A RESEARCH-BASED STATISTICS COURSE FOR TERTIARY STUDENTS

JIYOON PARK
University of Minnesota
parkx666@umn.edu

ROBERT C. DELMAS
University of Minnesota
delma001@umn.edu

ANDREW ZIEFFLER
University of Minnesota
zief0002@umn.edu

JOAN GARFIELD
University of Minnesota
jbg@umn.edu

ABSTRACT

The NSF-funded CATALST project has developed a radically different undergraduate introductory-statistics course that uses randomization and resampling approaches as the only methods for statistical inference. This course is based on research in cognitive science, mathematics and engineering education, as well as in statistics education. A carefully designed sequence of activities was used to enable students to develop an understanding of randomness, chance models, randomization tests and bootstrap coverage intervals as well the authentic models of statistical thinking. This paper outlines the research foundations that were used to develop the course, the year-long teaching experiment that was used to pilot, study, and revised the course, and plans for gathering large sets of quantitative and qualitative data to evaluate the impact of the course on important student outcomes.

Keywords: *introductory-statistics course; randomization and resampling, statistical inference, MEA*

1. INTRODUCTION

The CATALST course, developed as part of a three-year NSF-funded teaching experiment (Collaborative Research: The CATALSTS Project, Change Agents for Teaching and Learning Statistics, DUE-0814433) incorporates a different approach to introducing and building ideas regarding statistical inference. Inspired by a talk given by George Cobb (2005), the CATALST course was created to use the ideas of simulation and randomization-based methods to introduce students to ideas of statistical inference. The CATALST course immerses students in the process of statistical inference from the first day of the course. From the beginning, students create models to simulate samples of data, examine distributions of statistics from repeated simulated samples, and make informal, then formal inferences regarding particular observed sample statistics. In this paper we describe the development, pilot testing, and preliminary evaluation of the CATALST course, which was designed to develop students' inferential reasoning and to introduce a non-traditional approach for formal inference. We also describe the research and evaluation methods planned to determine the impact on important students outcomes.

2. BACKGROUND AND RESEARCH FOUNDATION

In science, *catalysis* is the acceleration (increase in rate) of a chemical reaction by means of a substance called a catalyst. In more general terms anything that accelerates a process may be called a "catalyst." The authors of this paper associate catalysts with actions that lead to some profound or major change and use the acronym CATALST to represent the goal of accelerating change in the teaching and learning of statistics.

The CATALST project created curricular materials based not only on Cobb's ideas regarding randomization-based inference, but also on research in cognition and learning, and instructional design principles. The project was based on the following, important foundations:

Adapting and Implementing Innovative Materials in Statistics: The AIMS Project This exploratory project adapted and implemented activities based on several types of innovative materials (textbooks, software, web resources, and special simulation tools) produced in the past few years for introductory statistics courses. The development of 26 detailed lesson plans and one to three activities per lesson were built on the GAISE guidelines for teaching introductory statistics courses as well as on instructional design principles based on cognitive theory (Cobb & McClain, 2004). These principles worked well for creating effective activities that engage students in active learning and in developing and expressing their reasoning. Despite positive results, we have noted that one thing missing from AIMS is a structure that interconnects the activities and also fosters the development of inferential reasoning. Based on research findings from studying model eliciting activities in mathematics education (Lesh & Doerr, 2003) and cognitive studies on the role of invention in conceptual development (Schwartz et al., 2007) we expect the following components to address these limitations and lead to improved students outcomes.

Model Eliciting Activities (MEAs) Improving Engineering Students Learning Strategies through Models and Modeling Project MEAs are open-ended problems that are designed to encourage students to build mathematical models in order to solve complex problems, as well as provide a means for educators to better understand students' thinking. MEAs are created to look like authentic, real-world problems and require students to work in teams of three to four students to generate solutions to the problems via written descriptions, explanations and constructions by "repeatedly revealing, testing, and refining or extending their ways of thinking" (Lesh, Hoover, Hole, Kelly, & Post, 2000, p. 597). See Lesh et al. (2000) for the six principles and details about MEAs. We expect that MEAs will help students unify basic ideas in major units of the course, lead to better conceptual understanding, problem solving, retention, and transfer. MEAs should also help develop students' problem solving skills and ability to work together in solving problems.

Invention for Transfer: The work of Dan Schwartz We think that the MEAs can be enhanced by utilizing principles of invention from the work of cognitive psychology (NSF REC grant, *Inventing to Prepare for Learning,* Schwartz et al., 2007). This work examines students' prior knowledge and intuitions that conflict with what they are learning, something that has often been studied in statistics education research. Trying to replace these concepts with a new piece of knowledge does not lead to a better understanding of the concept. Schwartz's work reveals the need for learning activities that help students work through inconsistencies in their prior knowledge and intuitions. Though it may be possible to teach people the right performance, the underlying inconsistencies may remain and yield misconceptions that interfere with students' abilities to learn new material. We utilize an instructional model (developed by Schwartz and his colleagues) that directly confronts students' misconceptions and intuitions, as they solve carefully designed problems.

Assessments Integrated and Aligned with Instruction (ARTIST and CSI Projects) Assessment is an important part of teaching and learning statistics. Innovative Items and instruments from the NSF-funded ARTIST and CSI projects, that better focus on key conceptual understanding, were adapted and modified for assessment. In addition, the MEAs offer a valuable way to gather authentic assessment data that reveals students' reasoning and developing statistical thinking (Lesh & Lamon, 1992; Lesh et al., 2000; Wiggins 1998). We used both individual and group assessments, building on the highly successful group quiz structure used in the AIMS curriculum. This feature emphasizes and reinforces the positive benefits of cooperative learning and problem solving, a key aspect of the MEAs and the AIMS activities.

3. THE CATALST COURSE

The content of the coures originally was inspired by Cobb (2005, 2007) and has been further influenced by the work of additional research in statistics education (e.g.., Zieffler et al., 2011). The course emphasizes a conceptual approach to topics, while still immersing students in statistical thinking and, in the paraphrased words of Alan Schoenfeld (1998), "teaching students to cook rather than only follow recipes".

The CATALST course consists of three units: (1) Chance Models and Simulation, (2) Models for Comparing Groups, and (3) Accuracy of Models. While inference is one of the key components of each unit, modeling and simulation are used to develop these ideas. Each unit begins with a model-eliciting activity (MEA; see Lesh & Doerr, 2003). The MEA, which is based on a real statistical inquiry, is an open-ended problem with no clear solution. Aside from engaging students in an interesting context, the MEA is designed to motivate and prepare students to learn the relevant content in each unit to stimulate statistical thinking. In order to solve the problem posed in the MEA, students invent and test models that promote statistical reasoning and thinking.

GENERAL SEQUENCE OF EACH UNIT

The MEA at the start of each unit presents students with a problem, such as how to decide if two groups differ meaningfully on some variable. Materials for the activity describe the problem and present data for the two groups. Students work in small groups to explore the data and then to create a way to test for a significant difference in order to address the research question. As students report their solutions, the instructor presents them with counterfactuals (e.g., additional data) where the proposed method does not lead to a clear decision, requiring students to further develop their solution so it accommodates each new instance. Working through the MEA sets the stage for the topics in that unit, which are developed through five to seven activities (modified CSI and AIMS activities) that allow students to discover the important ideas in that unit. The activities include simulations that generate multiple samples, so that students become familiar with sampling variability and how to examine distributions of sample statistics.

Each unit ends with an Integration MEA, or I-MEA. The I-MEA parallels the structure of the MEA introduced at the start of the unit. It requires students to integrate content from the follow-up activities in order to produce a solution for the original MEA. The purpose of the I-MEA is to evaluate students' ability to retain and transfer knowledge, and to view how their statistical reasoning and thinking are developing.

Also in each unit students experience a model-eliciting extension (referred to as an MXA). The MXA is presented as an item on an end-of-unit exam and is essentially a transfer problem that asks students to apply what they have learned in the unit to a slightly novel situation. For example, students learn how to test for a difference between group means in Unit 2. The MXA problem for Unit 2 asks students to apply what they have learned in a context that motivates testing for a difference between group standard deviations (e.g., are flight arrival delays more variable for one airline than another?).

4. METHODOLOGY

In this section we describe a teaching experiment conducted at the University of Minnesota with tertiary-level students enrolled in an introductory statistics course. Before this description, however, we first provide an introduction to the teaching experiment methodology (see Steffe, & Thompson, 2000).

4.1 TEACHING EXPERIMENTS

Preparing for the Teaching Experiment The first stage in a teaching experiment is preparation for the actual study. It is during this stage that the research team, which usually includes researchers and teachers, envisions how dialogue and mathematical activity will occur as a result of planned classroom activity. The researchers propose a hypothetical learning trajectory (Simon, 1995) that consists of a sequence of ideas, knowledge, and attitudes that they hope students will construct as they participate in the activities and classroom dialogue and plan instruction to help move students along this path toward the desired learning goals.

Actual Experimentation During the actual teaching experiment, the researchers test and modify their conjectures about the statistical learning trajectory as a result of their communication, interaction, and observation of students. The learning environment also evolves as a result of the interactions between the teacher and students as they engage in the content. The research team ideally meets after every classroom session to modify the learning trajectory and plan new lessons. These meetings are generally audiotaped for future reference. Because of the constant modification, detailed lesson plans cannot be made too far in advance.

Retrospective Analyses In some instances, the research team performs a retrospective analysis after each session to redirect the learning trajectory. In addition, the team performs a retrospective analysis after an entire teaching experiment has been completed. During this stage the team develops domain specific instructional theory to help guide future instruction. They also develop new hypothetical learning trajectories for future design experiments.

4.2 The CATALST Teaching Experiment

After two years of design, the first CATALST was taught during the fall 2010 semester to undergraduate students enrolled in a non-calculus based statistics course in the Educational Psychology department at the University of Minnesota. Thirty students were enrolled in this course, which was taught by a graduate student who was a member of the research team and an experienced teacher of statistics. One of the research faculty co-taught the class from time to time as needed, and various faculty and graduate students observed class sessions and took detailed notes about each activity.

In addition to the classes, which met twice a week, the graduate students and PIs at the University of Minnesota had weekly meetings to discuss, revise, and write the curriculum based on the student learning observed during the course. There were also weekly meetings among the instructional team to write assessments and grading rubrics, as well as to discuss the day-to-day pedagogical and content issues. The entire CATALST team of PIs (including colleagues in California and Ohio) and graduate assistants had monthly conference calls to discuss the course and prepare assessments.

During the first semester of this experiment, we studied three broad questions: (1) How would students respond to the demands of the course? (2) Could we prepare graduate students to teach the course? (3) How would students' reasoning about inference be developed in this setting?

In the spring semester, after making several revisions to the curriculum based on what we learned in the fall, the currciulum was again taught at the University of Minnesota, but this time in three sections by three different graduate students. It was also implemented by a tenure-track faculty member in a section of statistics for honors students at North Carolina State University. The entire instructional team, along with three CATALST PIs, met weekly to again write and revise material based on the instructors' and TA's perceptions of the students' learning in the course. Additional weekly meetings were used to plan and discuss day-to-day issues regarding the course.

Students worked in collaborative groups throughout the course where they were encouraged to make and test conjectures, ask questions, and constantly explain their reasoning (e.g., how would the average family size change if families were restricted to having only one son?).

The spring semester was spent studying (1) whether the revised sequence was more coherent and conceptually viable for students; (2) How effective the collaborative teaching model was in preparing instructors for teaching the CATALST course; (3) Whether we could take the experiences of these instructors and use them to help create lesson plans for future CATALST teachers; and (4) what could be learned from the assessment data about how our students are comparing to other students?

5. ASSESSMENT DATA

5.1 Assessments used for feedback and grades

Individual student assessments were used to provide formative evaluation information both to students as well as to the course instructors and team. Unit Tests created by the team were administered to groups of students, and were designed to see if students had achieved the overall unit learning goals. In addition, an unfamiliar "transfer" problem was included in each of the three unit exams.

5.2 Assessments used for Research and Evaluation

Two instruments were developed and used to gather summative data to be used to provide information to the project team, a comprehensive content-related exam and an attitudinal measure. A comprehensive exam consisting of two instruments—the Goals and Outcomes Associates with Learning Statistics (GOALS) and the Models of Statistical Thinking Assessment (MOST)–was developed to assess students' achievement of the broader course learning outcomes. The items used from the GOALS test included 30 multiple-choice items that assess students' statistical reasoning. The items included from the MOST test consisted of 10 open-ended items that assess students' statistical thinking. This comprehensive exam was administered at the end of the spring 2011 term to students enrolled in three sections of the CATALST course. Students also completed an 11-item attitudinal survey. The items were Likert-response items and were written to elicit students' attitudes and perceptions about the course content, as well as about the value of statistics.

6. RESULTS

Data of student learning outcomes were gathered during the semester as well as at the end of the semester. The results of these data are currently being analyzed and will be presented during the talk at the ISI.

7. SUMMARY AND NEXT STEPS

The original research questions have helped structure the teaching experiment and the collection of student outcome data. The preliminary data gathered so far suggest that we are achieving positive results and are helping students both think statistically as well as develop an appreciation for the value of statistics.

Now that the CATLST course has been piloted in four sections of tertiary statistics classes, the next step is to adapt and implement this course to other courses and institutions. Plans are being made to teach 12 versions f this course in the coming year, and to gather extensive data form students using the three measures developed for use this spring. Our plan is also to administer these instruments to many additional non-CATALST courses so that comparison data may be examined.

9. REFERENCES

Cobb, G. W. (2007). The Introductory Statistics Course: A Ptolemaic Curriculum? *Technology Innovations in Statistics Education, 1*(1). Retrieved September 28, 2010 from http://escholarship.org/uc/item/6hb3k0nz#page-1

Cobb, G. W. (2005). *The introductory statistics course: A saber tooth curriculum?* After dinner talk given at the United States Conference on Teaching Statistics.

Cobb, P. & McClain, K. (2004). Principles of instructional design for supporting the development of students' statistical reasoning. In D. Ben-Zvi and J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 375-395). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Konold, C. (2010). *TinkerPlots™ Version 2 Beta*. Key Curriculum Press.

Lesh, R., & Doerr, H. M. (2003). Foundations of a models and modeling perspective on mathematics teaching, learning, and problem solving. In R. Lesh & H. M. Doerr (Eds.), *Beyond constructivism: Models and modeling perspectives on mathematics teaching, learning, and problem solving* (pp. 3-33). Mahwah, NJ: Lawrence Erlbaum.

Lesh, R., & Lamon, S. J. (Eds.). (1992). *Assessment of authentic performance in school mathematics*. Washington, DC: AAAS Press.

Lesh, R., Hoover, M., Hole, B., Kelly, A., & Post, T. (2000). Principles for developing thought- revealing activities for students and teachers. In A. E. Kelley and R. A. Lesh (Eds.), *Handbook of Research Design in Mathematics and Science Education* (pp. 591-646). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Schoenfeld, A. H. (1998). Making mathematics and making pasta: From cookbook procedures to really cooking. In J. G. Greeno and S. V. Goldman (Eds.), *Thinking practices in mathematics and science learning* (pp. 299–319). Mahwah, NJ: Lawrence Erlbaum.

Schwartz, D., Sears, D., & Chang, J. (2007). Reconsidering prior knowledge. In M. Lovett and P. Shah (Eds.), *Thinking with Data* (pp. 319-344). New York: Erlbaum.

Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education, 26*(2), 114- 145.

Steffe, L. P., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In R. Lesh & A. E. Kelly (Eds.), *Research design in mathematics and science education (*pp. 267- 307). Hillsdale, NJ: Erlbaum.

Wiggins, G. (1998). *Educative assessment: Designing assessment to inform and improve student performance*. San Francisco: Jossey-Bass.

Zieffler, A., Garfield, J., delMas, R., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal, 7(*2), 40-58. Retrieved September 28, 2010 from http://www.stat.auckland.ac.nz/~iase/serj/SERJ7(2)_Zieffler.pdf

Zieffler, A., Garfield, J., delMas, R., Rossman, A., Chance, B., Holcomb, J., Cobb, G., Isaak, R., Le, L., & Ziegler, L. (2011, April). *It takes a village to effect change: The CATALST course teaching experiment.* Invited presentation made at the National Council of Teachers of Mathematics Research Presession, Indianapolis, IN.