DEVELOPING TERTIARY-LEVEL STUDENTS' STATISTICAL THINKING THROUGH THE USE OF MODEL-ELICITING ACTIVITIES

Joan Garfield, Robert delMas and Andrew Zieffler
University of Minnesota, United States of America
jbg@umn.edu

This paper reports on the development of specially designed Model-Eliciting Activities (MEAs) to help students develop statistical thinking. While MEAS have been successfully used in mathematics and engineering education (Lesh & Doer, 2003; Zawojewski, Bowman, & Diefes-Dux, 2008), their use in an introductory applied statistics course had not been investigated. The NSF-funded CATALST project has been studying the development and use of MEAs as a way of having students experience an authentic statistical problem that is based on a real data in order to expose students to the discipline of statistics and promote students' statistical thinking.

THE IMPORTANCE OF STATISTICAL THINKING

Statistical thinking has long been a topic of discussion and a generally agreed upon goal for statistics instruction (see Chance, 2002; Rumsey, 2002; Snee, 1999). In his influential report on the need to reform the teaching of introductory college statistics classes, Cobb (1992) wrote that:

> Any introductory course should take as its main goal helping students to learn the basic elements of statistical thinking. Many advanced courses would be improved by a more explicit emphasis on those same basic elements.

Those elements were described as: the need for data; the importance of data production; the omnipresence of variability; and the quantification and explanation of variability (Cobb, 1992).

A few years later, Moore (1998) suggested that statistics involves distinctive and powerful ways of thinking, writing that, "Statistics is a general intellectual method that applies wherever data, variation, and chance appear. It is a fundamental method because data, variation, and chance are omnipresent in modern life" (p. 134). In their landmark paper, Wild and Pfannkuch (1999) provided an empirically-based model of statistical thinking, that described the processes involved in the statistical practice of data-based enquiry from problem formulation to conclusions.

CHALLENGES IN DEVELOPING STATISTICAL THINKING

Despite agreement on the need for statistical thinking, there have been no empirically tested instructional materials or methods that can be shown to develop such an important learning outcome. In fact, studies of student outcomes in a first statistics class show an alarming lack of statistical reasoning and thinking (see delMas, Garfield, Ooms & Chance, 2007). The Guidelines for Assessment and Instruction in Statistics Education (GAISE)–endorsed by the American Statistical Association (2005)–include the goal of developing students' statistical thinking. The report also offers examples of what this type of thinking may look like and advocates the importance of teaching and modeling statistical thinking in the introductory statistics course. Furthermore, the report suggests that merely learning statistical content (i.e., terms, formulas and procedures), even with real data and research studies used as a context, does not appear to lead students to think more like statisticians along the lines described by Wild and Pfannkuch (1999). What is needed is a radically different approach that is designed to help students to begin to think statistically and to build on this thinking to understand and appreciate the discipline of statistics.

A CURRICULUM DEVELOPED TO BUILD STATISTICAL THINKING

The CATALST (Change Agents for Teaching and Learning Statistics) project was funded by the National Science Foundation (DUE-0814433) to develop and study a unique way to introduce students to statistics and to develop and assess their statistical thinking. The method of building this thinking is by engaging students in a series of carefully crafted *Model-Elicting Activities* (MEAs) that challenge students to create and test models in order to solve a complex statistical problem.

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

MODEL-ELICITING ACTIVITIES (MEAS)

MEAs were developed in the mathematics education area, and are open-ended problems that encourage students to build mathematical models in order to solve complex problems, as well as provide a means for educators to better understand students mathematical thinking (Lesh & Doerr, 2003). 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, Hole, Hoover, Kelly, & Post, 2000, p. 597). MEAs are based on the following six principles (Lesh, et al., 2000).

1. *Model construction principle:* problems must be designed to allow for the creation of a model dealing with elements, relationships and operations between these elements, and patterns and rules governing these relationships.
2. *The reality principle:* problems must be meaningful and relevant to the students.
3. *Self-assessment principle:* students must be able to self-assess or measure the usefulness of their solutions.
4. *Model documentation principle*: students must be able to reveal and document their thinking processes within their solution.
5. *Model share-ability and reusability principle:* solutions created by students should be generalizable or easily adapted to other situations and usable by others.
6. *Effective prototype principle:* ensures that the model produced will be as simple as possible yet still mathematically significant.

Model-eliciting activities can be designed so that they lead to significant forms of learning (Lesh et al., 2000), and have been used with dramatic and positive results in mathematics and engineering education (Moore, Diefes-Dux & Imbrie, 2006; 2007; Diefes-Dux, Imbrie & Moore, 2005; Zawojewski, Bowman & Diefes-Dux, 2008). Furthermore, because significant forms of conceptual development occur during relatively brief periods of time, using MEAs, it often is possible to observe the processes that students use to extend, differentiate, integrate, refine, or revise the relevant constructs. Consequently, to investigate cognitive development, it is possible for researchers to go beyond descriptions of successive states of knowledge to observe the processes that promote development from one state to another (Lesh, 1983).

While some of the MEAs already created focus on some statistical ideas, they were developed using "realistic" data and focus on the underlying mathematical models. Therefore, the CATALST project decided to develop a new type of MEA that:

a. Reflects a real type of statistical problem (e.g., making predictions using multiple variables or figuring out a way to classify objects to make predictions, based on known data)
b. Has a current and engaging context that will motivate students to work on a solution and also illustrate the relevance of statistics to their everyday lives
c. Uses real data either gathered in a research study or gathered for the purpose of the MEA.

We believe that MEAs with these properties will help prepare students to learn specific statistical content for different units of a course, and therefore, 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.

CONNECTIONS OF MEAS TO STATISTICAL THINKING

As we began to design, pilot, and study the use of MEAs in an introductory statistics class, we realized that they appeared to be stimulating students to think statistically in important ways. We observed wonderful examples where students' considered data production, dealt with the messiness of data, operationalized constructs, generated different possible solutions, saw the importance of testing their models on new data, assessed how well their methods were working, and explained and justified their solutions. We think that these are important components of statistical thinking that are not often promoted by typical activities in an introductory statistics class.

For each MEA, students are provided with a media article and individually respond to a set of readiness questions to become familiar with the context of the problem and also to begin to engage with the problem. Then, in teams of three or four, students are given the problem statement, work on a team solution to the problem, and, as a team, write their solution in a letter or memo to the client who issued the problem statement. Each team presents their solution to the class, which then discusses the different solutions, the statistics involved, and the effectiveness of the different solutions in meeting the needs of the client. Here we provide examples of two MEAs developed by our project, and point out how they lead students to think statistically.

Determining if an iPod Shuffle tool is generating random playlists This MEA asks students to work in a team to respond to a person who has written a letter to Apple to complain that the shuffle feature on his iPod is not generating random playlists because he believes it is producing playlists in which some artists are played too often and others are not played enough. Each team of students is given a set of 25 randomly generated playlists from the client's library and asked to write down characteristics of randomly generated playlists that will help them create a set of rules that flag playlists that do not appear to have been randomly generated. They are then provided with five additional randomly generated playlists on which to test and revise their rules. After they have their revised set of rules, the teams of students are provided with the client's original three playlists and asked to judge whether there is convincing evidence that the client's iPod Shuffle feature is producing playlists which do not seem to be randomly generated. Each team writes a letter on behalf of Apple, in which their set of rules is clearly outlined and also responding to the initial complaint.

Designing a SPAM filter for e-mail messages This MEA asks students to work in a team to develop a set of rules that can be used to program a SPAM filter that would determine whether or not each e-mail for a specific client is SPAM or a legitimate message based on examining the subject line of incoming email messages. Students are given a sample of 50 SPAM and 50 non-SPAM subject lines to develop a set of rules that will classify SPAM and NON-SPAM emails using only the subject line. After they have a set of rules that they believe works well, the teams are provided another set of 50 SPAM and 50 NON-SPAM email subject lines to test and revise their set of rules. Based on their revised set of rules, the teams of students are asked to come up with a numerical measure to quantify how well their method (model) works. Finally, each team writes a report describing how their model works and how well it performed on the test data.

By considering a real-world problem, a set of data, and the need for a solution, students experience the statistical enquiry cycle (see Wild and Pfannkuch, 1998). We believe (and have seen in our pilot work) that both MEAs described above engage students in the types of thinking that Wild and Pfannkuch (1998) describe as "fundamental to statistical thinking" – recognition of the need for data; transnumeration; variation, and the use of models. Both MEAs require the use of statistical concepts that are unspecified by the problem and yet unknown to the students. The iPod Shuffle MEA provides students with experience that sets the stage for exploring ideas such as randomness, stating and refining hypotheses, and testing hypotheses. The SPAM MEA sets the stage for the consideration of statistical ideas such as fitting a model (the set of rules), judging the fit of a model with new data, modifying the model, and types of error in prediction. Furthermore, the students must come to solutions through making sense of the data provided and from their experiences with the problem context. As Wild and Pfannkuch remind us, "One cannot indulge in statistical thinking without some context knowledge. The arid, context-free landscape on which so many examples used in statistics teaching are built ensures that large numbers of students never even see, let alone engage in, statistical thinking" (p. 228).

PLANS FOR BUILDING AND ASSESSING STATISTICAL THINKING DURING UNITS OF INSTRUCTION

Each of the MEAS described above are used to introduce a unit of instruction in an introductory course in statistics primarily offered for social science students. The iPod shuffle MEA launches a unit on *Chance Models and Simulation* that includes ideas of informal inference. The SPAM MEA begins a unit on methods of *data classification* that includes ideas of hits and misses (mis-classification), and types of errors. The initial MEA in each unit is followed by four to six more class sessions filled with activities and discussions, aided by the use of technology tools, to build the ideas of that unit. By the end of the unit, students will also have experienced one possible expert solution to the problem given in the MEA and also solved a parallel problem using a modeling tool.

Assessment is an important part of teaching and learning statistics. 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). Our assessments, which are under development, will include an activity similar to an MEA that allows students to transfer what they have learned in a unit and to apply and illustrate their statistical thinking. This will serve as an important assessment for each unit as well as a way to measure the development in their statistical thinking. We are currently examining the use of these types of assessment for use in both individual and group settings.

CONCLUSION

There is room for much improvement in the teaching and learning of introductory statistics at the tertiary level. There is a need for new materials, instructional approaches, and assessments to help study and evaluate the effectiveness of different curricula in helping students develop valued learning goals such as the ability to think statistically.

The CATALST project is exploring the use of MEAs to develop and reveal students' statistical thinking as part of a sequence of units that build on each MEA and lead students from informal ideas of modeling and chance to the use of powerful statistical methods. A detailed evaluation and research plan will be implemented during the upcoming two years of the project to determine the impact of this approach in a variety of institutions and courses. In the meantime, we are sharing materials at our CATALST website (http://www.tc.umn.edu/~delma001/CATALST/) and encourage interested statistics educators to try an MEA in their class, as well as one or more of the follow-up activities and assessments, and to share with use what they learned.

REFERENCES

American Statistical Association. (2005). GAISE college report. http://www.amstat.org/education/gaise/GAISECollege.htm.

Chance, B. L. (2002). Components of statistical thinking and implications for instruction and assessment. *Journal of Statistics Education, 10*(3). Online: www.amstat.org/publications/jse/v10n3/chance.html

Cobb, G. (1992). Teaching statistics. In L. Steen (Ed.), *Heeding the call for change, MAA Notes No. 22* (pp. 3-34). Washington, DC: Mathematical Association of American.

delMas, R. C., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal, 6*(2), 28-58.

Diefes-Dux, H. A., Imbrie, P. K., & Moore, T. J. (2005). *First-year engineering themed seminar - A mechanism for conveying the interdisciplinary nature of engineering.* Paper presented at the 2005 American Society for Engineering Education National Conference, Portland, OR.

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

Lesh, R., Hoover, M., Hole, B., Kelly, A., & Post, T. (2000). Principles for developing thought-revealing activities for students and teachers. In A. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 591-646). Mahwah, NJ: Lawrence Erlbaum.

Lesh, R. & Lamon, S. (Eds.) (1992). *Assessments of Authentic Performance in School Mathematics*. Washington, DC: American Association for the Advancement of Sciences Press.

Moore, D. S. (1998). Statistics among the liberal arts. *Journal of the American Statistical Association, 93*, 1253-1259.

Moore, T. J., Diefes-Dux, H. A., & Imbrie, P. K. (2006). *The quality of solutions to open-ended problem solving activities and its relation to first-year student team effectiveness.* Paper presented at the American Society for Engineering Education Annual Conference, Chicago, IL.

Moore, T. J., Diefes-Dux, H. A., & Imbrie, P. K. (2007). How team effectiveness impacts the quality of solutions to open-ended problems. Distributed journal proceedings from the International Conference on Research in Engineering Education, published in the October 2007 special issue of the *Journal of Engineering Education*, *96*(4).

Rumsey, D. J. (2002), Statistical Literacy as a Goal for Introductory Statistics Courses, *Journal of Statistics Education*, *10*(3). Online: www.amstat.org/publications/jse/v10n3/rumsey.html.

Snee, R. (1999). Discussion: Development and use of statistical thinking: A new era. *International Statistical Review*, *67*(3), 255-258.

Wiggins, G. P. (1998). *Designing assessments to inform and improve student performance.* San Francisco, California. Jossey-Bass.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67,* 223-265.

Zawojewski, J., Bowman, K., & Diefes-Dux, H. A. (Eds.). (2008). *Mathematical modeling in engineering education: Designing experiences for all students*. Rotterdam, the Netherlands: Sense Publishers.