

BEING CRITICAL ABOUT APPROACHES TO RESEARCH IN STATISTICS EDUCATION

Anna Reid and Peter Petocz
Macquarie University, Australia
anna.reid@mq.edu.au

Teachers undertaking educational research for the first time usually begin their explorations by evaluating some aspect of their practice. By contrast, experienced researchers will start from an argued research question supported by a defined theoretical framework. In this paper, we use a critical discourse approach to explore various interpretive research paradigms that are commonly used to investigate aspects of statistics education. By considering the underlying epistemological positions and critiquing the approaches and methods used to explore human action in social situations, we become more critical in the design, implementation and reporting of research in statistics education.

INTRODUCTION

Statistics education research is represented through language that is situated in the world of statistics teachers and lent an external credibility through the use of academic discourse (such as citations for authority, institutional affiliation and academic levels). How can we make sense of the different sorts of research that comprise statistics education research? What do these research orientations say about the focus and nature of research in statistics pedagogy? What constitutes legitimacy in research practice of statistics education? A critical discourse approach enables us to treat language as a form of social practice. Texts – in this case, texts representing a sample of research in statistics education – are situated within the activity of the writers and readers: thus, the research writings can be interpreted as representative of those situations. In this paper, we will examine a sample of texts from ICOTS-6 (and elsewhere), including our own work using phenomenography, that represent the discourse of statistical education researchers. We will look at the locus and focus of the studies, and the relationships between the central research questions, the methods used to explore those questions and their outcomes. We aim to establish a framework for reflective teachers and active researchers to examine their own research assumptions and evaluate the ‘trustworthiness’ of their findings.

Much mainstream research in higher education uses qualitative methods to examine learning situations, often represented through analysis of some sort of linguistic elements. For statisticians, who deal primarily with variation using numerical methods, qualitative approaches can seem arcane: within statistics, rigour is represented by the use of statistical techniques. A recent talk by Shelley (2005) claims that rigour in educational research in statistics (and elsewhere) can only be established by the use of “proper scientific methods,” and describes US Government legislation to “foster scientifically valid research.” Here, Shelley legitimises quantitative experimental designs for the study of social situations through a careful elaboration of language common to the core of statistics as a research method in its own right. The use of terms such as ‘scientific,’ ‘experiment,’ ‘gold standard,’ ‘quantification,’ ‘random trials,’ ‘valid control groups’ and even ‘power’ all seem consistent with the world inhabited by statisticians.

However, those who are involved with pedagogical experiments find themselves questioning the *lived* experience of the participants and the implications for change in learning *relationships*. Researching pedagogical situations provides an opportunity to explore using ‘improper’ methods, and to acknowledge the *qualitative* variation found amongst people, their orientations to learning and their expectations for their futures. Indeed, statisticians have become aware of the benefits of including qualitative as well as quantitative aspects of research in statistics education (e.g., Batanero *et al.*, 2001). Moss (2004) suggests that a critical discourse approach can bridge barriers created by the inter-subjective nature of statistics and the social realities represented through hermeneutic approaches. Guba and Lincoln (1989) propose that the notion of trustworthiness as a qualitative equivalent of rigour demands that four criteria are met: the research must be credible, transferable, dependable and confirmable. These criteria lend authenticity to the research.

FROM TEACHERS TO RESEARCHERS

Our own ICOTS community provides us with a rich source of information regarding the place of pedagogical research within our own discipline. As opposed to the claims of Shelley (2005), where educational researchers are exhorted to involve statisticians in their practice, at ICOTS, statisticians *are* the educational researchers. Thus, it is reasonable to look at the range of topics and approaches that have previously been reported through our conference proceedings. We can subject such statistics education papers to a critical discourse analysis (Polkinghorne, 1995). This particular form of analysis uses artefacts generated from within a particular social milieu (in this case, the statistics education community) and examines them for hints about their 'location' within the milieu and the 'spaces' between them.

Kreber (2002) presents a continuum of types of engagement with research in higher education. She acknowledges that teaching "*is seen as a very time-consuming but also scholarly activity in that it requires sound knowledge of one's discipline as well as a good understanding of how to help students grow within, and perhaps even beyond, the discipline*" (p.9). At one extreme, papers characteristic of an observational and experiential approach contain descriptions of statistical problems in the classroom and provide practical teaching advice on their solution (e.g., Kunte, 2002; Connor, 2002). In one sense, the underlying epistemology here is one of observation, and assumes that the problems described are common and that other practitioners will see the utility of the solutions. Within pedagogical theory, this sort of approach could be considered 'reflection in or on action' (Schön, 1983). Moving along the continuum, we find that Kreber's description of 'scholarship in teaching' (after Boyer, 1990) represents another group of papers. These papers hold their practice up to reflective scrutiny, acknowledging that change and improvement in the quality of learning and teaching is always possible (e.g., Chauchat, 2002; Rossman and Chance, 2002). In one sense, this approach is more systematic than the experience and observation group. Looking through the abstracts and papers presented at ICOTS-6, we can see that only a few move beyond these two frames (the observational and the reflective) towards a designed research effort (e.g., Lavigne and Lajoie, 2002; Hoyles and Noss, 2002).

Let us look at some of these papers in more detail. Kunte (2002) demonstrates an awareness of the embedded and implicit context of teaching statistics and presents a solution to a particular problem using the language of a statistician within the dominant scientific paradigm of the profession. However, elements of language provide us with surety that the question is derived from a pedagogical issue – in this case the inadequacy of statistical texts to properly explain the notion of correlation. The solution is presented mathematically, with an implication that this will assist students (and perhaps teachers!) in their developing understanding of statistical concepts. Connor (2002) reports on the implementation of the *CensusAtSchool* project, giving details of the development and implementation of a solution to the pedagogical problem of providing students with relevant and interesting data that can be examined in a modern technological setting. These papers are representative of Kreber's observation and experience category.

A more explicit form of reflective practice is shown by Chauchat (2002). Twenty years' experience with postgraduate students has enabled him to develop and refine authentic learning in which students are required to find real situations where statistics is used and to solve associated problems. Here, students see statistics in action beyond a classroom experience. Indeed, every six years, they work on the statistical aspects of a national election, providing them with the real pressure of work-related deadlines. There is no 'assessment task,' but Chauchat suggests that their learning is 'just-in-time' and profound. Rossman and Chance (2002) describe a project to develop curriculum for a first course in statistics for students with a strong mathematical background. The discussion is set in the context of a broad reflection on developments in statistics education and national priorities for mathematics education. Each of these papers goes beyond observation and experience to represent Kreber's scholarly reflection category.

Lavigne and Lajoie (2002) describe a focused pedagogical research design with which they investigate the effect of social dimension as a factor in the success of group-work projects with school children. They used their experience in the classroom with support from a wide literature to formulate the question for investigation. Observational techniques and analysis of actual student dialogue provide them with evidence regarding the role of human interaction and knowledge for group-based statistical decision making. It is interesting, as the main results (which

are qualitative) are supported by the use of tables which enable it to ‘fit’ with a statistical audience. Thus, Lavigne and Lajoie acknowledge their own situation as joint researchers/teachers /statisticians and use language that inhabits all three fields. In the broader context of investigation into the role of (explicit and hidden) mathematical activity in professional life, Hoyles and Noss (2002) examine ways in which paediatric nurses view average and variation. They combine a cognitive approach with a recognition of the sociocultural dimension of statistical understanding, utilising a combination of ethnographic observation, clinical interviews and a ‘teaching experiment.’ These papers are examples of Kreber’s third category of designed research.

Overall, these papers provide examples of the range of ways in which our community approaches their pedagogic research processes. Figure 1 (below) can help illuminate some of the inherent disjunctions through an explanation of epistemological positions: it demonstrates the relationships between some common research methods and the dominant research paradigm in higher education. Of course, the list is not exhaustive, but merely indicative of the range of different research approaches that can be taken. Awareness of the underlying epistemology is important as this provides an indication of how the research outcomes are represented to the community, and in turn how those outcomes are understood and appreciated. Similarly, awareness of dominant disciplinary values enables any of these research approaches to be situated meaningfully within common discourse. Confusion arises (as in Shelley’s assertions) by insisting that rigour and validity can be established using the same criteria for all research orientations.

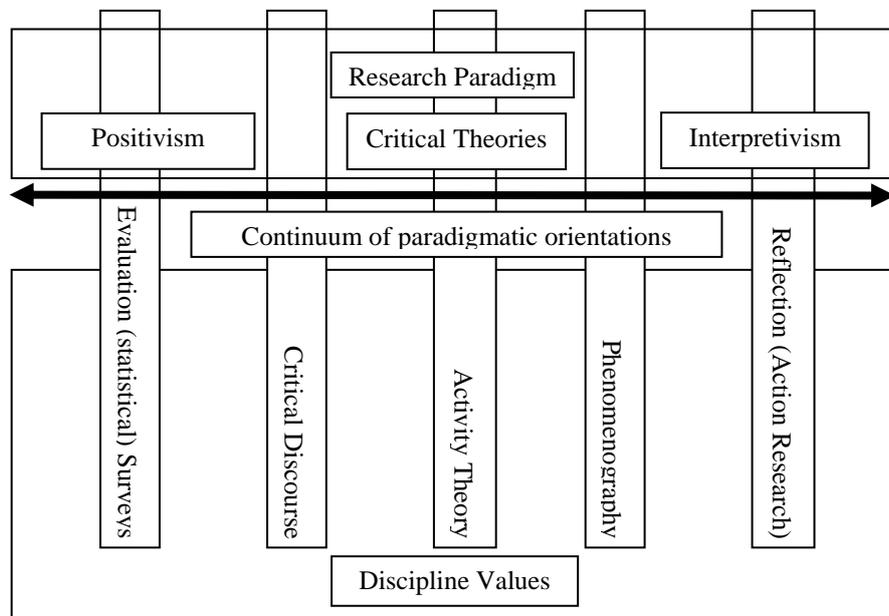


Figure 1: Relationships between common research methodologies in higher education

PHENOMENOGRAPHY – A QUALITATIVE RESEARCH FRAMEWORK

We have been using a theoretical framework known as phenomenography for research into pedagogy in statistics (and other disciplines). Phenomenography is a qualitative orientation to research that is often used to describe the experience of learning and/or teaching (Marton and Booth, 1997). It focuses on aspects that are critically *different* within a group involved in the same situation: these differences make one way of seeing the situation qualitatively different from another. In this context, it is interesting to observe that phenomenography and statistics each focus on the study of variation, though from quite different viewpoints! Data are typically collected through in-depth, open-ended interviews that allow each person to fully describe their experience. In the analytic phase, categories that describe the essential variation between one way of experiencing something and another are formulated: this is known as the *outcome space*. The next stage is a return to the transcripts to check that the categories are supported by the raw data. If there is no direct evidence from the transcripts (in the form of extended quotations) then the category is abandoned. Software such as *NVivo* (QSR International, 2002) may be useful for this

phase. The categories are usually arranged in a hierarchy from narrower, fragmented views to broader, holistic ones. People who express the broader conceptions are also aware of the narrower conceptions, and can make use of them if necessary: however, people who express the narrower conceptions seem unaware that other conceptions exist. It is for this reason that we as educators favour the broader, more expansive conceptions over the narrower, more limiting ones.

Phenomenographers assume that people's experiences vary, and that the description of the essential variation enables targeted change in teaching practice to take place. Thus, the initial question is usually derived from a pedagogical problem, explored using individuals' experience, analysed using a strict protocol, checked through further interrogation of the data, and then presented in a form that could be used for curriculum and teaching change. In this case, the emphasis lies on internal validity determined by the participant group. As an example, Reid and Petocz (2002) reported on the results of an empirical study of students' conceptions, or understanding, of statistics. The data consisted of transcripts of interviews with 20 students who were majoring in some area of the mathematical sciences (including statistics). Six qualitatively different conceptions of statistics were described, in three broad groups that can be described as a focus on techniques, a focus on data, and a focus on meaning. The pedagogical implications of the phenomenographical results were discussed: for instance, the importance of recognising that such variation exists in statistics classes, and the importance of designing a learning environment that encourages students' awareness of the broader role of statistics in their lives.

Once again, we return to the challenge provided by Shelley (2005). If we are going to disagree with the premise that quality research in statistics education must be experimental in the scientific sense then we need to justify a qualitative approach such as phenomenography in terms of its reliability and internal validity (Sandberg, 1996). For instance, there is a presumption with Shelley's model that rigour is established through the definition of normality and the measurement of the amount change through a direct intervention group and a control group. A recent research study in Thailand tried just such an approach in an investigation of online versus traditional pedagogy in a business statistics class (Suanpang *et al.*, 2004). However Suanpang determined that this approach alone was not enough to answer the real research question – what does it take for students to learn? She established an alternate approach through the use of hermeneutic interviews that revealed a broader and richer view of the same experience.

Of course, there is an ongoing debate about the distinction between qualitative and quantitative approaches to research and their combinations: Fielding and Schreier (2001) and the papers in the same volume of the journal explore the issues in depth, and Johnson and Onwuegbuzie (2004) present a justification for 'mixed methods' from pragmatic philosophy. Taking account of the essential features of the debate, we find that we need to ask these questions of any research project in statistics education in order to establish its trustworthiness:

- Why would you (or others) trust the findings of this particular research project?
- Can the results of this project, involving people in one situation, be transferred to another?
- How does this particular research orientation address the issues of reliability and validity?

ALTERNATIVES, EXTENSIONS AND COMBINATIONS

The debate about 'mixed methods' usually refers to combinations of quantitative and qualitative research approaches. However, it is equally interesting to look at hybrids of distinct qualitative approaches: here, we will illustrate some approaches that we have used in our own investigations. Firstly, it may be possible to use an alternative method to analyse data collected for a phenomenographic study. In a project investigating statistics' educators ideas about teaching service statistics, we subjected transcripts to both thematic and phenomenographic analysis (Gordon *et al.*, 2005). Thematic analysis has a different epistemological basis – one where common ideas rather than differences are sought. Rather than establishing underlying *variation in meaning* from the transcripts (which is the focus of phenomenography), thematic analysis observes common phrases and emphases made by the participants. These common areas are then put into groupings and the groups given a descriptor. The outcome has no particular sense of ordering or complexity, but shows the range of common experiences. The validity of the approach is established through the identification of short quotations.

Secondly, it is sometimes useful or necessary to utilise extensions of a basic approach. The investigation of students' conceptions of statistics (Reid and Petocz, 2002) that we have described earlier represents a classical phenomenographic approach. In contrast, our investigation of the relationship between students conceptions of learning statistics and their expectations of their teachers (Petocz and Reid, 2003) required us to extend the methodology. We needed to assign individual students to each of the phenomenographic categories developed to describe the range of conceptions of learning statistics and of the role of the statistics teacher in order to discover an interesting fact. Students who showed the broadest views of learning statistics (as using statistics to understand other areas beyond the discipline, or to change ones view of the world) also showed the full range of conceptions of the teacher's role (from providing essential materials, structure and motivation to being a catalyst for change and growth), while students who showed the narrowest views of learning statistics (as doing the required activities) only showed the narrowest conceptions of the teacher's role. Assigning respondents to phenomenographic categories is an extension of the basic methodology, and is subject to caveats concerning the specific time and situation of the experience: nevertheless, it was a necessary step in order to reach our conclusions.

Thirdly, we have made use of various combinations or hybrids of phenomenography with other qualitative methodologies. For example, we have investigated and reported the whole cycle of development of phenomenographic categories describing students' conceptions of statistics, and incorporating this knowledge in practical statistics pedagogy using an action research approach (Reid and Petocz, 2003). Action research (McNiff, 2002) is a form of enquiry that focuses on a critical approach to practice, undertaken by practitioners from within rather than by researchers outside the practice: thus, its epistemology is based on the validity of experience.

CONCLUSION: THE CRITICAL APPROACH

We have been involved for several years researching the experience of learners and teachers in statistics. We started using the emergent hermeneutic approach known as phenomenography, which focuses on identifying variation in human experience: This approach could be considered as being at one extreme of the epistemological continuum. However, the use of a single methodology often results in outcomes that can only be appreciated from within that framework. For pedagogical research to have an impact, research outcomes need to be accessible to a wide range of potential users. Adopting complementary methods can help us develop alternate interpretations of social data, while at the same time challenging us to explain and justify the relationships between the methodologies.

In particular, as statisticians as well as statistics education researchers, our own discipline values can seduce us into an uncritical acceptance of the superiority of quantitative approaches even in situations that are best investigated using hermeneutic methodologies. On the other hand, our use of qualitative research approaches can feed back into our notions of our own discipline, highlighting those aspects of statistical studies, for instance the selection and wording of questions for a survey or the decisions of what variables to measure, that are subjective and not amenable to quantitative approaches. Ultimately, research in statistics education must be usable and useful, and the research questions and answers are more important than the methodologies.

REFERENCES

- Ashworth, P. and Lucas, U. (2000). Achieving empathy and engagement: A practical approach to the design, conduct and reporting of phenomenographic research. *Studies in Higher Education*, 25(3), 295–308.
- Batanero, C., Garfield, J. and Ottaviani, M. (2001). Building a research agenda for statistics education. *Statistical Education Research Newsletter*, 2(2), 9–13.
- Boyer, E. (1990). *Scholarship Reconsidered: Priorities of the Professoriate*. Princeton, NJ: The Carnegie Foundation for the Advancement of Learning.
- Chauchat, J. (2002) Survey sampling: Learning by doing. A twenty years graduate teaching experience. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics*, Cape Town. Voorburg, The Netherlands: International Statistical Institute.

- Connor, D. (2002). CensusAtSchool 2002: Creation to collation to classroom. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics*, Cape Town. Voorburg, The Netherlands: International Statistical Institute.
- Fielding, N. and Schreier, M. (2001). Introduction: On the compatibility between qualitative and quantitative research methods. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 2(1), <http://www.qualitative-research.net/fqs-texte/1-01/1-01hrsg-e.htm>.
- Gordon, S., Reid, A. and Petocz, P. (2005). How important are communication skills for 'good' statistics students? – An international perspective. In L. Weldon and B. Phillips (Eds.), *Proceedings of the ISI/IASE Satellite on Statistics Education and the Communication of Statistics*, <http://www.stat.auckland.ac.nz/~iase/publications/14/gordon.pdf>.
- Gordon, S., Reid, A. and Petocz, P. (submitted to *Statistics Education Research Journal*). The experience of teaching service statistics.
- Guba, E. and Lincoln, Y. (1989). *Fourth Generation Evaluation*. Newbury Park, CA: Sage.
- Hoyles, C. and Noss, R. (2002). Problematising statistical meanings: A sociocultural perspective. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics*, Cape Town. Voorburg, The Netherlands: International Statistical Institute.
- Johnson, R. B. and Onwuegbuzie, A. J. (2004). Mixed methods research: A paradigm whose time has come. *Educational Researcher*, 33(7), 14–26.
- Kreber, C. (2002). Teaching excellence, teaching expertise, and the scholarship of teaching. *Innovative Higher Education*, 27(1), 5–23.
- Kunte, S. (2002) Understanding the sign of a correlation coefficient: An example. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics*, Cape Town. Voorburg, The Netherlands: International Statistical Institute.
- Lavigne, N. and Joloie, S. (2002). Factors affecting student decision making on statistics projects. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics*, Cape Town. Voorburg, The Netherlands: International Statistical Institute.
- Marton, F. and Booth, S. (1997). *Learning and Awareness*. Mahwah, NJ: Lawrence Erlbaum.
- McNiff, J. with Whitehead, J. (2002). *Action Research: Principles and Practice*. London: Routledge Falmer.
- Moss, G. (2004). Provisions of trustworthiness in critical discourse narrative research: Bridging intersubjectivity and fidelity. *The Qualitative Report*, 9(2), 359 – 374.
- Petocz, P. and Reid, A. (2003). Relationships between students' experience of learning statistics and teaching statistics. *Statistics Education Research Journal*, 2(1), 39–53.
- Polkinghorne, D. (1995). Narrative configuration in qualitative analysis. In J. Hatch and R. Wisniewski (Eds.), *Life History and Narrative*, (pp. 5-23). London: Falmer Press.
- QSR International. (2002). *NVivo 2.0*. Doncaster, Victoria, Australia, www.qsrinternational.com.
- Reid, A. and Petocz, P. (2002). Students' conceptions of statistics: A phenomenographic study. *Journal of Statistics Education*, 10(2), <http://www.amstat.org/publications/jse/v10n2/reid.html>.
- Reid, A. and Petocz, P. (2003). Completing the circle: Researchers of practice in statistics education. *Mathematics Education Research Journal*, 15(3), 288–300.
- Rossman, A. and Chance, B. (2002). A data-oriented, active learning, post-calculus introduction to statistical concepts, methods, and theory. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics*, Cape Town. Voorburg, The Netherlands: International Statistical Institute.
- Sandberg, J. (1996). Are phenomenographic results reliable? In G. Dall'Alba and B. Hasselgren (Eds.), *Reflections on Phenomenography*, 109, (pp. 129-140). Goteborg Studies in Educational Sciences.
- Schön, D. (1987). *Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Professions*. San Francisco: Jossey Bass.
- Shelley, M. (2005). Education research meets the gold standard: statistics, education, and research methods after “No Child Left Behind.” Presented at the Joint Statistical Meetings, Minneapolis, August.
- Suanpang, P., Petocz, P. and Kalceff, W. (2004). Student attitudes to learning business statistics: comparison of online and traditional methods. *Educational Technology and Society*, 7(3), 9–20.