

Using qualitative methods to make and support claims in physics education research

Cody Sandifer
Towson University
csandifer@towson.edu

Andy Johnson
Black Hills State University
andyjohnson@bhsu.edu

This session was designed to stimulate conversations about the use of qualitative methods in physics education research. To start the session, we presented a general overview of qualitative research. Then, to provide a context for discussion, we conducted a mini research activity; in this activity, we introduced data (interview, video transcripts, and student work) from a university physics course for preservice teachers. Participants were given the task of examining the data and deciding whether a particular claim was sufficiently supported by the data. A rich discussion ensued, in which many research-related issues were raised. These issues, which we suggest might serve as topics of discussion for future sessions, are listed and briefly editorialized at the end of this paper.

Part 1: Overview of Qualitative Research

In simplest terms, qualitative research is a type of research that answers research questions by making highly believable arguments based on non-numerical data. This is in contrast to quantitative research, which instead utilizes numerical data and statistical tests of significance to accept or reject statistical hypotheses within a given level of statistical confidence.

Our understanding of qualitative research can be deepened by recognizing that qualitative studies share a number of characteristics that go beyond the simple "non-numeric" characteristic described above. That is, qualitative research [1]

- Recognizes the importance of educational context and background in analyzing and interpreting data
- Attempts to understand the world (physical phenomena, the learning environment, etc.) as viewed by the participants, rather than as viewed by the researchers
- Often focuses on both the learning process and learning outcomes
- May involve small numbers of subjects

In other words, the qualitative researcher views physics learning as a complex, ongoing evolution of ideas that can be influenced by such factors as teacher expectations, the student's prior knowledge and experience, the student's attitude towards science and science knowledge, the physical set-up of the classroom, and a host of other student-, context-, and/or teacher-related factors.

Another characteristic of qualitative research that follows from the recognition of the complexities of the learning process and learning environment is that, in qualitative research, the formulation of data categories and identification of important factors is not done *a priori*, as is often done in quantitative research. Rather, in qualitative research, category formulation and factor identification is carried out during the very process of analysis, at which time the categories and factors emerge as the researcher carefully evaluates and re-evaluates the research data [2].

What form does this non-numerical research data take, and how is this data obtained? Given the qualitative researcher's need to investigate students' thoughts, concepts, and perceptions, and the relationship of these intangibles to various factors in the learning environment, the qualitative researcher requires data in the form of

(a) students' out-of-class self-reflections (on themselves, their knowledge, their learning environment, etc.), which are obtained through face-to-face interviews, (b) students' in-class questions, comments, discussions, and interactions, which are obtained through classroom observations, and (c) student work, which is photocopied and kept on file.

Consequently, when the qualitative researcher presents his or her highly believable argument in the form of a research article, it is natural for the article text to include detailed descriptions of the learning environment, direct quotations from students and teachers, and samples of students' written work. This is done so that the reader can make informed judgments as to whether the conclusions of the qualitative research are both valid and complete.

We believe that qualitative methods, because of the depth of information gathered, can provide meaningful answers to complex questions; however, there are two possible problems with the validity of qualitative data: (1) students' perceptions may be inaccurate, and (2) the researcher's perceptions of the interview and classroom data may also be inaccurate. For example, while students' words can reveal insights into their emotions, the ways in which they have organized their world, their sensory perceptions, their thoughts about what is happening, and their past experiences [3], students' words can also reveal an emotion that isn't really there, the ways in which participants have inaccurately organized their world, incorrect sensory perceptions, inaccurate thoughts about what is happening, and misremembered past experiences -- either because the students incorrectly self-reported these things, or because the researcher misinterpreted the students' originally accurate self-reports.

Three different methods are recommended to avoid the danger of inaccurate perception of the data [4]: (1) data triangulation, which is the use of multiple data sources (e.g., the use of both written work and student interviews) to verify the accuracy of the researcher's interpretation of the data, (2) keeping an open but critical mind to different interpretations of the data, and (3)

obtaining feedback (if possible) from the research participants as to whether the researcher's interpretation of the data is accurate and complete.

The problems with self-reporting, and the subsequent need for validity-control methods such as data triangulation and keeping an open but critical mind, can be seen in two examples -- one relating to everyday life, and the other relating to science education in the classroom.

Example 1. One day, the four-year-old son of one of the authors (CS) overheard his parents discussing housing prices, after which the young boy commented that "Wow, \$200,000 is a lot of money!" The author's response to his son's astute observation was one of intense pride: "Yes," the author answered, "\$200,000 is a lot of money!" It was only when the son left to retrieve \$200,000 from his piggy bank that the author realized that his son's understanding of large sums of money was not quite the understanding that he had assumed -- and that perhaps the author should have kept a more open mind when considering his son's understanding of two hundred thousand dollars.

Example 2. When asked, students in the author's (CS's) early childhood methods courses are often able to provide the standard definition for science: science involves hypotheses, predictions, and so forth. Yet, when asked to actually "do" science, many of the students tend to show scientific demonstrations rather than perform scientific experiments. For instance, a student might throw a marble across the room as an example of gravity, rather than perform an experiment to determine how certain variables affect the motion of the marble projectile. Based on these actions, it is obvious that the students' self-reports concerning their understanding of science are not fully accurate, and that the request for these students to actually "do" science is a necessary triangulation that reveals the mismatch between these students' self-reports and their actions.

Part 2: Mini Research Activity and Discussion

The purpose of the second part of the session was for the session participants to contribute their own expertise to a discussion of qualitative methods.

To achieve this goal, we felt that the group needed a common "qualitative" experience to which everyone could refer. We provided this common experience in the form of a mini research activity.

In this activity, the presenters handed out nine pages of video and interview transcripts and student work from a study one of us (AJ) did on learning in magnetism. We introduced the course setting (an inquiry course on static E&M), and made the claim that, initially, one of the students "thought that uncharged metals are negative". The session participants were then asked to decide whether the data they looked at were sufficient to support that claim.

We chose this approach for two reasons. First, we wanted to spark a group discussion about what researchers can do to assure the validity of their data and of the claims they make, and so it was appropriate to introduce a research claim that may or may not be adequately supported by a set of data. Second, we researchers rarely get to see the analysis process unfold in another researcher's work -- we only get to hear the final results in papers and presentations. Therefore, we thought that, by referring to a common data set, everyone could benefit from a nuts-and-bolts discussion aimed at evaluating a research claim.

During the research activity, the session participants read through the transcripts, talked in small groups about their concerns, and recorded questions that they wanted to bring up for whole-group discussion. Once the small groups finished their analysis and discussion, everyone participated in a whole-group discussion of the group questions. Many issues were raised (listed below). We think that these constitute some of the big issues facing PER today.

Issues raised in the discussion:

Researchers need the classroom context. Many of the session participants said that they needed more information about the students and about the particulars of the class. They felt that the data provided (nine dense pages of descriptions, transcripts, and student work) were insufficient -- either to make sense of what was going on, or to support the prototype claim that had been offered.

Many participants in the session wanted more information on what the students were like, what the class was like, what other students were doing and so on to be convinced that the claim being made was the best choice. Without that contextual background, many of the participants did not agree that the proffered claim was well-supported in the nine pages of data.

This suggests that it is important for the researcher to be very familiar with the day-to-day activity in a classroom. A researcher who is immersed in the setting can take cues from everything that he or she notices. This seems to be necessary for making dependable interpretations of students' thinking as expressed in transcripts and for being convinced of conclusions.

How you collect data makes a difference. Researchers must balance the need for accurate and sufficient information with the simultaneous need to not influence the subject. How do you do this, particularly in interviews? Should one set a rigid interview protocol and stick with it? Is it a good idea to stray from the protocol to find out more about something that sounds interesting? Or would this be likely to influence the subject unduly? This topic needs much more discussion.

Many participants in this session, however, seemed to agree that it is very difficult to give a good interview. It takes time to learn how to talk and how to listen. Worse, you can't pay attention to every possible interpretation of the subject's statements. The presenters believe that interviewing is to be approached with care and humility.

Stability of students' thinking: Researchers want multiple pieces of evidence to make reliable claims that can be supported. This is triangulation. However, if you want to make claims about what students know or are thinking during a course, you are aiming at a moving target. In particular, the session participants were concerned as to whether the student in the study really had a stable concept of negativity -- perhaps her idea was evolving?

This problem is compounded by the question of the coherence of students' thinking, as well as whether we view students' thinking as knowledge

in pieces or as coherent systems. It may be that sometimes we can see students' thinking being stable and coherent long enough that we can make a claim about it as if it were consistent, and other times we won't see enough coherence to make that kind of claim.

The researcher's agenda and beliefs influence the selection of data and the findings. Many participants in the session noted that because the presenters offered a claim to be tested, their views of the data were tainted. Because they had been told that "one of the students thought that negative meant uncharged," the session participants started right in looking for evidence to support or disprove that claim. Some said that, because of this, they didn't pay attention to other issues available in the data.

This may be an ever present issue in PER. Not only does the researcher's agenda color how he or she analyzes the data, but it even determines which data are considered important. This issue is unavoidable, because it is impossible to collect a complete set of data on everything that happens in a classroom, and in analyzing data one must focus on only parts of it or never finish the analysis.

What is sufficient data? This important issue did not receive enough discussion. Some important questions raised were: What does sufficiently supported data look like? How do you make a sufficiently precise claim that can be supported by data?

This is a problem with reporting research results in talks and papers. It may be relatively easy to provide data to show that a particular claim is plausible, but it may not be possible to show convincingly that a particular claim is highly believable or the best choice given the data.

Final thoughts:

While many issues were raised by this session, none were discussed sufficiently or settled. With regard to doing PER, much remains to talk about. We hope that conversations begun in this session continue in the future.

References

- [1] Patton, M. P. (1980). Qualitative evaluation methods. Beverly Hills, CA: Sage.
- [2] LeCompte, M. D., & Preissle, J. (1993). Ethnography and qualitative design in educational research, second edition. San Diego: Academic Press.
- [3] Patton, M. P. (1980). Qualitative evaluation methods. Beverly Hills, CA: Sage.
- [4] Creswell, J. W. (1994). Research design: Qualitative and quantitative approaches. Thousand Oaks, CA: Sage.

Additional Readings

- Bogdan, R. C., & Biklen, S. K. (1992). Qualitative methods for education: An introduction to theory and methods. Boston: Allyn & Bacon.
- Erickson, F. (1986). Qualitative methods in research on teaching. In Handbook of Research on Teaching, 3rd edition (pp. 119-161). New York: Macmillan.
- Erickson, F. (1998). Qualitative research methods for science education. In B. Fraser & K. Tobin (Eds.), International Handbook of Science Education, vol. 2 (pp. 1155 - 1174). Dordrecht: Kluwer.
- Lincoln, Y. S., & Guba, E. G. (1995) Naturalistic inquiry. Beverly Hills, CA: Sage.
- Marshall, C., & Rossman, G. B. (1989). Designing qualitative research. Newbury Park, CA: Sage.
- Merriam, S. B. (1988). Case study research in education: A qualitative approach. San Francisco: Jossey-Bass.
- Millroy, W. L., Preissle, J., & LeCompte, M. D. (Eds.). (1997). The handbook of qualitative research in education. San Diego: Academic Press.
- Tesch, R. (1990). Qualitative research: Analysis types and software tools. New York: Falmer.