PREPARING FUTURE TEACHERS TO TEACH STATISTICS

Gail Burrill
Michigan State University, United States
burrill@msu.edu

As statistics and data analysis become increasingly important so does the need to prepare teachers to teach these topics. The paper addresses three aspects related to this work with prospective teachers. First, a theoretical framework related to fundamental statistical ideas in the school curriculum (Burrill & Biehler, 2011) might provide guidance for the preparation of teachers across the grade levels. Second is a description of a course for prospective elementary teachers at a large university, which is designed to introduce key statistical concepts using real data and relevant contexts. Third is a brief discussion of the preparation of prospective mathematics teachers at the secondary level and how the statistics courses they take prepare them, or not, to teach statistical content and reasoning processes, particularly the content suggested in the U.S. Common Core State Standards, which have a strong emphasis on informal inference at the upper secondary level.

The need to prepare students to take part in a world awash in data is changing what was once the "basics" – reading, writing and arithmetic. While these basics are still important, the workforce today is looking for people who also have the ability to examine data from an assembly line to determine if the process is in control, derive and interpret statistical models for monitoring cash flow or to design methods to collect and analyze data about efficient modes of transportation. Countries need citizens who can make reasoned decisions based on data about issues such as health choices and the environment. As a consequence, the role of statistics as a component of the school curriculum is gaining traction across the world. For example, in the US, the Common Core State Standards in Mathematics (CCSSM) (2010) call for all students to learn about the role of data in the statistical process (including how the data were collected), randomization to draw conclusions and compare treatments, and probability models to describe random processes. In New Zealand, the curriculum learning area is called mathematics and statistics (Ministry of Education, 2007). The Australian curriculum for grades 1 to 10 is organized around three content strands, one of which is statistics and probability (Australian Curriculum Assessment and Reporting Authority, 2012).

Given that statistical content is beginning to be embedded in the curriculum, the question then becomes, who is teaching this content. Unfortunately, "Few current teacher training programmes adequately educate teachers for the task of teaching statistics." (Batanero, Burrill, & Reading, 2011, Executive Summary). This became evident in the 18th ICMI Study, *Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education,* a joint project with IASE, which calls for connecting teacher education to teaching practice and to the reality of classrooms, promoting teachers' statistical literacy and reasoning and engaging them with real data and statistical investigations. In the context of this recommendation, the paper suggests a framework for fundamental statistical ideas that can guide the engagement of preservice teachers in doing and thinking about statistics.

A THEORETICAL FRAMEWORK FOR PREPARING PROSPECTIVE TEACHERS OF STATISTICS IN THE SCHOOL CONTEXT

The content of teacher training programs in statistics should focus on core statistical concepts and ways of thinking. Based on a variety of perspectives on teaching statistics including the recommendations in the Guidelines for Assessment and Instruction in Statistics (Aliaga et al, 2005; Franklin et al, 2005); an interpretation of statistical literacy for adults (Gal, 2001), and the processes involved in developing statistical thinking (Wild & Pfannkuch, 1999), the following might be considered fundamental statistical ideas (Burrill & Biehler, 2011) and as such, critical components in developing preservice teachers' content knowledge for teaching statistics.

1. Data – including types of data, ways of collecting data, measurement; respecting that data are numbers with a context;

2. Variation – identifying and measuring variability to predict, explain or control. The term "variability" is used for the general phenomenon of change and "variation" for describing the total effect of the change;

3. Distribution – including notions of tendencies and spread that are foundational for reasoning about statistical variables from empirical distributions, random variables from theoretical distributions; and summaries in sampling distributions;

4. Representation – graphical or other representations that reveal stories in the data including the notion of transnumeration;

5. Association and modelling relations between two variables – nature of the relationships among statistical variables for categorical and numerical data including regression for modelling statistical associations;

6. Probability models for data generating processes – modelling hypothetical structural relationships generated from theory, simulations or large data set approximations, quantifying the variability in data including long term stability;

7. Sampling and inference – the relation between samples and the population and the essence of deciding what to believe from how data are collected to drawing conclusions with some degree of certainty.

These ideas, however, invoke not only statistical but also mathematical constructs, often leading to a tension between the two views. Future teachers should recognize for example, that work with data goes beyond manipulating and operating with numbers to context-related reading between and beyond the data (Friel, Curcio, & Bright, 2001). Mathematics and statistics differ in representing data: 1) most statisticians begin with a graph; in mathematics, the emphasis is on numerical operations without paying attention to a visual representation of data; and 2) graphs in mathematics are often used in showing the same relationship in different representations (tables, graphs, symbols). In statistics, different graphs or representations are used to identify different aspects of the same data. While mathematics can be thought of as being exact and precise, statistics is about "noise", i.e., how to measure and control variability. Real data in statistics are contextual, containing uncertainty, and error while data in school mathematics classrooms are typically assumed to perfectly fit a mathematical model. Data collection plays no systematic role in going from a real situation to the mathematical model in most mathematics classrooms nor does comparing mathematical results to empirical data. In such models a statistical lens is typically absent, for example checking residuals or connecting the context and choice of a model.

Teachers should understand that a mathematical approach to proportional reasoning can undermine the statistical approach for reasoning from samples. Percentages in mathematics are often applied in simple contexts, where the reference is set and the units are clear and constant. Careful statistical statements made about margin of error and confidence intervals are replaced by simplistic "inferences" from "sample" to "population", assuming a perfect proportional relationship. Ignoring uncertainty and variability, sample results are reported in point estimates rather than interval estimates in many media reports. Preparing students for statistical thinking requires that discussions in mathematics classrooms make this difference explicit. In mathematics, deciding what to believe is straightforward: conclusions follow deductively from definitions and agreed on principles. In statistics, reasoning is partly inductive, and conclusions always uncertain. The degree of faith in a statistical conclusion depends on the integrity of the entire investigative process, while in mathematics a proof makes you certain. In statistics, how the data were collected and the role of randomness determines how you can interpret the results; in (pure) mathematics, the reasoning is independent of the data.

These differences between mathematical and statistical ways of thinking can be difficult for a teacher, responsible for both subjects, to navigate. And the focus of most teacher preparation programs is on the mathematics. Recognizing the need for a program to develop preservice teachers' knowledge of statistics, in 2007 Michigan State University, through a large grant focused on reviewing the university's teacher preparation programs, engaged two statisticians and a mathematics educator in designing and piloting a statistics course for preservice K-8 teachers. This led to the incorporation of the course into the curriculum as an elective for those preservice

teachers who choose a mathematics focus in their preparation program. (All elementary preservice teachers have to select an academic area of emphasis to complete their degree.)

THE COURSE

Future teachers are encouraged to take the course to prepare for teaching the CCSSM in their own classes, understand how to represent and analyze data that is essential for teaching science and social studies often part of an elementary teacher's responsibility in the U.S., and, in the current context of accountability in the US, to be able to interpret statistical measures and graphs to assess the instruction in their schools and to communicate with parents. The goals of the course have evolved over time but in general are compatible with those recommended in GAISE (Franklin et al, 2005):

- Emphasize statistical literacy and develop statistical thinking;
- Use real data;
- Stress conceptual understanding rather than mere knowledge of procedures;
- Foster active learning in the classroom;
- Use technology for developing conceptual understanding and analyzing data;
- Use assessments to improve and evaluate student learning.

In particular, the course objectives for students are to:

- develop a comfort level with statistics and probability that will allow them to be effective in the classroom and competent in handling the statistical issues that they encounter in their professional and personal lives;
- develop an understanding of statistical thinking;
- use technology as an aid for understanding data, probability and statistics and consider the role of technology for use in classroom instruction.

The content is organized into three sections: Data Collection, Data Analysis, and Probability. A typical schedule might look like that in Table 1 with readings from the text and assignments detailed for each week.

Table 1: Statistics and Probability for Prospective K-8 Teachers

Topic	Week	Topic	Fundamental Idea
Collecting Data	1	Sampling bias; Observational studies, surveys and samples	Sampling
	2	Sampling Variability; Margin of error	Variation
	3	Bias in surveys and experiments	Data/Sampling
	4	Exploring measurement and random error	Sampling
Data Analysis	5	Graphing categorical and quantitative data	Representation
	6	Review of Data Collection and Test 1	
	7	Distributions: Shape, center and spread	Distribution
	8	Two quantitative variables: Scatterplots and association	Association
	9	Linear regression	Association
Probability	10	Chance, Randomness, and Probability	Probability models
	11	Test 2; Law of large numbers, expected value	Probability models
	12	Simulations	Probability models
	13	Probability Models including Normal Distribution	Probability models
	14,15	Student presentations	

Class time is used for active investigation of course topics, in line with the GAISE (Aliga et al, 2005) recommendations, to help students understand the material both for teaching and using

statistical thinking in their professional life. Investigations involve problem contexts such as "Is swimming with dolphins therapeutic? What is the average word length in Abraham Lincoln's Gettysburg Address? How can we find a margin of error for a voter preference poll, critically reading media articles (Are kids who own dogs really more active?), investigating the number of students who skip class and the factors involved in their choice to attend or not to attend, and analyzing the achievement scores of fourth and eighth grade students on the National Assessment of Educational Progress. The textbook, currently *Statistics: Concepts and Controversies, Eighth Edition* (Moore & Notz, 2014), is used as a resource outside of class to solidify the ideas investigated in classroom activities. Homework consists of weekly assignments, two reports on activities related to computer investigations, and a project in which students design, implement, and analyze data from an observational study, experiment, or survey.

Technology investigations involve creating different representations of data to reveal different aspects of the story contained in the data. Technology is used to develop understanding of concepts such as box plots (Figure 1) or correlation (Figure 2) and to engage students in exploring variability in sampling distributions (Figures 3 and 4) through simulations.
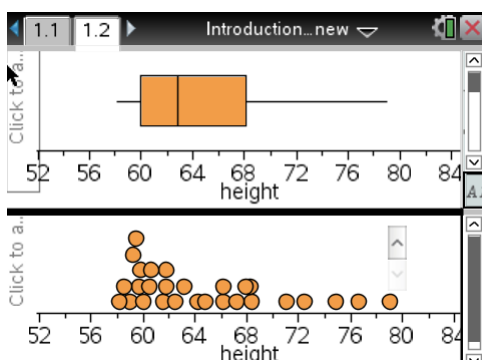

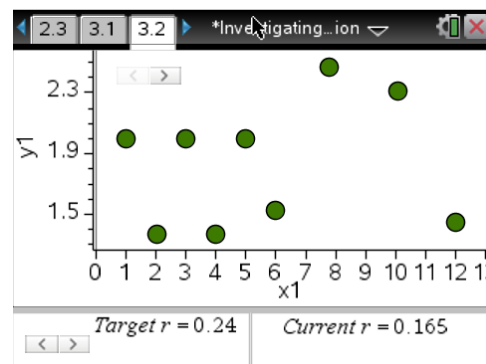Figure 5 Investigating box plots

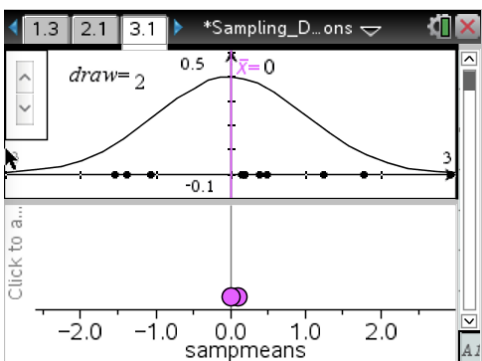Figure 6 Investigating correlation
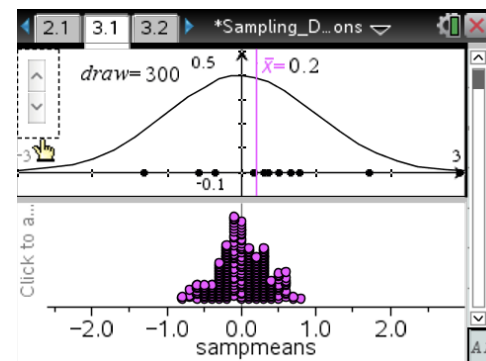

Figure 7 Distribution of sample elements

Figure 8 Simulated Distribution of Sample Means

The only measure of success of the course has been continued interest with an enrollment of about 25 students per year. No funding was available for classroom follow up.

PREPARATION OF SECONDARY TEACHERS TO TEACH STATISTICS

Prospective secondary teachers may or may not have to take a statistics course as part of their degree program, depending on university requirements. If they are not required to take such a course, they have only their own learning of statistics from informal experiences or, in rare cases, their own high school courses on which to draw. Additionally, if they are required to take such a course, the course might be mathematical in nature, in essence defining statistical concepts in mathematical terms, rather than beginning from data based experiences. Typically in such courses, prospective teachers learn procedures and formulas without context. Such an approach leaves them ill prepared to teach statistics as proscribed in the CCSSM or to address their students' difficulties.

As part of a year-long six credit methods course, prospective secondary teachers were asked to design a class discussion around a given task much as they would expect to do with their

students. Students worked in small groups on one of the seven tasks chosen to represent a range of statistics content and aligned with the fundamental ideas described above: a data analysis problem involving the amount of money in change students in class carried with them, a data analysis task involving the comparison of graduation rates by state, a question of difference involving data presented in a two-way table, a bivariate data task comparing cholesterol levels before and after a treatment, several probability problems approached through simulation, an informal approach to a confidence interval, a central limit theorem task and a task related to interpreting correlation. The students were seniors who would graduate as mathematics majors, and most of them had taken a mathematically based statistics course.

In the first task, the prospective teachers were asked to draw a graph of the amount of change in cents they had with them and describe the distribution. The graphs below are typical of those produced by students over five different cohorts of students.
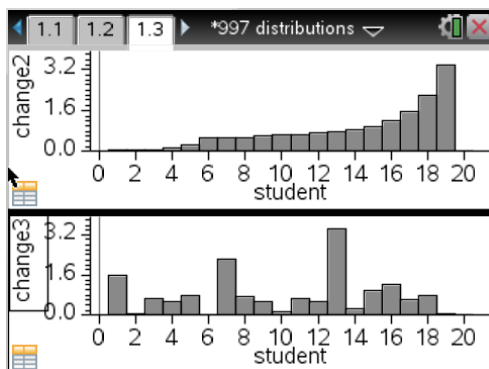

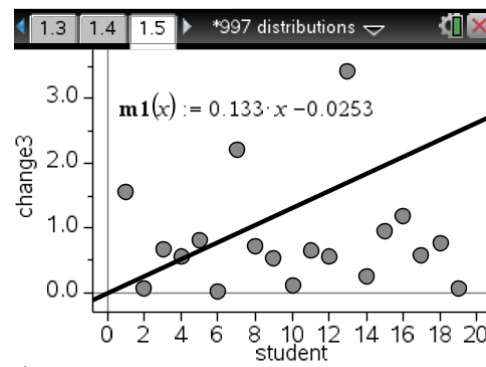Figure 9 Univariate displays of change (cents)   Figure 10 Plot of change vs student

The students had little understanding of making sense of data because they had not worked with any but were adept at finding means and standard deviations. They could find the probability of three or more boys in a family of 5 children but had little sense of variability and had never seen a simulation. They could not identify whether an observed outcome might have been due to chance or explain what a margin of error was, and had no ways to even think about what the concepts might mean. They knew about the Central Limit Theorem but were not sure when it applied.

CONCLUSION

Thus, we face a two-fold dilemma. Prospective K-8 teachers are rarely given an opportunity in their preparation to learn about statistical reasoning and making informed decisions based on understanding statistical reasoning. As well, secondary teachers may not be given such an opportunity, or if given the opportunity, it may be mathematical in nature. The U.S. CCSS describe a fairly ambitious set of statistics standards that all students are expected to learn. Teachers have to be prepared to teach to those standards, which means that preservice programs that do not prepare students for this teaching are not adequately preparing students for the work of teaching. The second edition of a document commissioned by the Conference Board of the Mathematical Sciences related to the preparation of teachers of mathematics (Conference Board of the Mathematical Sciences, 2012) contains strong recommendations for preparing future teachers of statistics beginning with a call for statistics departments to become involved and ensure that statistics courses for teachers are centered around statistical concepts and real-world case studies, and make use of technology in an active learning environment. The document suggests that the standard statistics courses for future engineers or science majors in many institutions might not be appropriate and advocates for content that will engage future teachers in developing statistical thinking that matures over time and help them understand ways in which understanding of statistical content and processes evolve. The authors argue for one semester of statistics, preferably beyond an introductory statistics course for middle school teachers and at two semesters for prospective high school teachers.

Moving forward and even implementing these recommendations, however, should not be done in isolation from the statistics education research community. Many open questions exist, and

investigating these can enhance our understanding of how to better prepare teachers. For example, in what ways can we more effectively use technology to develop understanding of statistical concepts? What depth of content knowledge do teachers themselves need to prepare all students to be "consumers" and some to be "producers" of statistics (Gal, 2002)? What strategies effectively engage statisticians in preparing teachers for the work of teaching K-12 statistics? Should statistics at the high school level be a "course" in its own right, or should statistical content be integrated into the mathematical content sequence? And what are the consequences of each pathway for the preparation of teachers?

Acknowledging the increasing role in the intended curricula for students in elementary and secondary schools presents a challenge for those in statistics education but also an opportunity. Thoughtful and careful study can help us make the transition, informed by research, one that has lasting impact for the students in classrooms throughout the world.

ACKNOWLEDGEMENTS

REFERENCES

Aliaga, M., Cuff, C., Garfield, J., Lock, R., Utts, J., & Witmer, J. (2005). *Guidelines for Assessment and Instruction in Statistics Education (GAISE)*. Alexandria VA: American Statistical Association. http://www.amstat.org/education/gaise/GaiseCollege_Full.pdf

Australian Curriculum and Assessment Reporting Authority (2012). *The Australian Curriculum: Mathematics*. Sydney: ACARA. www.australiancurriculum.edu.au

Batanero, C., Burrill, G., & Reading, C. (2011). *Teaching statistics in school mathematics-Challenges for teaching and teacher education.* ICMI Study volume 14. New York: Springer.

Burrill, G., & Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics - Challenges for teaching and teacher education*: A joint ICMI/IASE Study (pp. 57-69). New York: Springer.

Common Core State Standards Mathematics. (2010). Council of Chief State School Officers & National Governors Association. www.corestandards.org/

Conference Board of the Mathematical Sciences (2012). *Mathematics Education of Teachers II*. Washington DC: Conference Board of Mathematical Sciences & Mathematical Association of America.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education, 32*(2), 124-158.

Franklin, C., Kader, G., Mewborn, D. S., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-K-12 curriculum framework*. Alexandria, VA: American Statistical Association. Online: amstat.org/education/gaise/.

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review, 70*, 1-51.

Ministry of Education. (2007). New Zealand Curriculum, Learning areas, mathematics and statistics. http://nzcurriculum.tki.org.nz/The-New-Zealand-Curriculum/Learning-areas/Mathematics-and-statistics, 2007

Moore, D., & Notz, W. (2013). *Statistics: Concepts and controversies, eighth edition.* W.H. Freeman & Co.

New Zealand Ministry of Education (2006). *The New Zealand Curriculum: Mathematics and statistics*. Online: nzcurriculum.tki.org.nz/

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