Studying the Role of Simulation in Developing Students' Statistical Reasoning

Zieffler, Andrew
University of Minnesota, Department of Educational Psychology
Room 206 Burton Hall
178 Pillsbury Dr. SE
Minneapolis, MN, 55455, USA
E-mail: zief0002@umn.edu

Garfield, Joan B.
University of Minnesota, Department of Educational Psychology
Room 206 Burton Hall
178 Pillsbury Dr. SE
Minneapolis, MN, 55455, USA
E-mail: jbg@umn.edu

Introduction

Much of the literature on the teaching and learning of statistics has identified statistical reasoning as an important outcome in an introductory course in statistics (e.g., Garfield, 2002; delMas, 2004.). Current recommendations (e.g., GAISE guidelines, see Franklin and Garfield, 2006) for curriculum and assessment in introductory statistics courses have also focused on the development of students' statistical reasoning. As statistics educators adapt their courses to implement these recommendations, their curricula and assessments are changing to reflect this greater emphasis on conceptual understanding and statistical reasoning rather than computations and procedures. In particular, many of these changes involve using simulations in the classroom. This paper describes a teaching experiment that took place over a one semester introductory college statistics course, in which a series of simulation activities was implemented and the development of students' statistical reasoning was examined in an effort to help evaluate the course and materials.

Literature on Using Simulations to Teach Statistics

Simulations have become increasingly common in the tertiary statistics course. Both statisticians and researchers have recommended the use of simulation within an introductory statistics course to support specific instructional efforts. (see Chance and Rossman, 2006). Simulation is now an important tool used by statisticians to solve problems, and simulations can also help students visualize and build a deep understanding of difficult and abstract statistical concepts (Burrill, 2002; Maxara & Biehler, 2006). Simulations allow students to see dynamic processes, rather than static figures and illustrations. Simulations are one of the best ways to promote a deeper understanding of statistical concepts by allowing students to pose "what if" questions and test those questions using data (Garfield & Ben-Zvi, in preparation). For example, what happens to the shape, center, and spread of a sampling distribution if the sample size is increased? Simulations can also be used to help students understand random processes and outcomes, seeing that a random variable can have an unpredictable outcome yet have a predictable pattern over the long run.

Simulations allow students a way to informally address questions involving statistical inference, before formally studying this topic later in class. For example, in a class designed experiment to test whether students can correctly identify Coke or Pepsi in a blind taste test, students can compare the experimental results to what might have happened due to chance, determining whether their result is just due to chance, or leads them to believe something else. This can be done without any formal hypothesis testing, by simulating

what data would result if a student was just guessing.

Simulations provide a way to actively engage students in making and testing conjectures about data, developing their reasoning about statistical concepts and procedures. Simulations can encourage students to develop their analytical reasoning by requiring them to analyze processes and set up a series of steps for predicting outcomes (i.e., modeling) (Gnanadesikan, Scheaffer, & Swift, 1987). For example, students can be asked to predict what will happen if a "One Son" policy is enforced, limiting families to one son. They can reason about what the average family size might be and what the ratio of boys to girls would be. This can then be tested by simulating births, collecting simulated data on family size that can be graphed and summarized to answer the research question. The actual data can be compared to students' conjectures, motivating them to reason about differences between what they predicted and what might actually occur.

Despite the widespread agreement on the importance of simulation (see Scheaffer, 2001) and the numerous tools and activities available for use in teaching statistics (see Mills, 2004) there is a lack of formal evidence to support the use of simulation in developing students' statistical reasoning. In two experimental studies, Lane and Tang (2000) and Wackerly and Lang (1996) found that students who had used simulations outperformed students who had not. Schwartz, Goldman, Vye, and Barron (1998) reported a modest increase in students' performance with the use of simulations. Other research on simulation has suggested that simulation was especially effective when coupled with questions that focused students' attention to relevant features or characteristics of the simulation (e.g., delMas, Garfield, & Chance, 1999).

Overview of the Study

The main research question used to guide this study was: What is the pattern of change in students' development of statistical reasoning for students taking an introductory statistics class which involves simulations throughout the curriculum? Four sections of an undergraduate class of liberal arts students (n=120) were taught an innovative introductory statistics course in which nine in-class, small group simulation activities were embedded. Data were gathered at different times in the course to examine students' growth in statistical reasoning as a way to evaluate the course, which depended heavily on the use of the simulation activities. Because the modeling of change requires individuals to be measured in temporal sequence, a repeated-measures design was employed. Students enrolled in the course were assessed on their statistical reasoning three times during the semester: on the first day of class, after the midterm exam, and on the last day of class.

Subjects/Setting

The study participants consisted of n=120 undergraduate students, each enrolled in one of four sections of a one-semester, non-calculus based introductory statistics course taught in a College of Education at a large Midwestern university in the United States during the spring semester of the 2006/2007 school year. Three different instructors taught these four sections. The class met in a computer lab for an hour and fifteen minutes twice a week. The students were typically female social science majors (90% females and 10% males) who were enrolled in the course to complete part of their major requirements.

This particular introductory statistics course was designed so that it was aligned with recent Guidelines for the Assessment and Instruction in Statistics Education (GAISE) endorsed by the American Statistical Association (ASA; American Statistical Association, 2005). In addition, the course materials were based on what has been learned from the research literature on teaching and learning statistics (e.g., Garfield & Ben-Zvi, in preparation). The research guided both the structure of the course (i.e., scope and sequence) and the instructional methods used within the course.

Lesson plans for every instructional session, which included class goals, discussion questions and a sequence of activities, were used by the instructors to help provide greater consistency across the multiple sections of the course. The three instructors also met regularly throughout the duration of the course to further maintain consistency between the sections.

Simulation Activities

Nine in-class simulation activities were used throughout the course. These simulations took one of three forms: (1) physical simulations of data to answer a research question, (2) computer simulations to illustrate an abstract concept or process, or (3) simulations used to informally test whether an extreme sample result could be attributed to chance or to some other factor. The computer simulations used either the *Sampling SIM* software tool (delMas, 2001) or applets at the http://www.rossmanchance.com/applets website. The simulations were integrated throughout the entire course, from the first unit on Data to a final unit on Inference. Each of the nine simulation activities are described below in the order in which they were introduced in the class.

1. The *One-Son Modeling* Activity. Students address the question of what would happen if a country had a policy that limited each family to one son. How would this affect average family size and the ratio of boys to girls? Students first generate simulated data (boy and girl births) using slips of paper, then coin tosses, and finally a computer simulation program. The data are graphed and examined to answer the two questions of interest.

2. *Let's Make a Deal* Activity. Students explore two strategies to determine if one strategy has a higher chance of wining than the other. They model the two strategies using playing cards to generate data. They also simulate the activity using a Web applet to see if the trends become clearer with more data. The resulting data, which are often counterintuitive, are used to help students see that their intuitions about chance events may be incorrect, and further that simulation is a powerful and useful way to generate data to estimate probabilities.

3. *Sampling Words* Activity. Students take what they believe is a representative sample of words (a judgment sample) from a given passage (population). They also draw a random sample of words from the same passage. Each student's average word length is plotted for both samples and differences between the two plots are examined (especially in reference to the population average). Students then use a Web applet to simulate the drawing of several random samples of words which helps demonstrate several desired qualities of random samples.

4. *Cola Taste Test* Activity. Students design and carry out an experiment which consists of multiple blind taste tests of two colas, to determine if students in the class are able to correctly identify the colas. After the randomized experiment is compete and data are gathered, students use software to simulate data so that they have a distribution of possible results to compare their sample to, in order to judge if it is surprising or not.

5. *Sampling Reese's Pieces* Activity. Students make and test conjectures about the proportion of orange colored Reese's Pieces candies. They take physical samples from a population of colored candies (Reese's Pieces) and construct distributions of sample proportions. Students then use a Web applet to generate a larger number of samples of candies, allowing them to examine the distribution of sample proportions for different sample sizes.

6. *Generating Sampling Distributions* Activity. Students use two different Web applets to simulate samples and sampling distributions of sample means from two populations, one that is skewed right (word lengths) and one that is skewed left (dates on pennies). Students generate sampling distributions of the sample mean for increasingly larger sample sizes and use them to investigate and describe any predictable or common patterns that emerge, despite the differing original population shapes.

7. *Describing the Predictable Pattern: The Central Limit Theorem* Activity. Students use *Sampling SIM* Software to investigate the impact of sample size and population shape on the shape of the sampling distribution, as well as to distinguish between sample size and number of samples.

8. *Balancing Coin* Activity. Students gather data using sets of 10 coins balanced on their edges to count the proportion of times they fall landing head side up. This proportion is used to test a null distribution based on equally likely outcomes. The idea of the *p*-value is examined visually and conceptually, and then simulated *p*-values are found using *Sampling SIM* software. The argumentation metaphor is also used to help explain the logic of testing hypotheses.

9. *Reasoning about Confidence Intervals* Activity. Students develop their reasoning about confidence intervals by using simulation to make and test conjectures about factors that affect confidence intervals (e.g., level of confidence, sample size). They also have opportunities to discuss common misconceptions as they critique interpretations of confidence intervals.

Assessment Instrument

To measure change in students' statistical reasoning, the Statistics Thinking and Reasoning Test (START) was administered three times during the semester. This scale consists of 14 multiple-choice items from the Comprehensive Assessment of Outcomes in a First Statistics Course (CAOS; Garfield, delMas, & Chance). The 14 items were identified through a principal components analysis (using a Varimax rotation) performed on CAOS data gathered in Fall 2005 and Spring 2006 (n = 1470). An Alpha Coefficient based on these 14 items from that data set was calculated to be 0.74. A review of these items also suggested that they were representative of the content domain of CAOS and measured both important and difficult aspects of students' statistical reasoning.

Results

Data were gathered during Spring semester 2007 and were analyzed using linear mixed-effects models (LMM) to examine these patterns of students' development of statistical reasoning. The results will be presented at the ISI conference, along with a discussion and implications for teaching as well as for future research.

REFERENCES (RÉFERENCES)

American Statistical Association. (2005). *GAISE Endorsement*. Retrieved Feb. 9, 2007, from ASA Endorsement website: http://www.amstat.org/education/gaise/ASAEndorse.htm.

Burrill, G. (2002). Simulation as a tool to develop statistical understanding. Paper Presented at the Sixth International Conference on Teaching Statistics, Cape Town, South Africa.

Chance, B, & Rossman, A. (2006). Using simulation to teach and learn statistics. Paper Presented at the Seventh International Conference on Teaching Statistics, Salvador da Bahia, Brazil.

delMas, R. (2001). *Sampling SIM* (version 5). http://www.tc.umn.edu/~delma001/stat_tools/.

delMas, R. (2004). A comparison of mathematical and statistical reasoning. In D. Ben-Zvi and J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 79-95). Dordrecht, The Netherlands: Kluwer Academic Publishers.

delMas, R. C., Garfield, J., & Chance, B. L. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education, 7(3),* [Online].

Franklin, C. A., & Garfield, J. B. (2006). The *GAISE* project: Developing statistics education guidelines for grades pre-K–12 and college courses. In G. Burrill (Ed.), *Sixty-Eighth NCTM Yearbook: Thinking and Reasoning about Data and Chance* (pp. 345-375). Reston, VA: NCTM.

Garfield (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education, 10(3),* [Online].

Garfield, J., & Ben-Zvi, D. (in preparation). The Role of Technology in Helping Students Develop Statistical Reasoning. In J. Garfield & D. Ben-Zvi (Eds.), *Developing Students Statistical Reasoning: Connecting Research and Teaching*. Key College Press.

Garfield, J., delMas, R., & Chance, B. (n.d.). *Assessment Resource Tools for Improving Statistical Thinking*. Retrieved Feb. 9, 2007, from https://app.gen.umn.edu/artist/index.html.

Gnanadesikan, M., Scheaffer, J. L., & Swift, J. (1987). The Art and Techniques of Simulation. Quantitative Literacy Series. Palo Alto, CA: Dale Seymour Publications.

Lane, D. M., & Tang, Z. (2000). Effectiveness of simulation training on transfer of statistical concepts. *Journal of Educational Computing Research, 22(4),* 383-396.

Maxara, C., & Biehler, R. (2006). Students' probabilistic simulation and modeling competence after a computer-intensive elementary course in statistics and probability. Paper Presented at the Seventh International Conference on Teaching Statistics. Salvador da Bahia, Brazil.

Mills, J. D. (2004). Using Computer Simulation Methods to Teach Statistics: A Review of the Literature. *Journal of Statistics Education 10(1),* [Online].

Scheaffer, R. (2001). Statistics education: Perusing the past, embracing the present and charting the future. *Newsletter for the Section on Statistical Education, 7(1),* [Online].

Schwartz, D. L., Goldman, S. R., Vye, N. J., Barron, B. J., & The Cognition Technology Group at Vanderbilt. (1998). Aligning everyday and mathematical reasoning: The case of sampling assumptions. In S. P. Lajoie (Ed.), *Reflections on Statistics: Agendas for Learning, Teaching and Assessment in K-12* (pp.233-273). Mahwah, NJ: Lawrence Erlbaum.

Wackerly, D., & Lang, J. (1996). ExplorStat. A paper presented at the annual meeting of the American Statistical Association in Chicago, IL., August.

RÉSUMÉ (ABSTRACT)

We present three different uses of simulation in an introductory tertiary statistics course that is aligned with current Guidelines for Assessment and Instruction in Statistics Education (GAISE, http://www.amstat.org/education/gaise/). These three types of simulations are: using random devices to model real world phenomena, using computer simulations to illustrate abstract concepts, and using computer simulations to informally test hypotheses. The repeated use of these three types of simulations throughout an introductory course is described. Assessment data will be presented to examine the effectiveness of a course which embeds these simulation activities.