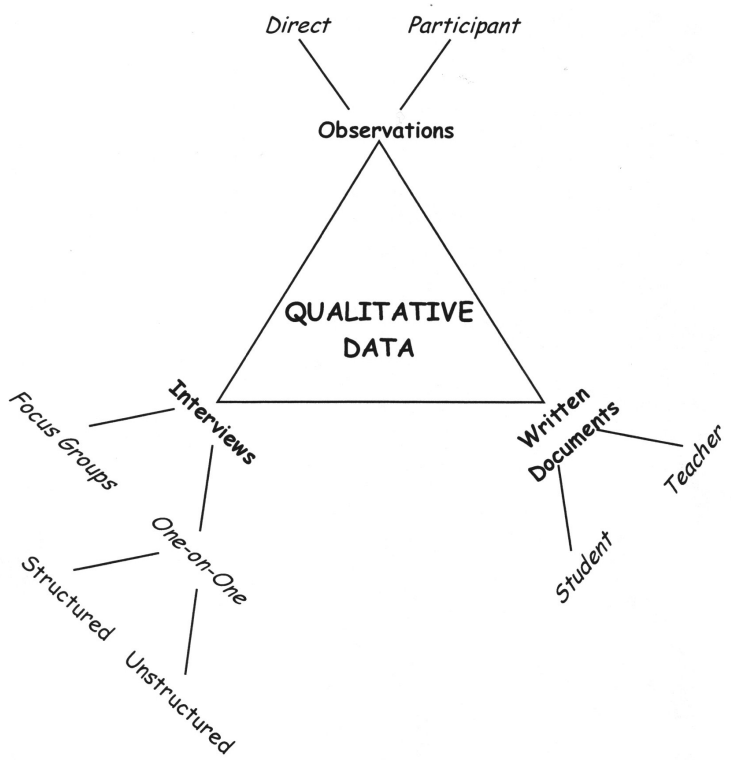

Figure 1. Schematic diagram of qualitative data, including the three main types, and subcategories.

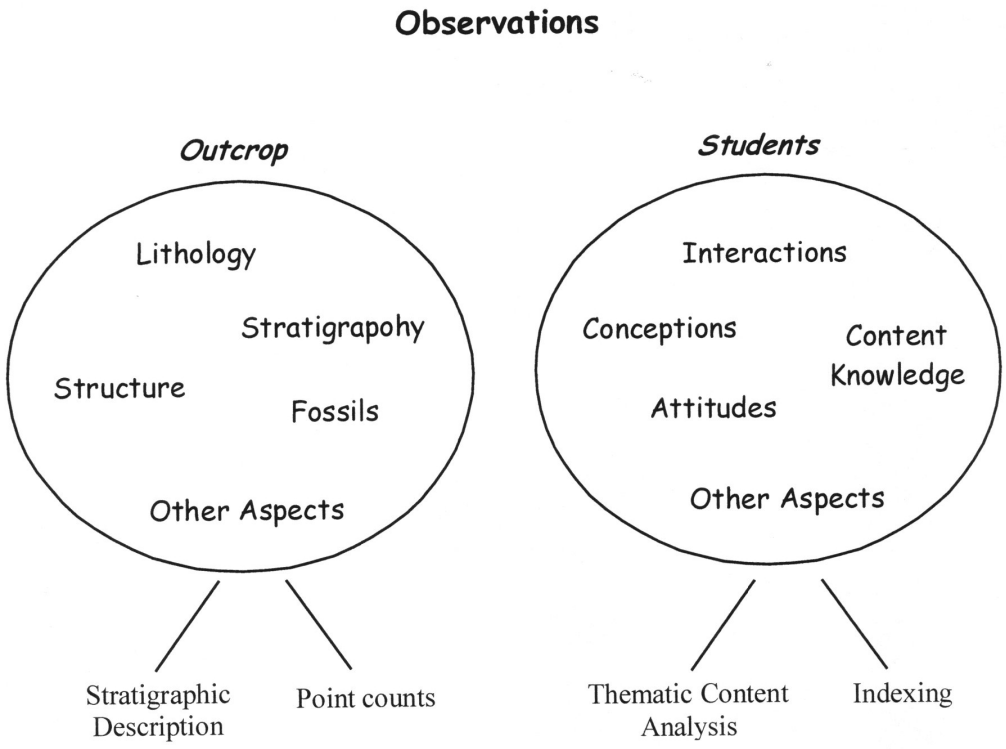

Figure 2. Comparison of techniques used in stratigraphic analysis with content analysis of qualitative data. Terms in the circles represent characteristics of an outcrop, for stratigraphy, and students, for qualitative analysis. In this analogy, stratigraphic description and thematic content analysis have common characteristics, as do point counting and indexing.

tunity to identify all of the variables inherent to that context. The final decision about the transferability of a study, however, ultimately falls on the shoulders of future researchers and is outside the control of the original investigator.

Dependability - The reliability of a study depends upon three factors: repeatability, stability over time, and similarity between measures (Kirk and Miller, 1986; Hoepfl, 1997). In qualitative research, one could argue that reliability can almost never be achieved, as subjects will certainly change over time and identical data can never be collected twice (Lincoln and Guba, 1985). However, a careful review of the process of data collection and the research product, especially as time evolves, can help establish reliability. For instance, comparison of observations carried out by two different researchers, as well as their analyses, can help determine the dependability of the qualitative process and study conclusions.

Confimability - A certain degree of objectivity is required in any scientific study, although perfect objectivity is certainly never attainable, even in the "hard" sciences. Science is always subject to some degree of disagreement and controversy! In qualitative research, the investigator must be able to demonstrate that interpretations are free of subjectivity, and that potential biases have been controlled. Keeping careful records, including original notes, transcriptions, and analyses will allow other researchers an opportunity to review the interpretation process. Although this type of review rarely occurs, it is useful to get into the habit of leaving a paper trail, just in case the day arrives when you make an interpretation that goes against the dominant paradigm!

QUALITATIVE DATA IN THE SCIENCES

Although most scientists are used to thinking about data quantitatively, the geosciences actually has a long history of using data and analytical techniques which are similar to those described above. For instance, paleontologists spend a significant amount of time describing (i.e. observing) the features of fossils, and then determining the fossil's systematic position. Students in paleontology classes can often be found divided into two camps: "lumpers" and "splitters", where splitters see many species while lumpers see few. In some ways, this methodology is similar to thematic content analysis.

Stratigraphy offers another analogy with qualitative data analysis (Fig. 2). Stratigraphers spend a significant amount of time observing outcrops and taking field notes. All features of the outcrop are documented, usually including sketches and any detail that may prove to be significant later. As discussed earlier, observers in a classroom should make similar sketches and note as much detail as possible. Additionally, stratigraphers and qualitative observers both approach field excursions with 1) an idea of the question being asked; and 2) a plan for the type of data that will be collected (Fig. 2). The analytical techniques used by stratigraphers also mirror some qualitative techniques. Stratigraphers typically do not publish field notes, but rather provide a synthesis of their observations. For instance, descriptions such as "a ten meter section of 1.0-0.5 m thick coarse sandstones interbedded with 0.5-0.1 m thick fine-grained shales" are common. This description provides a conceptual model of the observed outcrop by providing thickness ranges and using well-understood terminology, such as coarse and fine-grained. However, it is quite probable that the sandstones contain some larger or smaller grains; this detail is not vital for the ultimate interpretation of the outcrop's history, and is therefore left out. Similarly, coding of observations or other materials using content analysis removes extraneous information and highlights the most important themes.

CONCLUSIONS

Qualitative research can be a source of contextually rich information, especially when a study is carefully planned and executed. Tony's approach to qualitative research was quite good, and we saw only a few places for added improvement. To help you in your research endeavors, we have created a checklist to aid you in your qualitative design (modified from Lincoln and Guba, 1985):

A. Define your research question - What are you trying to determine with this research? Essentially, define your study boundaries. At this point you may not know much about your area of research, so keep in mind that the question may change later. Additionally, make sure from the very beginning that qualitative data, rather than quantitative, is going to assist you in achieving your research goals.

B. Conduct a literature review - Has this question been asked by other researchers? How much research has been published in this area of study? What can you do that will be unique and different? Researching a well-understood question is uninteresting unless you can look at it from a new angle. Make sure to look in professional journals, books, and convention proceedings (Libarkin and Kurdziel, 2001).

C. Clarify your research question - Based upon your literature review, rewrite the research question to a) incorporate new information you may have learned from the review; and b) ensure that you are asking a unique question.

D. Define your research plan - What methods will you use to answer the research question? Will you use several different methods? If necessary, what will your control group be? For instance, if you are looking at the effect of student collaborations on learning, you may want to choose one additional class to study that is strictly lecture-oriented. Determine the order in which data will be collected. You may decide to conduct short interviews to gather preliminary data, and then focus your study at a later stage. Also, decide if you will be using any equipment, such as a tape recorder, and make sure you have a plan for accessing that data. Transcribing a tape manually is time consuming, so make sure you have the time or can hire someone before collecting the data. Additionally, make sure to include a plan for data analysis, including the methodologies that will be used and the people who will be conducting the analyses. After your research plan is in place, submit a copy to your institution's Human Subjects Committee (most institutions require Human Subjects approval prior to conducting any research that involves human subjects.)

E. Validity - After completing the above steps, review your research proposal with an eye towards the four criteria of validity discussed above.

F. Share your research agenda with another person Other people can often add new insight and a fresh perspective that may be useful for further clarification of your research idea.

G. Conduct the research! - Remember that as you enter the preliminary stage of your research you may want to modify your research plan. This modification must be done carefully so that you will still be able to use your earlier research. For instance, you may begin a project by observing students and then decide that interviewing would be more useful. If you continue with your observations and simply add in interviews, then you will have a basis for comparing your earlier work (observations) with the later work (observations + interviews).

H. Assess how well you have answered the research question - This should be done throughout the course of the investigation. You may find that your research has answered an entirely different question than the one you initially posed or that the answer is unexpected. Once you believe you have completed the research, you can stop. Often a new set of questions will evolve out of this research, giving you a basis from which to start your next research project.

I. Report your results - It is important to share your research with both the science and education communities. This can be done at national science and/or education meetings, or in professional journals. The experience of writing a qualitative research paper is somewhat different from hard science writing and we will cover this issue in a future column. In the meantime, use the Journal of Research in Science Teaching as a guide!

REFERENCES

Bogdan, R.C., and Biklen, S.K., 1992, Qualitative Research for Education: An Introduction to Theory and Methods: Needham Heights, MA, Allyn and Bacon, 276 p.

Denzin, N.K., and Lincoln, Y.S. (editors), 1998, Strategies of Qualitative Inquiry: Thousand Oaks, CA, Sage Publications, 346 p.

Eisner, E.W., 1991, The enlightened eye: Qualitative Inquiry and the Enhancement of Educational Practice: New York, Macmillan Publishing Co., 264 p.

Golden-Biddle, K., and Locke, K.D., 1997, Composing Qualitative Research: Thousand Oaks, CA, Sage Publications, 145 p.

Hoepfl, M.C., 1997, Choosing qualitative research: A primer for technology education researchers: Journal of Technology Education, v. 9, p. 47-63.

Kirk, J., and Miller, M.L., 1986, Reliability and Validity in Qualitative Research: Beverly Hills, Sage Publishing, 87 p.

Libarkin, J.C., and Kurdziel, J.P., 2001, Research Methodologies in Science Education: Strategies for Productive Assessment: Journal of Geoscience Education, v. 49, p. 300-304.

Lincoln, Y.S., and Guba, E.G., 1985, Naturalistic Inquiry: Beverly Hills, Sage Publishing, 416 p.

Miles, M.B., and Huberman, A.M., 1994, Qualitative Data Analysis: An Expanded Sourcebook, 2nd Ed.: Thousand Oaks, CA, Sage Publications, 338 p.

Patton, M.Q., 1990, Qualitative Evaluation and Research Methods, 2nd Ed.: Newbury Park, CA, Sage Publications, 532 p.

Ponsford, B.J., and Masters, L.A., 1998, How to use focus groups in higher education research: College and University, v. 73, p. 2-9.

Seidman, I., 1998, Interviewing as Qualitative Research: A Guide for Researchers in Education and the Social Sciences, 2nd Ed.: New York, Teachers College Press, 119 p.

Trochim, W.M.K., 2001, The Research Methods Knowledge Base, 2nd Ed.: Cincinnati, OH, Atomic Dog Publishing, 363 p. (Http://trochim.human.cornell.edu/kb/index.htm)

Note: The authors are supported by National Science Foundation Postdoctoral Fellowships in Science, Mathematics, Engineering, and Technology Education (grants DGE-9906479 and DGE-9906478)