Validity in qualitative research: Application of safeguards[1]

Katrina M. Daytner

Western Illinois University

[1] Presented at the 18th annual meeting of the Ethnographic and Qualitative Research in Education conference, Cedarville, OH, on June 10, 2006. Please address correspondence to Dr. Daytner at this address: Western Illinois University, EIS Department, 115Q Horrabin Hall, Macomb, IL 61455. E-mail: KM-Daytner@wiu.edu.

Abstract

The construct of validity has received considerable attention in qualitative methods literature (Denzin, 1989; Erickson, 1986; Geertz, 1973; Goetz & LeCompte, 1984; Howe & Eisenhart, 1990; Maxwell, 1992; Smith & Glass, 1987). Much of the attention has been focused upon the issue of whether qualitative results and interpretations accurately reflect the meanings, perceptions, and beliefs of the participants, programs, and settings about which they are written. This conceptual paper will describe six different safeguards designed to increase the accuracy and validity of qualitative results. These safeguards include: triangulation, establishing a data trail, acknowledging researcher subjectivity, member checks and participant review, prolonged engagement, and consideration of disconfirming evidence and contradictory interpretations. The author will discuss the application of these safeguards to a qualitative case study.

Validity in qualitative research: Application of safeguards

Validity, defined as the trustworthiness of inferences (Eisenhart & Howe, 1992; Lincoln & Guba, 1985), has received considerable attention in qualitative methods literature (Denzin, 1989; Erickson, 1986; Geertz, 1973; Goetz & LeCompte, 1984; Howe & Eisenhart, 1990; Smith & Glass, 1987). Much of the attention has been focused on whether qualitative results and interpretations accurately reflect the meanings, perceptions, and beliefs of the participants, programs, and settings about which they are written.

In a case study research project investigating a teacher's instructional practices in language arts in a rural kindergarten classroom affected by poverty, six safeguards were implemented to increase the validity of the project's results. The purpose of this paper is to explore the issue of validity by describing the six specific safeguards and how they were applied in the case study project.

Theoretical Framework

The trustworthiness of data interpretations are directly related to the methodological and analytical processes of the project. Researchers have identified a number of specific strategies, or safeguards, to ensure that the methodological and analytical steps produce accurate results (Denzin, 1989; Lincoln & Guba, 1985; Mabry, 1998; Merriam, 1998; Peshkin, 1988; Stake, 1995). The six safeguards to be addressed in this paper include (a) triangulation, (b) establishing a data trail, (c) acknowledging researcher subjectivity, (d) member checks and participant review, (e) prolonged engagement, and (f) consideration of disconfirming evidence and contradictory interpretations.

Triangulation

Triangulation involves using multiple data sources, investigators, theories, or methods, to confirm a warranted interpretation or conclusion. The idea is that the trustworthiness of inferences is increased when multiple examples of support are available. One of the most commonly quoted sources regarding the use of triangulation in qualitative research is *The Research Act* by Norman Denzin (1989). In his book, Denzin articulates four types of triangulation for enhancing the credibility of qualitative research. These are (a) source triangulation, (b) investigator triangulation, (c) theory triangulation, and (d) methodological triangulation. Source triangulation involves the use of multiple data sources. Researchers who use source triangulation include several different participants, programs or settings in trying to understand a phenomenon. For example, a researcher interested in understanding the context of a classroom might interview the classroom teacher, the students in the classroom, as well as the building administrator, and then compare these multiple perspectives to develop one written description of the context. Investigator triangulation involves using multiple researchers for collecting and analyzing data. By having more than one researcher collect and analyze the data, comparisons can be made to determine the consistency of results across multiple people. Theory triangulation utilizes multiple theories and perspectives for understanding qualitative data. When using theory triangulation, researchers attempt to understand and interpret research findings from multiple theories, such as using Piaget and Vygotsky to interpret a child's cognitive development. Finally, methodological triangulation incorporates within-method and between-method types. Within-method triangulation involves using different data collection modes, such as observation, interview, and document analysis while between-method triangulation involves using different designs, for example case study and survey methods, to increase the validity of the results and interpretations.

Establishing an Audit Trail

Lincoln and Guba (1985) provide a detailed description of Halpern's (1983) concept of external audit. The external audit consists of two parts, the audit trail and the audit process. The audit trail involves making the written data (both raw data and interpretations) available for review by others while the audit process involves a 5-step evaluation of the data to ensure the results can be confirmed. The audit trail consists of six categories (a) raw data, (b) data reduction and analysis products, (c) data reconstruction and synthesis products, (d) process notes, (e) materials relating to intentions and dispositions, and (f) instrument development information. The five steps of the audit process include (a) preentry, (b) determination of auditability, (c) formal agreement, (d) determination of trustworthiness, and (e) closure. The presence of an external audit brings integrity to the study by making the data available for inspection by others.

Acknowledging Researcher Subjectivity

Subjectivity has been viewed as both a strength and a weakness of qualitative research. Subjectivity is the influence of a researcher’s experiences, knowledge, training, and emotions on research. Many quantitative researchers, such as Campbell & Stanley (1963), view subjectivity as one of the main weaknesses of qualitative research. Subjectivity is viewed as the inserting of value judgments and researcher bias into research, resulting in distorted "truth" about phenomena and constructs. In contrast, many qualitative researchers view subjectivity as a potential asset of qualitative research. Subjectivity is seen as a useful and personal quality of a researcher, resulting in unique insights into the understanding of a phenomenon. Jansen and Peshkin (1992) provide a useful review of the various qualitative perspectives related to subjectivity.

Like many qualitative researchers, I believe subjectivity is an asset to qualitative research if it is acknowledged and shared in research reports. Responsible qualitative researchers monitor their subjectivity throughout the research process. By providing readers knowledge about the processes and decisions utilized during a research endeavor, qualitative researchers give their readers a sound knowledge base for evaluating the usefulness of their research. Because of my life history and professional training, I provide a unique perspective on understanding phenomena. I believe my readers need to be aware of this unique perspective.

Just as debate exists surrounding the utility of subjectivity in qualitative research, there is little consensus as how best to deal with the topic of subjectivity. Peshkin (1988), who is arguably one of the most influential writers on this topic, argues for a formal, systematic monitoring of one's subjectivity throughout the research process. He believes the results of this monitoring should be shared in research reports so that readers can make informed decisions about what they have read.

Member-Checking and Participant Review

Member-checking involves seeking feedback from representatives of the stakeholding groups involved in or affected by an investigation (Lincoln & Guba, 1985). This process involves asking a selected group of participants to review and respond to data and interpretations. Generally, not all participants have the opportunity to be involved in the review of the research. Mabry (1998) has offered a more comprehensive two-step process, which I call participant review, that provides each informant an opportunity for critique. In the first stage, all participants are offered copies of relevant data that has not been interpreted, such as interview transcripts or observation field notes, to correct errors or to provide additional information that might improve the accuracy of the data. This stage is designed to increase the descriptive validity

of the results. The second stage, building from the first, involves sending a draft of one's initial interpretations to participants for their reactions. This stage is designed to increase the interpretive validity of the results.

Prolonged Engagement

Research conclusions are more credible if they are the result of repeated observations suggesting the same findings (Denzin, 1989; Lincoln & Guba, 1985; Merriam, 1998). Prolonged engagement is the investment of a sufficient amount of time in the research setting. Determining how much time is sufficient depends upon the setting. Lincoln and Guba (1985) argue that the amount of time can be determined by ensuring that certain purposes are achieved. These purposes include (a) learn and understand the culture (context), (b) test for misinformation introduced by distortions of the researcher or the respondents, (c) build trust and rapport with informants, and (d) identify those characteristics and elements in the situation that are most relevant to the problem or issue being investigated.

Consideration of Disconfirming Evidence and Contradictory Interpretations

The researcher who does not seek disconfirming evidence can be accused of searching only for information that strengthens interpretations (Erickson, 1986). Disconfirming evidence and contradictory interpretations should be sought and critically appraised so that inadequacies within the original interpretations can be made (Eisner, 1991; Erickson, 1986). As Eisner (1991) states, this does "…not mean that educational critics are obliged to provide readers with every possible interpretation and appraisal of a situation they write about" (p. 111). Rather, he suggests that researchers, in order to be fair to their readers, should consider alternative interpretations that are reasonably credible.

Dealing with Mismatch Case Study

The project used to apply the six safeguards described above used qualitative case study as its methodology. The project drew from an interpretative orientation. The goal was to understand, explain, and interpret through "thick description" (Geertz, 1973) the phenomenon of interest within its natural setting, embedded within a culture, through the voices of the people involved. Two ethnographic principles, thick description and progressive focusing, were used to guide the collection and analysis of the data.

The specific topic being investigated in this qualitative case study was the relationship between mismatch and effective language arts instruction in a kindergarten classroom. Mismatch was defined as the lack of congruence between the reading skills that children bring with them to the classroom and the classroom teacher's beliefs about those skills. The study, which lasted an entire school year, was located in a rural kindergarten classroom affected by poverty. The main data collection methods included classroom observations, interviews of key informants, including the kindergarten teacher, the principal, and the speech and language pathologist, and analysis of curriculum-related materials. A holistic approach to analysis was used for this investigation. A holistic approach was chosen because it attempts to understand phenomena as interwoven layers of complexity (Mabry, 1998). A holistic approach appreciates data within their contexts. Also, rather than being prescriptive in nature, a holistic approach to analysis encourages the emergence of results from the data. Throughout this investigation, every effort was made to allow this emergence to occur. The general results that emerged from the data were (a) an identification of the types and extent of mismatch in the classroom; (b) a description of the three main strategies the teacher used to deal with the mismatch; (c) an evaluation of the effectiveness of those strategies; and (d) a discussion of the extent to which match was ultimately

achieved in the classroom. The six safeguards described above were implemented throughout the research process to increase the validity of the results. A description of the specific application of these six safeguards is provided below.

Triangulation

In determining the design for the case study, special efforts were made to include triangulation. The first type of triangulation used was source triangulation. There were four main data sources for this case study project. The main data source was the kindergarten teacher, Ms. Ward. Seven interviews were conducted with Ms. Ward to obtain information about her expectations, knowledge, and practices regarding language arts. Two additional data sources were the principal of the school and the speech and language pathologist. The principal was interviewed to obtain information about the school, the students and their language difficulties, and the supports available for teachers. The speech and language pathologist was not formally interviewed, but several informal discussions were held to learn about the students' language and literacy challenges as well as Ms. Ward's concerns and needs. I was the final data source. I took detailed field notes each day that I observed in the classroom.

The second type of triangulation used in the case study was methodological triangulation. Three different data collection methods were used, including interview, observation, and document analysis. As mentioned above, eight formal interviews were conducted with two of the participants, as well as several informal interviews with the speech and language pathologist and Ms. Ward. In addition, 56 observations were made in the classroom and computer lab. The content of the observations included the general educational activities that occurred throughout the day. All “specials” (i.e., music, art, physical education, and computer lab) were excluded from observation with the exception of computer lab. Computer lab time was included in the

observations because Ms. Ward was the teacher for this period and therefore, had opportunities to instruct her students on various language arts skills. The final data collection method was document analysis. The documents analyzed included: the kindergarten curriculum (state curriculum guide, district curriculum guide, and a teacher's curriculum resource guide) and documents created by Ms. Ward for the purposes of the classroom (a weekly note that was sent home to parents regarding what had been covered that week in class, lesson plans for the Title I lessons that Ms. Ward had created, and monthly calendars that displayed weekly themes). Documents were read and analyzed, and outlines were developed regarding the content of the documents.

The final type of triangulation used was theoretical triangulation. Throughout the research process, attempts were made to use multiple theories and perspectives to interpret the data. One set of findings described how the teacher dealt with the mismatch that occurred in her classroom. Several theories were used to interpret these strategies and their effectiveness. For example, the teacher's instruction was analyzed using Bandura's (1986) observational learning and modeling, Vygotsky's (1934/1986) zone of proximal development, and Bruner's (Wood, Bruner, & Ross, 1976) scaffolding.

Throughout the research process and during the final analysis at the end, all three types of triangulation were used to increase validity. By having multiple sources and multiple data collection methods, I was looking for points of consistency and inconsistency. Consistency meant that the interpretation was robust, and therefore, it was a valuable part of understanding the case. Inconsistency meant that I needed to pursue the interpretation in more depth. Inconsistencies were not automatically dismissed as unimportant. Rather, they were further analyzed to determine their source, importance, and relevance. The theoretical triangulation

ensured that multiple theoretical viewpoints were used as the "lens" for interpretation.

Establishing a Data Trail

Due to time and financial constraints a full external audit was not conducted with this project. However, efforts were taken to establish a well-organized and thorough audit trail. All observation field notes, interviews, and document outlines were typed and organized in binders chronologically. Although this was a laborious task, it helped to make the data analysis phase more efficient. In addition, all narratives, interview transcripts, and outlines contained separate notes regarding my thoughts and insights about what I was seeing or hearing. That is, the data and the interpretations were contained in the same documents, yet they were clearly separated to indicate their differences. Furthermore, attempts were made to archive my thoughts about interpretations, themes, and insights by writing my ideas in a journal. There has been no request to view the written documentation that I created for this project, yet the option is available.

Acknowledging Researcher Subjectivity

In this research project, subjectivity was addressed in several ways. First, I provided an explanation of my professional background. My training is in both educational and developmental psychology, with a particular emphasis on language development in young children. This training provides me with a particular lens for looking at the world. Development and its impact on children is always a part of how I interpret the world. Second, I presented the conceptual framework with which I entered the study. The conceptual framework provided a clearer description of the specific areas of theory and research that have informed my understanding of language arts instruction. By acknowledging this conceptual framework, I am more aware of the influence that it has on my interpretations. In addition, sharing the conceptual framework with my readers helps them to better analyze my interpretations. Third, I

acknowledged, in writing, two assumptions that I had when I entered the project. First, I assumed that mismatch was going to be an issue for this classroom. This assumption was based upon multiple sources, including (a) the fact that Ms. Ward was a first year teacher and therefore limited in her experiences, (b) previous reading research conducted in the school that revealed a sizable percentage of children at-risk for reading problems, (c) the literature base on poverty and reading success, and (d) the literature base on teachers' expectations. The second assumption that I made prior to entering data collection was Ms. Ward's need to modify her expectations as well as change her curriculum and instruction in order to overcome the mismatch that existed between her expectations and the students' skills. Finally, I attempted to reduce the subjectivity in my study by using the words of my participants whenever possible rather than paraphrasing what was said. By using the participants' words, readers have the advantage of making their own interpretations of what the participants have said.

Member Checks and Participant Review

The first stage of Mabry's (1998) two-stage participant review involved two types of data, observation narratives and interview transcripts. It was determined it would have been entirely too time-consuming for Ms. Ward to read all 56 observation narratives. Therefore, a sample of the narratives was sent to Ms. Ward for her review. The sample included five narratives that were sent in two different mailings. Inclusion criteria for the sample of narratives were need for clarification or background information and representation of the entire school year. With each mailing, Ms. Ward was asked to review the narratives carefully, noting their accuracy. She was also encouraged to add comments for clarity. The mailings were sent after the completion of the data collection phase of the project because I did not want the narratives to influence Ms. Ward's teaching. I knew this decision could affect the accuracy of Ms. Ward's

review because of the time elapsed, but I determined that the risk of influencing Ms. Ward's teaching was greater. The narratives that were sent did not contain any initial interpretations or comments. The reviewed observation narratives were returned to me with relatively few comments. In general, the comments provided clarity regarding why certain instructional techniques or materials were used. Ms. Ward's comments were added to the narratives, highlighted by bold and italicized type, and were used during the data analysis phase.

All transcribed interviews, with the interpretations removed, were sent to interviewees for review. Interviewees were asked to review the transcripts for clarity and accuracy. Ms. Ward returned her transcripts with a few minor comments inserted, which I noted on the revised transcripts and incorporated into my interpretations. In general, the comments provided additional information about the topics discussed, as well as notations regarding inaccuracies in the transcription. For example, an error was found with regards to the name of a program. Ms. Moon, the principal, did not return her transcript; I assumed that she had no corrections or additions to make.

The second stage was not implemented in a formal manner. At the end of data collection, I shared with Ms. Ward some of my initial ideas about the strategies that I observed, in particular the strategy of scaffolding. She agreed it was a strategy she tried to incorporate into her classroom. Due to the passage of time between the data collection and the written report, I determined that a formal analysis of the final report would be a cumbersome task for Ms. Ward due to the size of the report, as well as the amount of time that had transpired since the data was collected. I acknowledge that this decision can be seen as a potential weakness of the study as it weakens the interpretative validity of the study.

Prolonged Engagement

During the school year this project took place, an extensive presence in the classroom was maintained. Fifty-six full day observations took place an average of two days a week for the entire school year. Observations began at 7:45 a.m. each day and lasted until 3:15 p.m., with the exception of a lunch break from 11:00 a.m. until 12:20 p.m. The extensive amount of time spent in Ms. Ward's classroom and in discussion of her ideas made me feel more confident about understanding her approach to language arts instruction. In addition, the extensive presence helped to reduce the level of participant bias. During the first few observations, it was noted that my presence in the classroom was affecting the routine. The students were constantly looking up at me or coming over to get attention. Also, initially Ms. Ward seemed nervous in my presence. On occasion, she directed her instruction to me instead of the students. By the third or fourth week, both Ms. Ward and the students began to ignore my presence. As the school year progressed, the observation narratives contained fewer and fewer instances of students approaching me or Ms. Ward directing her attention toward me during instruction.

Consideration of Disconfirming Evidence and Contradictory Interpretations

Throughout the research process, I attempted to consider and account for both discrepant and confirming cases. The first point at which disconfirming evidence became apparent was during my attempts at triangulation. When sources or methods resulted in inconsistencies, I sought to understand why. The second point where disconfirming evidence became evident was during the formal analysis process at the end. In coding the data and then organizing the codes into themes, I noted any discrepancies and entertained possible alternative explanations. For example, by February of the school year, Ms. Ward’s instructional strategies matched the language and literacy skills of her students for the most part. However, I began to notice that the

mismatch was not completely eliminated. In trying to understand this "discrepancy" in the data, I learned that Ms. Ward was being influenced by the expectations of those people around her. Some parents and the first grade teachers were pressuring her about her students' readiness for first grade. They felt she needed to be teaching more at a quicker rate.

Conclusions

Validity is a constant struggle in qualitative research. As researchers we have a responsibility to ensure that our results accurately reflect the phenomenon of interest. I have found the most success in increasing the validity of my work by incorporating various safeguards into my research design. In this paper, I have sought to explain several safeguards that I have found particularly useful and are supported in the literature on validity in qualitative research. Furthermore, in sharing examples of these safeguards in my own work, I hope that other researchers will share their experiences and unique safeguards so that I might learn additional ways to increase the validity of my current and future research.

References

Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice Hall.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton-Mifflin.

Denzin, N. K. (1989). *The research act: A theoretical introduction to sociological methods* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Eisner, E. W. (1991). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice*. New York: Macmillan.

Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 119-161). New York: Macmillan.

Geertz, C. (1973). *The interpretation of cultures: Selected essays*. New York: Basic Books.

Goetz, J. P., & LeCompte, M. D. (1984). *Ethnography and qualitative design in educational research*. New York: Academic Press.

Howe, K., & Eisenhart, M. (1990). Standards for qualitative (and quantitative) research: A prolegomenon. *Educational Researcher, 19,* 2-9.

Jansen, G., & Peshkin, A. (1992). Subjectivity in qualitative research. In M. D. LeCompte, W. L. Millroy, & J. Preissle (Eds.), *The handbook of qualitative research in education* (pp. 681-725). San Diego, CA: Academic Press.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.

Mabry, L. (1998). Case study methods. In A. J. Reynolds & H. J. Walberg (Eds.), *Advances in educational productivity* (Vol. 7, pp. 155-170). Greenwich, CT: JAI Press.

Maxwell, J. A. (1992). Understanding and validity in qualitative research. *Harvard Educational Review, 62*, 279-300.

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass Publishers.

Peshkin, A. (1988). In search of subjectivity – One's own. *Educational Researcher, 17*, 17-22.

Smith, M. L., & Glass, G. V. (1987). *Research and evaluation in education and the social sciences*. Englewood Cliffs, NJ: Prentice-Hall.

Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*, 89-100.

Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT Press. (Original work published in 1934).