DOCUMENT RESUME

ED 474 052 SE 066 922

AUTHOR	Chance, Beth; Garfield, Joan; delMas, Robert
TITLE	Developing Simulation Activities To Improve Students' Statistical Reasoning.
PUB DATE	2000-12-00
NOTE	10p.; In: Proceedings of the International Conference on Technology in Mathematics Education (Auckland, NZ, December 11-14, 2000). For full proceedings, see SE 066 920.
AVAILABLE FROM	Aukland University of Technology, PB 92006, Auckland, NZ. Tel: 64-9-907-999 X 8405; Fax: 64-9-307-9973; e-mail: mparker@aut.ac.nz or University of Auckland PB 92019, Auckland, NZ. Tel: 64-9-373-7599 X 5886; Fax: 64-9-373-7457; e-mail: minlee@math.auckland.ac.nz.
PUB TYPE	Reports - Research (143)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Statistics; Computer Uses in Education; Educational Change; Higher Education; Innovation; *Mathematics Education; *Teaching Methods
IDENTIFIERS	*Representations (Mathematics)

ABSTRACT

This paper describes a collaborative classroom-based research project at two American universities. Our research goal has been to investigate how technology can best help students understand, integrate, and apply fundamental statistical concepts, such as sampling distributions. We describe the three-year evolution of the software, activities, and assessment instruments we used to measure the impact of technology on students' conceptual understanding and to investigate effective implementation of such technologies. Key findings include the need to establish cognitive dissonance with student predictions. We hope our study serves as a model of classroom-based research for investigating the impact of technology on student learning. (Author)

Developing Simulation Activities to Improve Students' Statistical Reasoning

by

Beth Chance
Joan Garfield
Robert delMas

DEVELOPING SIMULATION ACTIVITIES TO IMPROVE STUDENTS' STATISTICAL REASONING

Beth Chance
Calif. Polytechnic State Univ.
<bchance@calpoly.edu>

Joan Garfield
The Univ. of Minnesota
<jbg@tc.umn.edu>

Robert delMas
The Univ. of Minnesota
<delma001@tcumn.edu>

This paper describes a collaborative classroom–based research project at two American universities. Our research goal has been to investigate how technology can best help students understand, integrate, and apply fundamental statistical concepts, such as sampling distributions. We describe the three–year evolution of the software, activities, and assessment instruments we used to measure the impact of technology on students' conceptual understanding and to investigate effective implementation of such technologies. Key findings include the need to establish cognitive dissonance with student predictions. We hope our study serves as a model of classroom–based research for investigating the impact of technology on student learning.

Introduction

In recent years, there has been a shift in the focus of introductory statistics courses, emphasizing skills such as the ability to interpret, evaluate, and apply statistical ideas rather than procedural calculations. Calls for reform, similar to those in mathematics and science education, also emphasize that instruction should fully incorporate genuine data, technological tools, and active learning (e.g. Cobb, 1992). Technology offers us many ways of accomplishing these goals, from finding current data through the world–wide web to more authentic statistical analyses to use of interactive, visual computer simulations. In fact, numerous visualization programs are now available (e.g., ConStatS, Hyperstat, Visual Statistics, StatPlay).

Despite the availability of these technological resources, little is known about the impact on student learning of such technologies. There are no accepted methodologies for measuring what students are gaining from their interactions with these technologies and how they are affecting students' conceptual understanding (Hawkins, 1997). Part of the problem is the need for more informative methods of student assessment. Traditional assessment too often emphasizes the final answer over the process (Garfield, 1993) and may not provide informative data either for evaluation of student performance or for research studies on the effectiveness of new instructional techniques. Instead, we need more focus on *why* a particular interaction with technology works, *how* students' understanding and reasoning are affected by the learning experience, and *implications* for how teaching practice should be changed. In this paper, we provide an example of a collaborative classroom–based research study on the effectiveness of computer simulations in guiding student construction and visualization of one fundamental statistical concept in particular – the behavior of sampling distributions. We present not only results of this study, but also an example of how classroom–based research studies can effectively inform our understanding of students' interaction with technology.

The research question

Researchers and educators have found that students and professionals often misunderstand foundational statistical ideas. Many students develop a shallow and isolated understanding of important concepts such as sample, population, distribution, variability, sampling, and sampling variability. We were concerned that many students who pass a statistics course do not develop the deep understanding needed to integrate these concepts and apply them in their reasoning. A particularly difficult topic for our students has been the concept of sampling distributions. We found their failure particularly troublesome as this topic is the gateway to understanding the process of statistical inference. We felt that a visual simulation program could be an effective way to improve student learning about sampling distributions.

The *Sampling Distributions* program, developed by delMas (see website below), allows students to visually explore sampling distributions, in a dynamic, interactive environment. Students change parameters and then run simulations in order to directly see the effects of these changes. Development of this software was guided by literature on conceptually enhanced simulations (e.g. Nickerson, 1995; Snir, Smith, & Grosslight, 1995). An accompanying activity was developed to guide students through the interaction with the software based on ideas from literature in learning and cognition (e.g. Holland, Holyoak, Nisbett, & Thagard, 1987; Perkins, Schwartz, West, & Wiske, 1995). The three authors began using the software and activity in our introductory statistics service courses, allowing us to compare results from diverse institutions: a private liberal arts college, a College of Education, and a Developmental Education College. A wide variety of student majors and backgrounds enroll in these courses. In all three settings, students were expected to have read the appropriate textbook chapter on sampling distributions and the Central Limit Theorem prior to the activity. Students also engaged in a hands–on simulation demonstrating the Central Limit Theorem during the class period prior to using the program. Our goal was to document the learning gains of the students from use of the *Sampling Distributions* program, beyond what they learned from our normal textbook and lecture instruction.

Stage one

To assess the effects of the program and activity on students' conceptual understanding of sampling distributions, we initially focused on students' ability to demonstrate a visual understanding of the Central Limit Theorem's implications. Students were provided with a picture of a population distribution, and were asked to choose among several candidate graphs as a resulting (simulated) sampling distribution for a sample mean from that population. They then chose again for a different sample size. This was done for several population shapes. In pilot tests, students were asked to explain their selection. These open–ended responses were then categorized into several common explanations. In later tests, students were asked to choose among these potential explanations for their graph choice. From student responses and graph choices, we were able to identify several different types of reasoning.

- Correct Reasoning: Students chose the correct histograms and explanations.
- Good Reasoning: Students made reasonable choices (e.g., the sampling distribution for the larger sample size was more normal looking and had less variability than the sampling distribution for the smaller sample size) but demonstrated minor errors in their thinking (e.g., choosing a graph that looks like the population when $n>1$).
- Larger to Smaller Reasoning or Smaller to Larger Reasoning: Students attended to the change in variability but did not correctly predict the amount of variability or did not correctly pick the normal shape of the sampling distribution.

These categories covered about 80–90% of the responses for each problem, but there were also a variety of other, less frequent, responses (e.g. choosing the same histogram for both sample sizes). To determine the change in understanding due to interaction with the *Sampling Distribution* program, students were given a pre–test before using the program (but after standard classroom instruction), and then a post–test of comparable items.

After pilot testing the assessment instrument with students, revised instruments were administered to 79 students at the private college and 22 students at the College of Education during Winter, 1997. Eighty–nine students who gave responses to all pretest and posttest items were used for the analyses. (See delMas, Garfield, and Chance, 1999 for more details.) Over five different population shapes, the average percentage of "correct" or "good reasoning" choices on the pretest was 22%. This increased to 49% on the posttest. While this is considerable improvement, students were still demonstrating some definite misconceptions; e.g., confusion between the sample distribution and the sampling distribution, and interpretation of "variability." We learned that well–designed software with clear directions does not ensure sufficient student engagement or change in conceptual understanding.

Stage two

The above results led to alterations in the software and the accompanying activity. The main adjustment, inspired by a model of conceptual change (Posner, Strike, Hewson, and Gertzog, 1982), was to use the pre–test to guide student interaction with the software. Research indicates that people are generally resistant to change and are likely to find ways to either assimilate information or discredit contradictory evidence rather than restructure their thinking in order to accommodate the contradictions (Lord, Ross, & Lepper, 1979; Jennings, Amabile, & Ross, 1982; Ross & Anderson, 1982). Modern information processing theories (e.g. Holland, Holyoak, Nisbett, & Thagard, 1987) suggest that it may be necessary to direct attention toward the features of the discrediting experience in order for the contradictory evidence to be encoded. Left to their own devices, people will attend only to those features predicted by their current information structure. Adapting this approach, we had students make predictions on the pre–test, and then use the software to compare their answers by embedding the assessment instrument into the activity (e.g. students were asked to comment on how the correct graph compared to the graph they chose). When students discover that their prediction is incorrect, this creates cognitive dissonance between the students' current knowledge or expectation and what they are seeing. Students are then able to utilize the software to identify and correct their misconceptions.

Assessment results for a total of 141 students using the new activity at both schools showed that on average, students used correct or good reasoning on 16% of the pretest items (similar to before), but correct or good reasoning on 72% of the posttest items (delMas, Garfield, & Chance, 1999). These results agree with other research results that students learn better when activities are structured to help students evaluate the difference between their own believes and actual results (e.g. delMas and Bart, 1989). Furthermore, the activity allowed us to better track student misconceptions, and what knowledge was lacking in their understanding of sampling distributions. We then altered the activity to better address the most prevalent misconceptions.

Stage three

Our results indicated that students still struggled with the notion of sample, variation, and even histogram. We feel that without these concepts, students are not able to develop a deep understanding of sampling distributions. To help determine whether students are cognitively ready to learn about sampling distributions, we developed a pre–test of basic skills that highlights common misconceptions in prerequisite knowledge. For example, our studies had shown that students often confuse "bumpiness" of a histogram with "variability," and may not properly use statistical terminology such as "normal" vs. "even." The pre–test assessment allows the instructor to correct these misconceptions before using the *Sampling Distribution* software. We also embedded the activity into a contextual example in order to help students learn to apply the implications of the Central Limit Theorem. We again administered the post–test in our different institutional settings and compared post–test scores (55 students) on the graphic based questions for two population shapes to scores from previous versions of the activity. However, these results were not as impressive with only about 60% of students demonstrating good or correct reasoning. Some possible explanations include:

- Insufficient development and definition of sampling distributions in lecture prior to use of the computer program (this varied at the three schools).
- A decreased level of student engagement with the "prediction questions." In Stage Two, the pre–test questions were turned in to the instructor for marking before students used the program. In Stage Three, the activity relied on the student to invest sufficiently in the activity to create significant dissonance.
- The longer contextual activity may have required students to attend to more information than is feasible in one interaction with the software.
- The Stage Three activity did not include as many "prediction questions."

Stage four

Last year, interviews were conducted with students to gain a more in–depth understanding of their statistical reasoning about variability, samples, and sampling distribution (see also Garfield, 2000). The students were enrolled in a graduate–level introductory course in the College of Education and Human Development at the University of Minnesota. Interviews, which lasted from 45 to 60 minutes, asked participants to respond to several open–ended questions about variability and sampling and were guided through an interactive activity with the *Sampling Distributions* software. The interviews were videotaped, transcribed, and viewed many times as we tried to determine students' initial understanding of how sampling distributions behave and how feedback from the computer simulation program helped them develop an integrated reasoning of concepts. We found ourselves identifying stages that the students went through as they progressed from faulty to correct reasoning about sampling distributions. This led us to propose a framework that describes the development of students' statistical reasoning about sampling distributions. This framework is an extension of one developed by Graham Jones and colleagues to capture the statistical thinking of middle schools students (Jones, Langrall, Thornton & Mogill, 1997; Jones, Thornton, Langrall, Putt, & Perry, 1998; Tarr & Jones, 1997).

Level 1: Idiosyncratic Reasoning: The student knows words and symbols related to sampling distributions, uses them without fully understanding them, often incorrectly, and may scramble them with unrelated information.

Level 2: Verbal Reasoning: The student has a verbal understanding of sampling distributions and the Central Limit Theorem, but cannot apply this to actual behavior. For example, the student can select a correct definition, but does not understand how key concepts such as variability and shape are integrated.

Level 3: Transitional Reasoning: The student is able to correctly identify one or two dimensions of the sampling process without fully integrating these dimensions; e.g., the relationship between the population shape and the shape of the sampling distribution, the fact that large samples lead to more normal looking sampling distributions, the fact that larger samples lead to narrower sampling distributions.

Level 4: Procedural Reasoning: The student is able to correctly identify the dimensions of the sampling process but does not fully integrate them or understand the process. For example, the student can correctly predict which sampling distribution corresponds to the given parameters, but cannot explain the process, and does not have full confidence in predictions.

Level 5: Integrated Process Reasoning: The student has a complete understanding of the process of sampling and sampling distributions, coordinates the rules and behavior. The student can explain the process in their own words and predicts correctly and with confidence.

The current stage

Our current research focuses on the validation and possible extension of the above framework to other areas of statistical reasoning and to students at the secondary and tertiary level. We believe that in order for students to fully understand sampling distributions, they need to experience a variety of activities: text or verbal explanations, concrete activities involving sampling from finite populations, and interactions with computer–simulated populations and sampling distributions when the parameters are varied. This contradicts some of the psychology research that argues for teaching specific training rules.

We are currently developing activities that integrate the *Sampling Distribution* software earlier in the course. One aim is to provide the students with more familiarity with the program prior to the sampling distribution topic. We hope this will allow students to better focus on the statistical concept, having already learned the software. The second aim is to use the visualization capabilities of the program to between develop a correct and full understanding of foundational concepts, e.g. variation, sample distribution vs. sampling distribution. Students will construct prerequisite knowledge using a predict–and–test environment throughout the course. We are also trying to explore activities that help students develop the ideas "process" and "model" earlier and throughout the course. Finally, we are also expanding our collection of follow–up application questions to test students' ability to apply the knowledge gained from their interaction with the software in new settings.

Research and assessment

The above research presents an example of classroom–based research (e.g., Cross & Steadman; 1996, see also Kelly & Lesh, 2000) in the context of an introductory statistics course. We believe this is an exciting and productive model for research on the effects of technology as an instructional tool. Classroom–based research provides on–going, systematic evaluation in the classroom setting, narrowing the bridge between theory and practice. While classroom–based research is grounded in evidence, results are continually tied to existing theory and generative of new theory. It is a dynamic process that allows the questions to change in response to results and feedback, while simultaneously focusing on curricular development, instruction, and assessment.

While our students cannot be considered a random sample of all introductory statistics students, we have taken several steps to enhance the quality of our study. Working at different universities we have ensured multiple perspectives, diverse instructional settings and student audiences, and multiple time points. Our project, while focusing on our experiences as teachers, also combined our expertise in cognition, educational psychology, and statistics. We also brought in, and hopefully expanded, research results from other areas, such as cognition, learning theory, and information processing theory. While we have not identified a definitive approach to teaching sampling distributions, our research has provided substantial insight into students' misconceptions and their sources. We believe we are developing understanding about why an activity works, how students' understanding and reasoning are effected, and how prior knowledge affects their experience with the technology.

Furthermore our results have demonstrated the instructional uses of assessment. By embedding the assessment into the learning activity, we were able to strengthen the students' level of engagement with the technology. This assessment approach also takes advantage of the dynamic, immediate feedback nature of the technology. By indicating students' short–term and long–term understanding to the instructor, and by providing the students with more immediate feedback on their own understanding, assessment can provide a very powerful teaching tool.

Conclusion

Statistics instructors have been very excited about how advances in technology have dramatically changed what we can do in our courses. For example, shifting the computational burden to computers and calculators allows more time to focus on conceptual understanding and other reform goals. However, recent research is illustrating that quality programs and simulations are not enough to ensure cognitive change. For example, the establishment of cognitive dissonance appears to be a crucial component to effective interaction with technology, providing students with the opportunity to immediately test and reflect on their knowledge in an interactive environment. However, it is less clear what level of student engagement is necessary to promote cognitive dissonance. We also found that prerequisite knowledge plays a large role in students' ability to learn from technology. Indeed our sampling distribution research results have had numerous implications on instruction of topics earlier in the course (e.g. more emphasis on understanding of variability). We have also begun using a developmental model of reasoning to help us identify and improve a student's level of reasoning throughout the course.

As we continue to examine these issues, new assessment instruments need to be developed that better examine students' *process reasoning*, beyond their *verbal reasoning*. We also need to take full advantage of the role of assessment as an instructional and research tool. We encourage more classroom–based research done carefully, collaboratively, and over time, to effectively provide insight into why an interaction with technology works, improve understanding of the processes involved, and develop knowledge of similarities and differences across multiple instructional setting, while suggesting changes for improved teaching practice and ongoing research.

Acknowledgments

This research was supported by a Grant–In–Aid from the University of Minnesota and a National Science Foundation grant from the Division of Undergraduate Education (DUE–9752523). Materials produced under this grant can be obtained at http://www.gen.umn.edu/faculty_staff/delmas/stat_tools/index.htm.

References

Cobb, G. (1992). Teaching Statistics. In L. Steen (Ed.), *Heeding the call for change: Suggestions for curricular action* (pp. 3–43). Mathematical Association of America Notes, No. 22.

Cross, K. P., & Steadman, M. H. (1996). *Classroom research: Implementing the scholarship of teaching.* San Francisco: Jossey–Bass Publishers.

delMas, R. & Bart, W.M. (1989). The role of evaluation exercise in the resolution of misconceptions of probability. *Focus on Learning Problems in Mathematics*, 11(3), 39–54.

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

Garfield, J. (2000). The role of statistical reasoning in learning statistics. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, April, 2000.

Garfield, J. (1993). An authentic assessment of students' statistical knowledge. In N. Webb (Ed.), *National Council of Teachers of Mathematics 1993 Yearbook: Assessment in the Mathematics Classroom* (pp. 187–196). Reston, VA: NCTM.

Hawkins, A. (1997). Myth–Conceptions! In J. Garfield and G. Burrill (Eds.). *Research on the Role of Technology in Teaching and Learning Statistics*. Voorburg, The Netherlands: International Statistical Institute.

Holland, J. H., Holyoak, K.J., Nisbett, R.E., and Thagard, P.R. (1987). *Induction: Processes of Inference, Learning, and Discovery*. Cambridge, Massachusetts: The MIT Press.

Jennings, D., Amabile, T., and Ross, L. (1982). Informal covariation assessment: Data–based versus theory–based judgments. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases* (pp. 211–230). Cambridge: Cambridge University Press.

Jones, G. A., Langrall, C. W., Thornton, C. A., & Mogill, A. T. (1997). A framework for assessing and nurturing young children's thinking in probability. *Educational Studies in Mathematics*, 32, 101–125.

Jones, G. A., Thornton, C. A., Langrall, C. W., Putt, I. J., and Perry, B. (1998). *A framework for assessing and fostering students' statistical thinking*. Manuscript submitted for publication.

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

Lord, C., Ross, L., and Lepper, M. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology* 37, 2098–2109.

Nickerson, R. S. (1995). Can technology help teach for understanding? In D. N. Perkins, J. L. Schwartz, M. M. West, & M. S. Wiske (Eds.), *Software Goes to School: Teaching for Understanding with New Technologies* (pp. 7–22). New York: Oxford University Press.

Posner, G.J., Strike, K.A., Hewson, P.W., and Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211–227.

Ross, L. and Anderson, C. (1982). Shortcomings in the attribution process: On the origins and maintenance of erroneous social assessments. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases* (pp. 129–152). Cambridge: Cambridge University Press.

Snir, J., Smith, C., and Grosslight, L. (1995). Conceptually enhanced simulations: A computer tool for science teaching. In D. N. Perkins, J. L. Schwartz, M. M. West, & M. S. Wiske (Eds.), *Software Goes to School: Teaching for Understanding with New Technologies* (pp. 106–129). New York: Oxford University Press.

Tarr, J. E., and Jones, G. A. (1997). A framework for assessing middle school students' thinking in conditional probability and independence. *Mathematics Education Research Journal*, 9, 39–59.

ConStatS: http://www.tufts.edu/tccs/services/css/ConStatS.html

Hyperstat: http://davidmlane.com/hyperstat/index.html

Visual Statistics: http://www.mhhe.com/business/opsci/doane/

StatPlay: La Trobe University, Australia

Sampling Distributions: http://www.gen.umn.edu/faculty_staff/delmas/stat_tools/stat_tools_software.htm

NOTICE

Reproduction Basis

X	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").