

QUALITATIVE METHODS IN STATISTICS EDUCATION RESEARCH: METHODOLOGICAL PROBLEMS AND POSSIBLE SOLUTIONS

Sashi Sharma

Department of Mathematics, Science and Technology Education,
The University of Waikato, New Zealand
sashi@waikato.ac.nz

Despite being relatively new in statistics education research, qualitative approaches need special attention as attempts are being made to enhance the credibility and trustworthiness of this approach. It is important that researchers are aware of the limitations associated with these methods so that measures are put in place to try and minimize the effects of these limitations. Philosophical roots and features of this paradigm are outlined. Challenges faced by qualitative researchers in terms of reliability, validity and generability are considered. Uses of the interview approach in research literature as a data gathering tool are considered next. Advantages and disadvantages of the interview approach are outlined. An example of a research in statistics education is provided to illustrate methodological problems and solutions related to qualitative methods.

INTRODUCTION

In the last two decades, a new paradigm widely referred to as qualitative research paradigm has begun to dominate research in mathematics/statistics education. Merriam (1998) traces the philosophical roots of this paradigm to the interpretative school of thought which considers education to be a process and school a lived experience. Reality (knowledge) is constructed socially by individuals. There are multiple interpretations of single events and situations, hence knowledge is multi-layered and complex. This paradigm provides a general perspective on knowledge and research that allow researchers to select specific methods for particular projects (Ernest, 1997). There is often an emphasis on qualitative methods, although quantitative methods can also be used when required. For instance, Shaughnessy (2007) writes that there is a growing trend in statistics education for researchers to use both quantitative and qualitative approaches. In many studies (Amir & Williams, 1999; Fischbein & Schnarch, 1997; Shaughnessy, Watson, Moritz & Reading, 1999; Watson & Callingham, 2003) researchers have gathered and quantified results of surveys on statistics tasks in large numbers but have also conducted interviews with smaller number of students. Hypotheses generated about why students were answering survey questions in particular ways were validated in detailed interviews. Interviews often revealed lines of thought that were often missed in the survey data by researchers.

FEATURES OF QUALITATIVE RESEARCH

As with all research in education, there are different interpretations and definitions for qualitative research. It is alternatively called naturalistic inquiry, field study, case study, participant observation and ethnography (Merriam, 1998). According to Creswell (2008), researchers using qualitative approaches tend to attend to participants' views when doing their studies; ask general, open-ended questions; and collect data in natural settings as the study develops. For example, in a statistics classroom, data is collected as students interact in small groups, during whole-class discussions, interact with the teacher, or working individually. In qualitative approaches, researchers are regarded as data collection instruments and have an important role. Their values, assumptions, beliefs, and knowledge have an influence on the data that is collected. Qualitative approaches focus on tentative interpretations of what goes on in context, so it is important to acknowledge that the data will be glossed with the meanings and purposes of the participants.

Another feature of qualitative research is that the data can be analysed using an open ended approach (Cohen, Manion & Morrison, 2000). In such an open ended approach the data can be analysed with respect to the research questions and the theories posed by the researcher as well as theories of what is occurring that can come out of the data itself. Cohen et al. (2000) call this a *grounded theory* approach where the analysis of the data comes from looking at the data. Grounded theory data analysis can be done by a constant comparative approach. A constant comparison approach to data analysis looks to discover key and recurrent themes that may emerge from the

data. Themes can then be used to find, classify and compare incidents so a clear account can be given of the research.

CHALLENGES FACED BY QUALITATIVE RESEARCH METHODS

Despite the strengths attributed to qualitative research approaches, it is important that researchers be aware of the limitations associated with these methods so that measures are put in place to try and minimize the effects of these limitations. Anfara, Brown and Mangione. (2000) state that too often qualitative inquiry is evaluated against the positivist criteria such as validity and reliability and found to be lacking in some or all of these criteria. Gergen and Gergen (2000) write that there is a growing dissatisfaction among some qualitative researchers for moving away from scientific standards or being too positivist. They add that in response to this criticism many educational researchers do not provide adequate or clear justifications for their methods, findings and conclusions.

Reliability is a measure of consistency over time and over similar samples, the extent to which a procedure produces similar results under constant conditions on all occasions with the same person or across interviewers (Cohen et al., 2000). A question which produces one type of response on one occasion but a different response on another is considered unreliable. Cohen et al., (2000) point out that qualitative researchers tend to view reliability as a fit between what they record as data and what actually occurs in the natural setting rather than accurate measurements between different observations. Two researchers studying a single setting may come up with different data and produce different findings, but both studies can be considered reliable. According to Burns (1994) reliability in qualitative research can be enhanced if researchers outline the reasons for the research and the major question they want to address, they explicate their perspectives on the question, stating their research assumptions and biases, explain their data gathering procedures and categories developed for analysis. However, Anfara et al. (2000) explain that including information on the decisions that are made during the process of constructing qualitative research is part of responding to the question of whether not the findings are credible and trustworthy. They offer three suggestions for assessing the methodological rigor and analytical defensibility of the qualitative paradigm: designing interview questions that address research questions, conducting data analysis through code mapping and verifying findings and data triangulation.

Closely related to the concept of reliability is validity. Like reliability, it is concerned with errors that may occur in the research process. It is particularly concerned with whether an item measures or describes what it is supposed to measure or describe (Bell, 1993). McCormick and James (1988) write that qualitative researchers, because they try to understand the experiences of their informants in naturalistic settings, believe their methods bring them closer to reality and are, therefore, more valid or at least ecologically valid (that is valid for a specific group or place). Patton (2002) argues that validity in qualitative research methods is largely dependent upon the skill, competence and rigour of the researcher. For instance, threats to validity in interviewing can be minimised by decreasing the number of incorrect interpretations through asking interviewees whether inferences drawn from what has been said are correct. Hunting (1983) suggests steps to maximise content relevance and the representativeness of the selected tasks. He argues that content validity can be achieved by analysing proposed tasks in terms of the content stated to be appropriate for various age levels in prominent curriculum documents.

Researchers use the term generalizability to refer to whether the findings of a study hold up beyond the setting or individuals under study (Bell, 1993; Burns, 1994). Critics of qualitative research point out that a major disadvantage of in-depth qualitative study is that generalisation is not possible beyond the population under study. Many forms of qualitative research share the view that notions of validity and generalizability are established in different ways from traditional criteria used when referring to the transfer and application of findings from one setting to another. Qualitative research does not aim to generalize findings but to understand a specific situation. Attention is on the local setting and unique context, rather than on generalizing the findings. Bell (1993) argues that researchers concerned with the question of generalizability may draw upon other studies to establish the representativeness of what they have found, or they may conduct a larger number of less intense mini-studies. Most researchers highlight the importance of making the work of teaching and learning open to critique of the wider community (Drew et al., 2008; Watson,

2002). Making work available for public examination supports the improvement of the researcher's practice and increases the chance that the work becomes useful to other mathematics educators.

RESEARCH METHODS: THE INTERVIEW APPROACH

In qualitative research, interviewing is the major source of the qualitative data for understanding the phenomenon under study (Drew, Hardman & Hosp, 2008; Fontana & Frey, 2005). Interviews provide an opportunity for the researcher to investigate ideas and beliefs of participants further and to gather data which may not have been obtained by other methods such as observation or survey (Cohen et al., 2000; Shaughnessy, 2007).

The interview method takes the form of a dialogue in which the researcher seeks to elicit information from the subject about how the latter thinks. Usually, the discussion is centred upon a task or problem which has been carefully chosen to give the respondent every opportunity to display behaviour from which mental mechanisms used in thinking about that task can be inferred.

Depending on the purpose of the interview, interview questions can be highly-structured, semi-structured, or open-ended. Additionally, the researcher's familiarity with and knowledge about the exploratory or confirmatory nature of the study determines the structure of the interview questions. Important aspects in interviews include maintaining a relaxed manner, asking clear questions, note-taking, appropriate use of follow-up question or probes, establishing trust, and keeping track of responses (Cohen, Manion & Morrison, 2000; Drew et al, 2008).

It has been pointed out that the direct interaction of the interview is the source of both its advantages and disadvantages as a research technique (Guba & Lincoln, 1994). One of the advantages of using the interview is that it is flexible (Drew et al., 2008). So interviews have the potential to provide greater depth of information. In an interview, personal perspectives of the respondent is provided, meanings and feelings can be quite detailed. Clarification of questions is possible in an interview, and the researcher has the opportunity to probe what is being said by the participant by asking for clarifications and/or examples. Through interviews the researcher is able to discover experiences that may have taken place in students' lives which might have a bearing on their learning of mathematics now. The researcher can use this knowledge and information to test hypotheses or to suggest new ones (Guba & Lincoln, 1994; Cohen, Manion & Morrison, 2007).

Several variations of the same task may be presented to probe the strengths and limits of a students understanding and to provide additional insights into the child's cognitive structures. The response that a student gives on each variation determines what variation is introduced next (Hunting, 1983). It follows that this technique allows students to describe, predict and explain events at their own level of cognitive development.

Although the interview has a number of important advantages over other data collection techniques, it does have definite limitations as a research tool (Hunting, 1983). McCormick and James (1988) note that both interviewers and respondents are sources of bias, the former through the questions they ask, or their perceived role and presence, the latter through their conception of the interview, their memory of the event, their ability to answer and their motivation in taking part in the interview. They suggest that whatever procedure for collecting data is selected, it should always be examined critically to assess to what extent it is likely to be reliable and valid.

Another disadvantage is that interviews can be time-consuming in terms of travel time and time required for transcribing and interpreting information. Interviews require great skill and expertise of the interviewer ((Drew et al., 2008). An interview is susceptible to manipulation by the respondent. This results in collecting false or distorted information, which, in turn leads to false findings and conclusions. There are some specific issues with regards to interviews in the school context. Trust and respect are key issues in the research, especially with Pasifika research participants (Anae, Coxon, Mara, Wendt-Samu & Finau, 2001). Language is also an issue as participants in the research may have English as a second language.

AN EXAMPLE OF A QUALITATIVE RESEARCH IN STATISTICS EDUCATION

A concern about students' difficulties with statistics and probability and a lack of research in this area outside of western countries led to my study (Sharma, 1997). Information was gathered about how form five (14 to 16 year-olds) Fijian-Indian students construct ideas about statistics. Overall, the research addressed the question:

"What ideas do Form Five Fijian students have about statistics and probability, and how do they construct these?"

The study comprised a series of individual interviews in a typical secondary school with a sample of differing abilities and attitudes. Each student was interviewed individually by myself in a room away from the rest of the class. Each interview lasted between 40 to 50 minutes. All interviews were recorded on audio tape and transcribed for analysis. Data also comprised of classroom observations and student work.

The data revealed that many of the students held beliefs and used strategies based on prior experiences which partially inhibited the students' development of statistical ideas—these included superstitions and beliefs about both animate and inanimate objects. Some students based their reasoning on their non-mathematical in-school experiences, and on their cultural experiences which included religious and everyday experiences. Further difficulties arose when students interpreted a task differently from that intended and when some task words carried a mathematical meaning different from the usual everyday meaning. The interview approach allowed students to demonstrate statistical understanding and questioning which would not have been possible in say a multiple-choice format. The results provide evidence that students often gave correct answers for incorrect reasons. For example, when two dice were rolled some students believed that the chance of getting different numbers was more than the chance of obtaining same numbers. It would be easy to conclude from the results that the students had a well developed concept of compound events. However, the justifications provided by the students indicate that they had no rational explanations for their responses. None of the explanations indicated any representation of the corresponding sample space. The explanations seemed to indicate a view that chance is naturally unpredictable.

DEALING WITH ISSUES OF QUALITY CRITERIA

McCormick and James (1988) suggest that whatever procedure for collecting data is selected, it should always be examined critically to assess to what extent it is likely to be valid and reliable. This section outlines how these issues were addressed in my study.

To aid validity the study made use of some of the view points mentioned by Anfara et al. (2000). Research questions provided scaffolding for my study. Since in-depth interviews were used as primary data-gathering source, the interview questions were cross-referenced to the study's research questions. The tasks used in the study were analysed in terms of the content stated to be appropriate for various age levels in the curriculum document (Fijian Ministry of Education, 1994) and the construct being investigated. The appropriateness of these interview tasks for Fijian students was established by checking them with the class teacher and the HOD mathematics at the school.

Most of the aspects outlined by Burns (1994) and McCormick and James (1988) were incorporated in the present research. A rationale for the study (importance of statistics in society, difficulties in statistical reasoning and a lack of research in this area) and the major questions addressed were outlined in chapter one. Researcher perspectives on the project (background, expectations), together with the categories developed for analysis are outlined in the thesis. Additionally, to aid reliability and minimise interviewer bias, the study employed a semi-structured approach. Each student interviewed was initially confronted with an identical task. During the interview, care was taken to avoid leading the students towards any particular viewpoint, so responses to questions were accepted as they were given and probing questions were asked simply to ascertain the reasons for what the student thought. My prior experiences and familiarity with the interview approach enabled me to use supplementary questions to check whether the interpretations constructed by the me fitted with the student construction. In some cases, students were asked to comment on the transcripts to ensure that the meaning constructed by me was the same as that constructed by the students.

Researchers use the term generalizability to refer to whether the findings of a study hold up beyond the setting or individuals under study. I acknowledge that the findings of my study may not be generalizable to all settings because students in other areas and countries are likely to have very different experiences and hence their reasoning would differ. However, since I was alert to the possible threats to generalizability, the results from other studies have been used to establish the representativeness of the findings. The purpose of my research was to build on what had already

been done by others. Amir and Williams (1999), Jones, Langrall, and Mooney (2007) and Shaughnessy (2007, 1992) raised the concerns about the lack of probability research outside of western countries. Shaughnessy advocated small scale studies that incorporate the strong ethnic influence that culture has on students' probabilistic reasoning. My study addressed some of these shortcomings. Additionally, since one of the purposes of the study was to identify general patterns of reasoning, findings gained from the study are likely to be of some importance to other groups. One of the reasons for developing a statistical model on students statistical ideas in my study was to provide empirically grounded results that others can adjust to their local circumstances. It must be noted that any model remains hypothetical because each situation, each teacher and each class is different. However, patterns could be found in students' thinking that may be similar across different contexts. The findings will contribute to the refinement of conceptual models developed in earlier research and assist teachers by providing a rubric for teaching and assessing students.

The information gathered was initially written up as a doctoral thesis and then turned into articles to be sent to national and international journals for publication. It is likely that workshops, conference presentations, and articles that focus on issues relating to research methodology over the course of the research project may also eventuate. Since the work will be accessible, the wider community, scholars and policy makers will all benefit from this project. It could provide a catalyst for change-others could try it out in their own settings.

The data reported in my study was mostly obtained from the individual interviews. Data from classroom observations and student written work were used to check initial findings, to fill gaps, and to informally triangulate the findings. In some cases students moved from appropriate strategies to inappropriate ones when responding to the interview tasks. One of the factors that could have made students change in this way was the students' experiences as learners at school. Student reasoning is not questioned in class. The teacher's concern is whether they give correct answers. If the students give correct answers it is accepted by the teacher. Students are only questioned when they give wrong answers. It seems probable that in my research, the students interpreted my probing as an indication that something was wrong with their answers and so they quickly switched to a different strategy. Obtaining data from interviews and classroom observations helped me to more fully understand the particular incidents such as switching strategy.

CONCLUSION

One of the strengths of qualitative research is that participants' behaviour is recorded in natural settings, with all the intricacies of the environment operating. Qualitative research is also particularly useful for in-depth study of a small group of people. Despite the strengths attributed to qualitative research approaches, it has been criticised for lack of reliability, validity and generalizability. In qualitative research it may be difficult to understand what is going on due to the complexity of the natural settings. It is important that researchers be aware of the limitations associated with these methods so that measures are put in place to try and minimize the effects of these limitations. The paper offers some suggestions for enhancing the reliability, validity and generalizability of qualitative research methods.

REFERENCES

- Amir, G., & Williams, J. (1999). Cultural Influences on Children's Probabilistic Thinking. *Journal of Mathematical Behavior*, 18(10), 85-107.
- Anae, M., Coxon, E., Mara, D., Wendt-Samu, T., & Finau, C. (2001). *Pasifika Education Research Guidelines Final Report*: Auckland: Uniservices.
- Anfara, V. A., Brown, M., & Mangione T. L. (2002). Qualitative Analysis on Stage: Making the Research Process More Public. *Educational Researcher*, 31(7), 28-38.
- Bell, J. (1993). *Doing Your Research Project*. Buckingham, Philadelphia: Open.
- Burns, R. (1994). *Introduction to research methods*. Melbourne: Longman.
- Borko, H., Liston, D., & Whitcomb, J. A. (2007). Genres of Empirical Research in Teacher Education. *Journal of Teacher Education*, 58(1), 3-11.
- Cohen, L., Manion, L., & Morrison K. (2000). *Research Methods in Education* (5th Edition). London: Routledge Falmer.

- Cresswell, J. W. (2008). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Denzin, N. K., & Lincoln, Y. S. (1994). *Handbook of qualitative research*. London: Sage.
- Drew, C. J., Hardman, M. L., & Hosp, J. L. (2008). *Designing and Conducting Research in Education*. LA: Sage Publications.
- Ernest, P. (1997). The Epistemological Basis of Qualitative Research in Mathematics Education: A Postmodern Perspective *Journal for Research in Mathematics Education*. Monograph, Vol. 9, 22-177.
- Fijian Ministry of Education, Women, Culture, Science and Technology (1994). *Primary Mathematics Prescriptions*. Suva: Curriculum Development Unit.
- Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education*, 28, 96-105.
- Flick, U. (2007). *Designing qualitative research*. London: Sage.
- Fontana, A., & Frey, J. H. (2005). The interview: From neutral stance to political involvement. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage Handbook of Qualitative Research* (pp.695-725). London: Sage Publications.
- Gergen, M. M., & Gergen, K. J. (2000). Qualitative inquiry: Tensions and transformations. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 1025-1046). Thousand Oaks, CA: Sage Publications.
- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.) *Handbook of Qualitative Research* (pp. 105-117). London: Sage Publications.
- Hunting, R. (1997). Clinical interview methods in mathematics education research and practice. *Journal of Mathematical Behavior*, 16(2), 145-165.
- Hunting, R. P. (1983). Emerging methodologies for understanding internal processes governing children's mathematical behaviour. *The Australian Journal of Education*, 27(1), 45-61.
- Jones, G. A., Langrall, C. W., & Mooney, E. S. (2007). Research in probability: Responding to classroom realities. In F. K. Lester Jr (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 909-955). Reston: The National Council of Teachers of Mathematics.
- McCormick, R. and James, J. (1988). *Curriculum Evaluation in Schools*. London: Croom Helm.
- Merriam, S.. (1998). *Qualitative research and case study applications in education*. San Francisco: Josey-Bass.
- Patton, M. Q. (2002). *Qualitative evaluation and Research Methods* (3rd edition). London: Sage.
- Sharma, S. (1997). *Statistical ideas of high school students: Some findings from Fiji*. Unpublished doctoral thesis. Waikato University, Hamilton, New Zealand
- Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester Jr (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 957-1009). Reston: The National Council of Teachers of Mathematics.
- Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465-494). New York: Macmillan.
- Shaughnessy, J. M., Watson, J. Moritz, J., & Reading, C. (1999). *School mathematics students' acknowledgement of statistical variation*. Paper presented at the 77th Annual National Council of Teachers of Mathematics Conference, San Francisco, CA.
- Watson, J. M. (2002). Doing research in statistics education: More than just data. *The Sixth International Conference in Teaching Statistics* (ICOTS 6). Cape Town: South Africa.
- Watson, J. M., & Callingham R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, 2(2), 3-46.
- Battista, M, Boerst, T. Confrey, J, Knuth, E. Smith, M., Sutton, J. White, D., & Quander, J. (2009). Research in Mathematics Education: Multiple Methods for Multiple users. *Journal of Research in mathematics Education*, 40(3), 216-240.
- Hill, H., & Shih, J. (2009). Examining the quality of statistical mathematics Education Research. *Journal of Research in mathematics Education*, 40(3), 241-250.