New Approaches to Gathering Data on Student Learning for Research in Statistics Education

Beth L. Chance
Cal Poly, Department of Statistics
101 Grand Avenue
San Luis Obispo, CA, USA
bchance@calpoly.edu

Joan B. Garfield
University of Minnesota, Department of Educational Psychology
Burton Hall, 178 Pillsbury Drive S.E.
Minneapolis, MN, USA
jbg@tc.umn.edu

1. Introduction

The past decade has witnessed a pedagogical movement in statistics education aimed at shifting the focus of instruction away from theory and recipes toward statistics thinking, genuine data, conceptual understanding, and active learning. Much of this movement has been motivated by research in educational psychology, psychology, mathematics education and science education (see Garfield, 1995). However, there is still need for documented evidence of whether such changes enhance student learning, retention, and appreciation of statistics. This evidence is also needed to explore the most effective instructional techniques and to develop models of how students shape their statistical understanding. In order to advance the field of statistics education and achieve academic recognition, it is essential to have a well-developed research literature and research agenda (Batanero, et al 2000). Statistics education research will not be able to sufficiently impact policy or the practice of teaching until individual studies become grounded in a broader program of study.

2. Current Trends in Statistics Education Research

Statistics education is unique in that instructors typically come from a variety of backgrounds and fields of application. Similarly, statistics education research has been conducted by a variety of individuals who represent different disciplines, educational programs and training in research methods. For example, research in psychology has revealed ways people reason about statistical or probabilistic information, paying particular attention to faulty reasoning and misconceptions. Studies conducted by mathematics education researchers have focused on how children reason about particular topics such as averages or graphs. There has also been work on how K-12 teachers understand and reason about statistics. Recent statistics education research efforts have focused on comparisons of two types of instruction (e.g., computers vs. lecture) or prediction of achievement based on mathematical ability, attitudes, and other variables. Focus has also been directed at evaluating student ratings of new implementations. While all three types of research studies offer interesting results, they are often limited in their generalizability and validity.

A concern regarding many of these studies is they lack cross-disciplinary or cross-institutional collaboration. While there are some large conferences, there appears to be little connection among researchers in the years between these conferences. In an effort to coordinate these research efforts, there are several research study groups aimed at statistics education. While these organizations have been conducting and supporting statistics education research, they are relatively new (so support is not extensive) and they have not been as focused on establishing standards of research, or researching effective methods. Thus, further discussion is needed on how to legitimize statistics education as a research domain and how to train future researchers in statistics education.

3. Limitations of Current Research Activities

The discipline of statistics education is relatively young, with researchers being trained in a variety of different disciplines (Batanero, et al, 2000). Moreover, little is known about, or has been published on, the methodology of statistics education research (Jolliffe, 1998). As statistical education research becomes a more visible discipline, it is vital to consider appropriate research techniques. For example, randomized comparative experiments, while the standard in medical and physical science studies, may not be the most appropriate tool for statistics education. One reason is that pure randomization can seldom be achieved in an educational setting, especially for a long-term (even semester long) investigation. Consequently, variables such as instructor attitude and classroom culture have dramatic effects on student achievement and attitude and cannot be controlled as in a laboratory setting and self-selection into sections is not a sufficient substitute for randomization. Student mobility between sections and drop rates are also serious difficulties. It may also not be feasible to create the classroom implementation that is of interest. For example, in deciding what factors most directly affect student achievement, there are ethical issue involved in manipulating variables that could adversely affect student grades. These examples help illustrate that the classroom environment is a sufficiently complex and dynamic world that is often not well described by traditional research techniques.

Furthermore, this type of research too often ignores the integral role played by the classroom instructor. An outsider may also not have sufficient understanding of the subject matter to fully appreciate and document the students' experiences. In addition, the standard research model typically leads to significant time delays before the research can be applied to the classroom environment, and often the results from a controlled laboratory setting are not immediately relevant. Directly involving the classroom participants in the research (teachers, curriculum designers, students) and immersing of the researcher into the classroom environment allows for more in-depth study, more reflection, and better interplay between theory and practice.

How to measure student performance and attitudes has also been problematic. Students who do well on final exams can still demonstrate poor statistical reasoning. As Lesh and Lovitts (2000) caution, "Most existing high-impact standardized tests are poorly aligned with national standards for instruction and assessment." Current assessment methods are also not typically dynamic in nature and fail to inform the researcher of the learning processes involved. One-time measures of achievement also fail to explore the developmental nature of learning or provide concurrent feedback to the study, nor allow an iterative research approach. In particular, traditional assessment strategies do not tell us enough as to why a particular teaching method or activity works, how students' understanding and reasoning are affected or unaffected by the learning experience, or provide direction for how a teaching practice should be changed. Movement beyond multiple choice, standardized paper-and-pencil exams is crucial in order to limit the types of information we obtain and the types of research questions we can ask.

Thus, effective research requires a careful combination of appropriate research methods and assessment techniques. However, most researchers do not have background in these different areas. A statistician in particular has mostly been schooled in controlled experiments and may not be familiar with research in learning, assessment, and/or education.

4. New Directions

While the standards of research are constant, how they are met is not. Kelly and Lesh (2000) published a research handbook in an effort to inform science and math education researchers on alternative methodological approaches, such as classroom-based research, naturalistic observation, and clinical interviews. In particular, Moschkovich and Brenner (2000) outline how research standards are met in naturalistic research. For example, classroom-based research, or action research, has been defined as "ongoing and cumulative intellectual inquiry by classroom teachers into the nature of teaching and learning in their own classrooms" (Cross and Steadman, 1986).

Thus, the classroom teacher is an integral part of the research team, helping to develop the questions to be investigating and assisting in data collection. By directly involving the teachers in the research process, classroom research aims to incorporate their perspectives, insight, and understanding of the classroom culture into the analysis. This approach also allows for further probing into the student and instructor experience and adjustments in the evaluation process can be immediately implemented. The gap between theory and practice is narrowed and the evaluation becomes a dynamic process that changes in response to results and feedback, while simultaneously focusing on curricular development, instruction, and assessment. Many of these techniques can be applied equally well to statistics education research.

In considering a research methodology, it is paramount that the methodology match the research question and that a variety of tools should be employed. "Different techniques generate different types of information, and it is often the case that a single technique will not provide the breadth of information necessary to answer unequivocally the research questions under investigation" (Mestre, 2000). Relaxing the strict adherence to classical experimental methods will allow richer sources of information through complementary techniques and new research questions.

5. Example

We now provide an example of a classroom research project used to investigate how student interaction with a simulation program affects their statistical reasoning. Our goals were to understand student thinking and to inform other instructors about the use of such a program to teach the topic of sampling distributions. We began gathering data in three diverse college settings: a private college, a developmental college and a school of education. In the first stage of the research, we explored graphics-based test items to determine whether students could demonstrate a visual understanding of the implications of the Central Limit Theorem by choosing the appropriate pictures that corresponded to the sampling distribution for different sample sizes. Initially, students were asked to justify their choice of graphs and explain their reasoning. These responses were then categorized so that future instruments asked students to select which statement best represented their own reasoning. Students were given these test instruments before using the program and after using the program in order to isolate the change in understanding from interacting with the program.

In an effort to improve student performance further, we next incorporated a model of conceptual change. Students were asked to select their responses and then use the program to check their predictions. This forced students to more directly confront the misconceptions in their understanding and changes in their performance on the post-test were significant (delMas, Garfield, and Chance, 1999). This led to additional investigations in the role of pre-requisite knowledge, tools for identifying prevalent misconceptions, and further refinements to the activity. Currently, our research is focused on formulating a model of student development of statistical reasoning. Through video-tape analysis we are documenting student explanations of their reasoning using students at several levels of development to validate this model of reasoning.

This brief summary showcases how collaborative, classroom-based research can be used in statistics education. The researchers came from diverse backgrounds (education, cognition, statistics) and institutions. The subjects were their own students. This allowed direct access to the students, the ability to create the desired learning environment directly, as well as additional insight into the students' experiences. The students involved in the study were diverse in background, field of study, and age. We have employed a variety of research tools from open-ended questions, multiple choice paper-and-pencil instruments, in depth interviews, and video. The researchers have used their diverse perspectives in the development of the learning environments and to cross-check interpretations and evaluations. The investigations were tied to existing theory and are helping to generate new theory. The project has led to the development of new assessment instruments that are now available to other instructors. We feel we have contributed insight into why an activity works, have demonstrated transferability in the learning gains to other instructional settings, and are continuing to employ a variety of research methods to evaluate the progress of this project.

6. Summary

This paper has addressed new approaches to gathering data on student learning for research in statistics education. However, the gathering of data is closely connected to theoretical frameworks, prior research, and type of methods and design used. In considering ways to improve the information gathered in educational research studies in statistics, we need to keep in mind the context of the growing and developing discipline of statistics education. Toward this end, statistics education needs to establish standards for preparing researchers in statistics education. This includes recommended coursework in statistics, education, learning theory, measurement, and qualitative and quantitative research methods. The body of statistics education research also needs to become more visible and accessible across disciplines. This should include publication of examples of high quality research, literature reviews, bibliographies on certain topics, and even a high quality research handbook specific to statistics education research. Such tools will enable researchers to combine traditional methods with alternative approaches in order to best answer a wider array of research questions.

REFERENCE

Batanero, C., Garfield, J., Ottaviani, M. G., Truran, J. (2000). Research in Statistical Education: Some Priority Questions. *Statistical Education Research Newsletter*, 1(2), 2-6.

Cross, K.P. & Steadman, M.H. (1996). *Classroom Research: Implementing the Scholarship of Teaching*. San Francisco: Jossey-Bass Publishers.

delMas, R., Garfield, J. and Chance, B. (1999). Assessing the Effects of a Computer Microworld on Statistical Reasoning, *Journal of Statistics Education*, vol 7(3).

Garfield, J. (1995). How Students Learn Statistics. *International Statistical Review*, 63, 25-34.

Jolliffe, F. (1998). What is research in statistical education? In L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee, & W. K. Wong (Eds.), *Proceedings of the Fifth International Conference on Teaching Statistics* (pp. 801-806). Singapore: International Statistical Institute.

Kelly, A. & Lesh, R. (Eds.), (2000). *Handbook of Research Design in Mathematics and Science Education* (pp. 361-512). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Lesh, R., & Lovitts, B. (2000). Research Agendas: Identifying Priority Problems and Developing Useful Theoretical Perspectives. In A. Kelly and R. Lesh (Eds.) *Handbook of Research Design in Mathematics and Science Education* (pp. 45-71). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Mestre, J. (2000). Progress in Research: The Interplay Among Theory, Research Questions, and Measurement Techniques. In A. Kelly and R. Lesh (Eds.) *Handbook of Research Design in Mathematics and Science Education* (pp. 151-168). Mahwah, NJ: Lawrence Erlbaum Associates.

Moschkovich, J. N. & Brenner, M. (2000). Integrating a Naturalistic Paradigm into Research on Mathematics and Science Cognition and Learning. In A. Kelly and R. Lesh (Eds.) *Handbook of Research Design in Mathematics and Science Education* (pp. 361-512). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

RESUME

En évaluant la compréhension conceptuelle des étudiants, les capacités de raisonnement et les attitudes, on a besoin de méthodes alternatives pour rassembler des données sur les étudiants , données allant au-delà de la computation et manipulation des chiffres. Cette étude présente et critique des méthodes alternatives d'obtention de données de recherche sur la façon dont les étudiants développent une compréhension des statistiques, ceci y compris la recherche faite dans les classes et les observations et entretiens d'étudiants sur vidéocassettes.