MAKING SENSE OF STATISTICAL STUDIES: A CAPSTONE EXPERIENCE FOR SECONDARY STUDENTS

Roxy Peck[1] and Daren Starnes[2]
[1]California Polytechnic State University, United States of America
[2]The Lawrenceville School, United States of America
rpeck@calpoly.edu

Statistics and data analysis have become more visible in secondary school mathematics in the United States. In most cases, statistics and data analysis topics have been spread through the mathematics curriculum. However, many important statistical concepts are not mathematical in nature and are not easily integrated into existing mathematics courses. As a result, most students complete their secondary education having encountered some graphical and numerical data analysis techniques, but not having engaged in meaningful statistical reasoning. Because the addition of a separate statistics course to the secondary curriculum is unlikely, an alternate approach is proposed. A capstone experience for secondary students has been designed. Making Sense of Statistical Studies is a coordinated set of investigations that can be used in existing mathematics courses to promote students' conceptual understanding of the data analysis process.

INTRODUCTION

In the last decade, statistics and data analysis have become a more visible component of the secondary school mathematics curriculum in many countries around the world. In the United States, the document *Principles and Standards for School Mathematics* (2000) includes data analysis and probability as one of five main content strands. This document has prompted revision of state and local school mathematics standards, and many of them now include methods and concepts from statistics. Other countries have also integrated data analysis. For example, Holmes (2000) discusses the role of data handling in the national curriculum in England. Ottaviani and Rigatti (2005) describe the data and predictions component in the mathematics curriculum in Italy. Mathematics and statistics together have been designated as essential learning areas in the Revised New Zealand Curriculum (2007). Curriculum efforts that incorporate statistics are underway in many other countries as well.

Although statistics and data analysis now appears in the "official" curriculum standards in many of the U.S. states, there is still a gap between the stated curriculum, the taught curriculum, and the curriculum that is assessed. Statistics and data analysis have not been incorporated in a meaningful way in most state and national assessments, and so even in states where statistics and data topics appear in the stated curriculum, they are often not part of the taught curriculum. Moreover, where statistics and data analysis do appear in state curricula, it is generally in a scattered and piecemeal way.

In 2009, in an unprecedented act, the governors and chief school officers of 48 U.S. states agreed to pursue a "common core" curriculum in mathematics. Even though the *Principles and Standards for School Mathematics* (2000) were meant to serve as a national model for K-12 mathematics, there was no political clout behind them and individual states usually wrote their own mathematics standards. The development of the Common Core is interesting in that it has the political backing that could well lead to widespread implementation. While the extent to which data analysis and statistics will be incorporated into the Common Core standards still remains to be seen, at the time of writing this paper there is some reason for optimism. Probability and statistics are two of ten themes in the public draft of the college and workforce readiness standards that were released for public comment in late 2009. The very preliminary drafts of K-12 grade-by-grade standards do incorporate learning objectives from data analysis and statistics, although the extent to which they are incorporated continues to be controversial. The hope of the statistics education community in the U.S. is that if statistics and data analysis are incorporated in a meaningful way in national mathematics standards that have political support and around which assessments will be designed, these topics could become part of a widespread taught curriculum in both primary and secondary schools.

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

However, even if this happens, it is still likely that statistics and data analysis topics will be integrated into the secondary mathematics curriculum by being divided up and spread through various mathematics courses. While the integration of data analysis into mathematics courses has advantages, including providing motivation and generating student interest in the associated mathematics content and showing the connections between statistics and areas of mathematics, there are also disadvantages. Many important concepts of statistics and data analysis are not mathematical in nature and are not easily integrated into existing mathematics courses. Examples include the concept of sampling variability, good data collection practices in sampling and experimental design, an understanding of the role that the method of data collection plays in determining the scope of conclusions that can be drawn, the distinction between association and causation, and the reasoning of statistical inference.

The challenge will be to avoid what is currently happening in many states and school districts where statistics and data analysis are currently part of the taught curriculum. Beginning in early grades, students are introduced to graphical and numerical data summaries, but they tend to collect data in ad hoc ways. In later grades, more sophisticated summaries are introduced, and students may see least squares lines as a way of summarizing bivariate data, but again little attention is paid to data collection and the complex issues associated with sampling variability are often overlooked. The typical end result is a student who knows how to use some data tools, but who does not have an understanding of data analysis as a process that begins with the formulation of a research question, followed by the development of a thoughtful plan for data collection, followed by data analysis, and finally interpretation and communication of results. Students who are not exposed to data collection and concepts associated with sampling variability do not have the opportunity to construct a framework that allows for the development of mature statistical reasoning.

An example of an attempt to ensure that some of the important (but non-mathematical) statistical concepts are included in the secondary mathematics curriculum is provided in the *College Board Standards for College Success* (2006). But these recommendations suggest including surveys and random sampling in Algebra I and the inclusion of design of experiments, surveys and observational studies in Algebra II. Even though a stated goal of these standards is to "describe the teaching and learning of mathematics and statistics as an integrated collection of processes and content elements", content from statistics and data analysis appear as "add on" topics. The statistics topics have been placed in the Algebra I and II courses only because they need to go somewhere if statistical thinking and an understanding of the data analysis process is to be developed.

To date, recommendations such as those in the *GAISE Report* and the *College Board Standards for College Success* have had no visible impact on the taught curriculum of Algebra I and Algebra II courses in the United States. Important statistical concepts appear as additional topics and are easily overlooked or intentionally omitted by teachers unprepared to teach these topics. Most textbooks for Algebra I and II do not support the inclusion of topics such as sampling and design of experiments—and why should they? It is not a natural fit. Unfortunately, even successful development of the Common Core K-12 standards is not likely to fix this problem.

MAKING SENSE OF STATISTICAL STUDIES

The *GAISE Report*, the *College Board Standards for College Success* and potentially the Common Core standards set the stage for changes in the taught curriculum. But is successful (from a statistics education point of view) large-scale integration of important statistics content feasible? Like the blindfolded participant in a children's game who has been spun around until dizzy, teachers will understandably be hesitant to take that first step. To move forward, we need to remove the blindfold, point them in the right directions, and give them a gentle nudge. In an attempt to do this, a joint committee of the American Statistical Association (ASA) and the National Council of Teachers of Mathematics (NCTM), with support from ASA, proposed the development of a set of classroom investigations with accompanying instructional materials for students and with supporting materials for teachers. The result was *Making Sense of Statistical Studies* (Peck & Starnes, 2009).

Making Sense of Statistical Studies (MSSS) consists of an introduction and four distinct sections. Each section begins with an overview that contains essential background information for students. The remainder of the section is devoted to guided student investigations. These investigations start with a research question on some topic of interest. Students are then led through a series of questions that help them examine the study design, analyze data, and interpret results. Later investigations ask students to design, carry out, and analyze results from their own studies. A description of each section follows.

The *Introduction* presents the components of the statistical problem-solving process. It includes a discussion of the primary methods of data production—surveys, experiments, and observational studies—as well as the difference between a sample and a population. Ethical issues involved in data collection are also mentioned here. *Investigations*: Did You Wash Your Hands?

Section I: Observational Studies shows students that much can be learned from observational studies. The first two investigations in this section help students review the primary graphical and numerical tools for analyzing data. The remaining investigations incorporate random selection, which allows students to generalize the results of their data analysis to some larger population of interest. *Investigations*: Get Your Hot Dogs Here!; What's in a Name?; If the Shoe Fits…; Buckle Up; It's Golden (and It's Not Silence).

Section II: Surveys begins with two investigations that require students to examine data from and critique the design of surveys that have already been conducted. In the final investigation of this section, students are led through the process of administering their own survey. *Investigations*: Welcome to Oostburg!; Student Participation in Sports; Planning and Conducting a Survey.

Section III: Experiments starts with an investigation in which students practice using the terminology of experiments as they review the details of two studies involving dieting and weight loss. In the next investigation, students are guided through the process of designing an experiment to test the effect of listening to music on memory. Once they have collected the data, students must use the data analysis and interpretation skills they developed in Section I to help answer the research question. Students get to design, execute, and analyze results from their own experiments in the final investigation of this section. *Investigations*: Do Diets Work?; Distracted Learning; Would You Drink Blue Soda?

Section IV: Drawing Conclusions introduces students to the basic ideas of inference—estimating a population characteristic and testing a claim about a population characteristic. Simulation is used to quantify the sample-to-sample variability that occurs in repeated random sampling. Students explore how this chance variation is reflected in the margin of error for an estimate and in the decision-making process for evaluating the validity of a claim about some population characteristic. *Investigations*: The Internet—Information or Social Highway?; Evaluating the MySpace Claim; Are Teens the Same Everywhere?

The investigations and instructional materials were reviewed by both classroom teachers and university faculty. Most of the investigations were piloted with teachers in summer institutes in 2008. Feedback from teachers and reviewers informed the authors as they revised the materials for final publication.

By completing the investigations in MSSS, students see the complete data analysis process in three contexts (observational studies, surveys and experiments), and they are exposed to the ideas of sampling variability and the role that sampling variability plays in statistical inference. MSSS is designed as a stand-alone experience with the methods of designing statistical studies and analyzing the data from these studies. It is written for a secondary school audience having some background in basic probability and graphical and numerical summaries. It can be completed in its entirety as a capstone experience that provides the big picture view of data analysis as a process that is missing in earlier courses that focus primarily on graphical and numerical descriptive methods. A teacher following the recommendations in the *College Board Standards for College Success* could cover the Introduction and Sections I and II as a unit in Algebra I and Sections III and IV as a unit in Algebra II.

To help teachers who may not be familiar with some of the topics introduced in MSSS, a teacher's module has been produced which includes supporting resources. Teacher notes for each section and for each of the investigations within a section include the following: an overview that

discusses the big ideas presented in the section or investigation; a list of prerequisite knowledge and skills that students will need to complete the section or investigation successfully; a list of learning objectives that describe what students should know and be able to do as a result of completing the section or investigation; detailed teaching tips that include discussion of key terms and concepts, examples, and specific questions from the student investigations; suggestions regarding timing of the investigations; possible extensions for the teacher who might want to go further with his or her students; and suggested answers to the questions in each of the fifteen investigations.

USES AND IMPACT OF *MAKING SENSE OF STATISTICAL STUDIES*

In the year since its introduction, *Making Sense of Statistical Studies* has been used effectively in many ways and with a variety of audiences. Some examples include:

- For Grade 9 students enrolled in a Math 2 course (geometry and mathematical models), MSSS served as an introduction to designing and analyzing data from statistical studies. Students completed 10 of the first 12 investigations as part of a two-week unit. The module's focus on the statistical problem solving process closely paralleled students' experiences in their required science course, which emphasized study design and basic methods of univariate data analysis.
- Grade 11 and 12 students enrolled in a yearlong introductory statistics class tackled 12 of the 15 investigations at appropriate points during the course. The teacher used five of the first six investigations to motivate a unit on designing studies. Three investigations relating to surveys and two examining experiments were used to supplement the treatment of these topics in the core textbook for the course. Finally, students completed two investigations from Section IV: Drawing Conclusions as an introduction to inferential reasoning.
- Investigation #4: If the Shoe Fits…served as the basis for an American Statistical Association webinar highlighting connections between the *Guidelines for Assessment and Instruction in Statistics Education* and the MSSS module. Teachers from Grades 7 through 12 participated live in this online event; many more have subsequently viewed the webinar.
- As part of an in-service program for secondary school statistics teachers, several MSSS investigations were used to highlight important principles of study design, implementation, relevant techniques of data analysis, and the appropriate scope of inference. After doing the investigations in teams, teachers had an opportunity to clarify underlying statistical issues and to discuss effective implementation strategies for MSSS in their own courses.

CONCLUSION

The hope for *Making Sense of Statistical Studies* is that the instructional materials and investigations provided will help to point teachers who are being asked to incorporate statistics and data analysis in the right direction and that it will serve as a gentle nudge in making that difficult first step. Assisted by good instructional materials and accompanying supporting materials for teachers, this first step could be the beginning of a move from a piecemeal approach to a more coherent and big picture view of data analysis in the secondary mathematics curriculum. This will be especially important if the Common Core Standards in Mathematics affirm the place of data analysis and statistics in the mathematics curriculum at the national level in the United States.

REFERENCES

College Board (2006). *College Board Standards for College Success: Mathematics and Statistics*. Online: http://professionals.collegeboard.com/k-12/standards

Franklin, C., Kader, G., Mewborn, D.S., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre K-12 Curriculum Framework*. Alexandria, VA: American Statistical Association.

Holmes, P. (2000). *Statistics across the English national curriculum*. Online: www.rssce.org.uk

National Council of Teachers of Mathematics (2000). Principles and Standards for School Mathematics. Reston, VA.

Ministry of Education (2007). Revised New Zealand School Curriculum. Online: www.minedu.govt.nz

Ottaviani, M.G., & Rigatti, S. (2005). Data and predictions emerging as one of the basic themes in the mathematical curriculum of the first cycle in Italy. In Burrill, G., & Camden, M. (Eds.), *Curricular Development in Statistics Education: International Association for Statistical Education 2004 Roundtable.* Voorburg, the Netherlands: International Statistical Institute and International Association for Statistical Education. Online: www.stat.auckland.ac.nz/iase/publications

Peck, R., & Starnes, D. (2009). *Making Sense of Statistical Studies*. Alexandria, VA: American Statistical Association.