SHAPING K-12 STATISTICS EDUCATION IN THE UNITED STATES

Roxy Peck[1], Gary Kader[2] and Christine Franklin[3]
[1]California Polytechnic State University, [2]Appalachian State University, [3]University of Georgia
United States of America
rpeck@calpoly.edu

In recent years, three key documents have been influential in focusing attention on statistics and data analysis in the Pre-K-12 mathematics curriculum in the United States. We examine how these three documents come together with a collective potential to shape the future direction of Pre-K-12 statistics education, and we describe the specific contributions made by the document Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework.

INTRODUCTION

Increasingly, statistics is recognized as a fundamental component in the development of an educated citizenry (Utts, 2003). Consequently, over the past few decades, statistics has become an important part of school curricula, primarily as a component of the mathematics curriculum. Holmes (2000) discusses the role of statistics (Data Handling) as one of four major strands within the study of school mathematics in the National Curriculum for England (DEE & QCA, 2000). He describes the multi-disciplinary aspects of statistics and the potential to reinforce statistical thinking in other subjects apart from mathematics. Ottaviani and Rigatti (2005) describe efforts by the Italian Statistical Society in promoting the Data and Predictions component in the mathematics curriculum in Italy. In Australia, Chance and Data are included among seven areas with mathematics learning outcomes in the development of numeracy for learning across the curriculum (Curriculum Council, 1998). Mathematics and statistics together have been designated as essential learning areas in the Revised New Zealand School Curriculum (2007). Certainly, similar efforts are underway in many countries promoting the importance of statistics in the education of its citizens.

In the last decade, statistics and data analysis have become a more visible component of the Pre-K-12 mathematics curriculum in the United States. Three key documents have been influential in focusing attention on these topics:

1. *Principles and Standards for School Mathematics* (National Council of Teachers of Mathematics (NCTM, 2000). This document includes Data Analysis and Probability as one of its five content strands and has become the basis for reform of the mathematics curriculum in many states.
2. *The Mathematical Education of Teachers* (Conference Board of the Mathematical Sciences, 2001), known as the MET Report. This document focuses on teacher development in mathematics and makes specific recommendations for essential teacher knowledge in the areas of statistics and data.
3. *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework* (Franklin et al., 2007). This document addresses student learning objectives in statistics and data analysis and is consistent with both the NCTM standards and the MET Report recommendations for teacher development.

This paper will describe the motivation behind the GAISE report and indicate how it supports implementation of the NCTM Standards and the MET Report recommendations. Recommendations regarding teacher preparation for implementing the GAISE guidelines will also be provided.

BACKGROUND

Statistics is a relatively new subject for many Pre-K-12 teachers, who have not had the opportunity to develop sound knowledge of the principles and concepts underlying the practices of data analysis that they now are called upon to teach. Publication of the NCTM

standards for data analysis and probability and the MET Report recommendations for teacher development directed attention to this important area, but these two documents did not provide teachers or teacher educators with a coherent picture of a unified statistics curriculum based on a developmental sequence of learning experiences.

Statistical thinking is different from mathematical thinking in important ways, and statistical thinking must develop and mature over time in a way similar to mathematical thinking. To facilitate this, teachers need to recognize the ways in which understanding of statistical concepts and processes evolve. The GAISE Report was written in order to provide teachers and teacher educators with a developmental framework for Pre-K-12 statistics education. The GAISE framework introduces teachers to the statistical problem solving process and then shows how this process can be presented at each of three developmental levels (Levels A, B, and C). Although the basic structure of the process is the same at each level, the degree of sophistication in the types of problems addressed increases over the developmental levels.

THE GAISE FRAMEWORK

The GAISE framework illustrates the statistical problem solving process with an additional emphasis on understanding the role of variability in this process. Statistical problem solving is an investigative process that involves four components: (1) question formulation, (2) data collection design and implementation, (3) data analysis, and (4) interpretation. This process is similar to the handling data cycle described by Holmes (2000) and the four-stage PCAI model for statistical investigation proposed by Graham (2006). Understanding the role of variability in this process requires maturation in statistical thinking, and the GAISE framework describes this maturation over three developmental levels. Although these levels may parallel grade bands, they are based on development in statistical thinking, not age. For example, a middle school student who has had no prior experience with statistics will need to begin with Level A concepts and activities before being able to successfully move on to Level B. Learning is more teacher driven at Level A but becomes student driven at Levels B and C.

At Level A, questions posed are typically restricted to the classroom with data collected by taking a census of the classroom. The data are analyzed using simple picture graphs, tallies, frequency tables and bar graphs, or dotplots (lineplots). The mode is introduced as a summary of categorical data for the category having the highest frequency. Both the mean (developed as the fair share value) and median are introduced at Level A as summaries of center for numerical data. The range is introduced as a basic measure of spread. Interpretation is focused on comparing individual to individual variability and individual to group variability within the context of the question posed. Generalizations beyond the classroom are not expected.

Transitioning to Level B, the questions posed are broader and typically extend beyond the classroom. The concept of random selection is introduced. Using multiplicative and proportional reasoning, Level B students are able to transition to pictographs, circle graphs (pie charts) and relative frequency tables and bar graphs for summarizing categorical data, and histograms and boxplots are developed for summarizing numerical data. The interpretation of the mean is expanded to that of the *balance point* of the data distribution, and the notion of variation in data from the mean is introduced using the mean absolute deviation (MAD) as a transitional measure to the standard deviation developed at Level C. Interpretation is focused on both comparing within group variability and between group variability. Students also make use of comparative relative frequency tables and bar graphs, conditional percentages, and segmented bar graphs for categorical data, and comparative dotplots and boxplots for numerical data. Through the use of boxplots, students are introduced to the interquartile range (IQR) as a measure of spread. At Level B, students begin to understand that the ability to generalize conclusions is tied to the way in which the data are collected.

Transitioning to Level C, students can now ask questions that require generalization from a sample to a larger group. Random selection is used when collecting data. Levels A and B focus on interpreting variability through descriptive statistics. At Level C, through the use of simulation, students begin to think about sampling variability, the role of probability in statistical problem solving, and their impact on conclusions and generalization.

The GAISE developmental framework presents a conceptual structure for statistics education using a two-dimensional model. One dimension is defined by the components of the statistical problem-solving process along with a component corresponding to an understanding of the nature and role of variability. The second dimension is comprised of the three developmental levels. Table 1 summarizes key aspects of this two-dimensional framework.

Table 1. Problem Solving Across Maturation Levels

Process Component	Level A	Level B	Level C
Formulate Question	Questions restricted to classroom	Questions not restricted to classroom	Questions seek generalization
Collect Data	Census of classroom	Non-Random Sample Surveys Begin to discuss random selection	Sample designs using random selection
Analyze Data	Display variability within a group Compare individual to individual Compare individual to group	Quantify variability within a group Compare group to group (between) variability in displays Acknowledge sampling error Some quantification of association	Measure variability within a group Measure variability between groups Compare group to group using displays and measures of variability Describe and quantify sampling error based on a simulated sampling distribution with margin of error and p-values. Understand sampling variability. Quantification of association
Interpret Results	Do not look beyond the data No generalization beyond the classroom	Acknowledge that looking beyond the data is feasible Acknowledge that a sample may or may not be representative of larger population Basic interpretation of models for association	Look beyond the data in some contexts Generalize from sample to population Interpret models for association
Nature of Variability	Measurement variability Natural variability Induced variability	Sampling variability	Chance variability
Focus on Variability	Variability within a group	Variability within a group and variability between groups	Variability in model fitting

RELATING THE THREE KEY DOCUMENTS

A primary goal of the GAISE Report is to provide a conceptual and developmental framework for Pre-K-12 statistics education. The GAISE guidelines were designed to complement and clarify the student learning goals outlined in the NCTM *Principles and Standards for School Mathematics* (2000). However, these documents both focus on student

outcomes, and they will not have maximum impact on the Pre-K-12 taught curriculum until issues of teacher preparation are addressed. In fact, the MET Report notes the following: "statistics is the science of data, and the daily display of data by the media notwithstanding, most elementary teachers have little or no experience in this vitally important field" and "of all the mathematical topics now appearing in middle grades curricula, teachers are least prepared to teach statistics and probability." (p. 87)

The MET Report goes on to make specific recommendations for teacher preparation that are well-aligned with the student focused NCTM Standards and the GAISE recommendations. MET recommendations for teacher preparation, aligned to GAISE developmental levels, are shown in Table 2.

Table 2. MET Teacher Experience Recommendations Aligned with GAISE Student Developmental Levels

Process Component	Level A	Level B	Level C
Formulate Question	Experience in understanding the kinds of questions that can be addressed by data.		
Collect Data	Experience in creating data sets.	Experience in designing simple investigations and collecting data (through random sampling or random assignment to treatments)	Experience in designing surveys to estimate population characteristics. Experience in designing experiments to test conjectured relationships among variables.
Analyze Data	Experience in understanding shape, spread, center. Experience in using different forms of representation. Experience in comparing two sets of data.	Experience in using a variety of ways to display data. Experience in exploring and interpreting data by observing patterns and departures from patterns in data displays, particularly patterns related to variability.	Experience in exploring data, using a variety of standard techniques for organizing and displaying data in order to detect patterns and departures from patterns.
Interpret Results	Experience in choosing among representations and summary statistics to communicate conclusions. Experience in understanding variability. Experience in understanding some of he difficulties that arise in sampling and inference.	Experience in drawing conclusions with associated measures of uncertainty.	Experience in using probability models to draw conclusions from data and to measuring the uncertainty of those conclusions.

PREPARING TEACHERS TO TEACH STATISTICS

In the United States the preparation of mathematics teachers has historically been the shared responsibility of mathematicians and mathematics educators. Until recently, statisticians have not generally had a major role in teacher preparation. As noted by Franklin and Mewborn (2006), "the explicit inclusion of statisticians in the enterprise of teacher education is a relatively recent phenomenon. Given the increasingly prominent role of data analysis in both curricula and daily life, it is essential that statisticians, mathematicians, and mathematics educators collaborate on the design and implementation of teacher education programs for both pre-service and in-service mathematics teachers in pre-K-12…" (p. 335). The GAISE

Framework was written, in part, to provide guidance to all involved with teacher preparation in clarifying the curricula in data analysis required for all levels, Pre-K-12.

Although the three GAISE developmental levels require different experiences with statistical content, there are many common elements in teacher preparation for all three levels. First and foremost, preparation must provide teachers at each level with the statistical content, concepts, and practices that will enable them to adapt both statistical and pedagogical ideas to a classroom setting. Teacher preparation should stress problem solving and provide instructional activities that have the spirit of genuine statistical investigation. Ideas should evolve in contexts that encourage the development of the statistical reasoning skills appropriate for that level.

Preparation should provide teachers with opportunities to learn statistics in the same ways that we hope they will practice teaching statistics in their own classrooms, and teachers should learn using the same types of activities that their students will experience. Two types of activities employed should be: (1) activities based on statistical problem-solving, and (2) activities based on statistical concept development.

Problem solving activities are based on the statistical process described earlier. These types of activities are important because they emphasize the complete cycle in the statistical process, and teachers should have experiences utilizing the entire process. Teachers must develop an understanding of the types of questions that can be addressed with data at each level. They must have experiences identifying appropriate variables, designing a plan to collect data, and collecting data for a study. Teachers need practice forming a data driven argument that addresses the question at hand. Problem solving activities should provide teachers with experiences using multiple representations of data (tabular, graphical, and numerical summaries), and that will help teachers to understand the connections between these various representations. Finally, problem-solving activities should provide examples of the proper interpretation of results appropriate for their level.

Concept development activities develop statistical concepts underlying the problem solving process. In many presentations of statistical concepts, definitions just appear (often as formulas), without adequate explanation or justification. Teachers need to understand that statistical techniques are inventions. They begin with a need to describe or understand a statistical concept and evolve from initial ideas. Concept development activities should emphasize the evolution of statistical ideas and provide the foundation for understanding statistical definitions, formulas and techniques.

The GAISE Report provides several examples of both problem solving activities and concept activities at each of the three developmental levels.

GAISEING INTO THE FUTURE

The three documents described in this paper differ somewhat in focus but support one another in a way that makes their potential collective impact far greater than the sum of their individual potentials. Through its developmental framework and by providing specific examples that show maturation of statistical thinking over the developmental levels, the GAISE Report is helping teachers and teacher educators to view Pre-K-12 statistics education as a developmental process by providing a framework that shows how statistics and data analysis can be viewed as a coherent curriculum strand. This is a critical step in the move to successfully implement NCTM standards in Pre-K-12 mathematics curricula and to achieve MET recommendations for teacher development.

REFERENCES

Conference Board of the Mathematical Sciences (2001). *The mathematical education of teachers.* Washington, DC: American Mathematical Society and Mathematical Association of America.

Curriculum Council (1998). *Curriculum framework for kindergarten to year 12 education in Western Australia.* Online: www.curriculum.wa.edu.au.

Department of Education and Employment and Qualifications and Curriculum Authority (2000). *The National Curriculum of England.* Online: www.nc.uk.net.

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.

Franklin, C., & Mewborn, D. (2006). *The statistical education of grades pre-K-12 teachers: a shared responsibility.* In G. Burrill (Ed.), *Thinking and reasoning with data and chance: 68th NCTM Yearbook.* (pp. 335-344)*.* Reston, VA: NCTM.

Graham, A. (2006). *Developing thinking in statistics*. London: Paul Chapman.

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

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.

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 G. Burrill, & M. Camden (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.

Utts, J. (2003). What educated citizens should know about statistics and probability. *The American Statistician, 57*(2), 74-79.