Using Statistics to Improve Education: A Dilemma

Gail Burrill
Division of Science and Mathematics Education, Michigan State University, East Lansing, Michigan, USA
burrill@msu.edu

Abstract

The ultimate goal of teaching statistics is to foster an adult population capable of reasoning from and about data and making informed decisions based on quantitative information. This paper describes several examples of the difficulty educators have in applying these skills in the context of their work as educators. The dilemma is how to prepare those involved in the design and delivery of education to understand reasoning with data to interpret and use information about their schools, teachers, and students to improve what they are doing as an educational system.

The Problem

Citizens are being called upon to make increasingly complex decisions about policies and practices in the socio-political, workplace, and consumer arenas. (Franklin & Garfield, 2006; Garfield, 2002; Kader & Perry, 2006). Crucial skills to make informed decisions include the ability to explain, decide, judge, evaluate, and analyze information (Rumsey, 2002). Growing evidence suggests that students who learn statistics this way find it difficult to apply statistical concepts in real settings where the concepts are clearly applicable, and their use could help prevent costly errors in decision-making (Garfield & Ben Zvi, in preparation). And many involved in the educational system have never had a course in statistics or quantitative literacy that would give them the background and tools to make sense of the data about student achievement and understanding with which they are confronted. Even those who have taken a course often do not leave with any lasting understanding of statistical concepts (Clark et al, 2003). They have not grasped that measures of center without corresponding measures of spread do not give an accurate picture of the situation, the necessity of interpreting the variability in graphs of data, the need to have a "standard" in order to measure both attainment and improvement.

The discussion in this paper draws on professional development work with teachers in several settings, including a large Mathematics Science Partnership project funded by the National Science Foundation, Promoting Rigorous Outcomes in Mathematics and Science Education (PROM/SE). The project, involving over 60 school districts and 3,000 teachers at five different sites in Ohio and Michigan and directed by William Schmidt and Joan Ferrini-Mundy, collected baseline data from all of the teachers and students in the project schools. This paper describes some of the issues, activities and results of efforts to enable teachers and administrators to understand data about their students, teachers and schools in addition to issues raised in a research project on the use of graphing calculators in beginning algebra.

Do you use a graphing calculator on homework?

		N	Mean	Std. Deviation
Percent correct on posttest	no	106	43.4486	17.84579
	yes	134	52.6580	18.37777
	Total	240	48.5905	18.67794
Percent correct on Posttest Procedure w/o Connection items	no	106	57.8167	25.27840
	yes	135	66.1376	24.78913
	Total	241	62.4778	25.29402
Percent correct on Posttest Procedure with Connection items	no	106	44.2208	19.93464
	yes	135	56.2449	19.87902
	Total	241	50.9563	20.74290
Percent correct on Posttest Doing Math Items	no	106	27.4259	17.85066
	yes	134	31.7697	20.15205
	Total	240	29.8512	19.25240

Table 1: Average percent correct on Posttest Items by Cognitive Demand of Items

Understanding the Relationship between Measures of Center and Spread

"The median home price is $179,000." "The car averages 32 miles per gallon." The structure of the curriculum, the media, and common usage all contribute to using measures of center without regard for how the data might vary around those measures (Shaughnessy, 1997). The standard deviation is not easily understood and often misunderstood (Garfield & Ben-Zvi, in preparation). Table 1 shows the results of a posttest test given to 13-14 year olds in a project studying the effect of graphing calculators on students' understanding of introductory algebra concepts (Breaux & Burrill, 2007). The items were clustered according to cognitive demand in the categories used by Stein and colleagues in their work (Stein et al, 2001). It is important to view the scores on the posttest for in terms of both mean and standard deviation, which shows a relatively large spread around the mean.

A box plot reveals more information about the data; showing the range for each category and how the data are clustered within that range (Figure 1). From this display it is possible to see that while the median percentage correct for the items without connections is about the same for those using and not using a calculator, the middle half of the group of students using a calculator scored higher, from about 50 percent to 82 percent while the no calculator group went from 38 percent to 71 percent; some students did not score well in either group; the top 25 percent of students using a graphing calculator scored about 11 percent higher than the no calculator use group. One inference might be that more needs to be done to reach low achieving students.

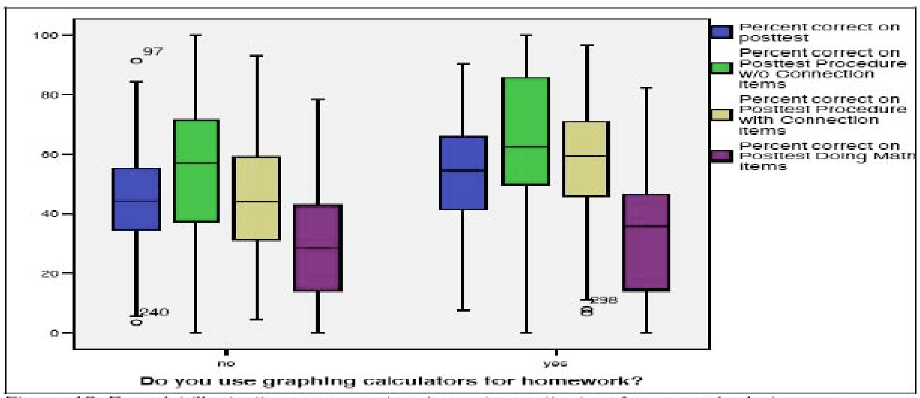


Figure 12: Box plot illustrating mean and variance in posttest performance by hgt use on homework

Figure1: Average Scores on Algebra Items by Level of Difficulty

Interpreting the results presented in box plots, however, was difficult for teachers as they have had little experience reasoning with variability.

Interpreting Box Plots

Unfortunately, our experience overall suggests that this is a common problem: some educators have trouble interpreting box plots, encountering many of the same difficulties that students have in understanding how to read them: linking length of segments or boxes with frequency of data, misunderstanding the lack of relationship between plot area and frequency, unable to reason from general characteristics of a set of data, difficulty perceiving how the data might be clustered within a quartile, problems comparing across groups (Bakker et al, 2006; Makar & Confrey, 2005; Bakker & Gravemeijer, 2004).

Box plots are a primary tool used to represent PROM/SE data (see Figure 2).

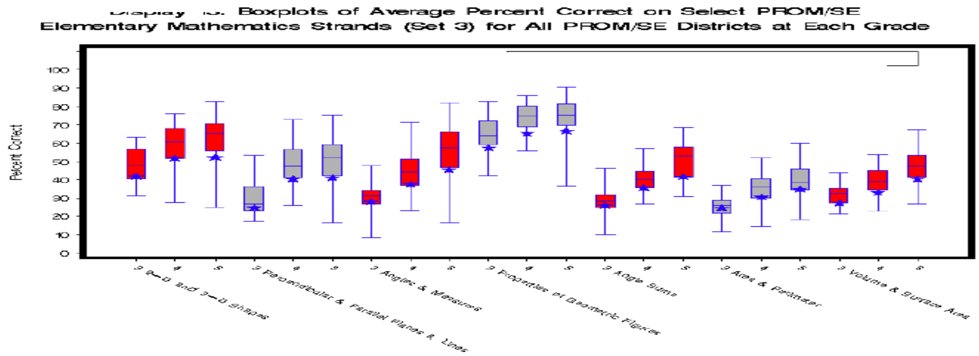
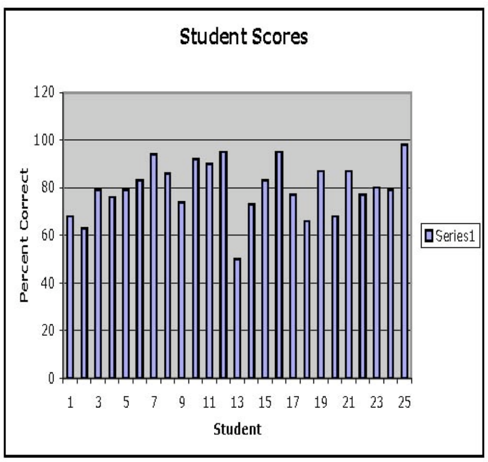


Figure 2: Box plots of Student Achievement on Geometric Concepts by Grades

To overcome the problems educators have in reading these plots, we designed a series of activities to help teachers, many of whom taught box plots to their students, think about how they might interpret box plots in the context of reports on student achievement. Participants considered bar graphs showing individual scores in a class, converted the data to dot plot on the number line and then imposed a box plot over the dot plot (Figures 3 and 4).



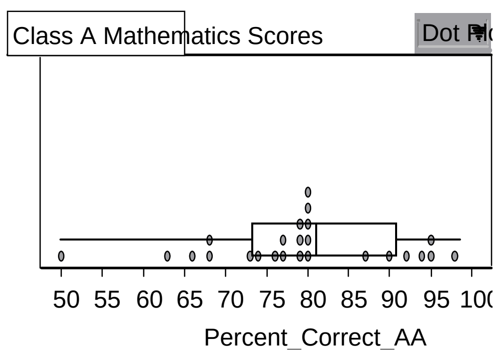
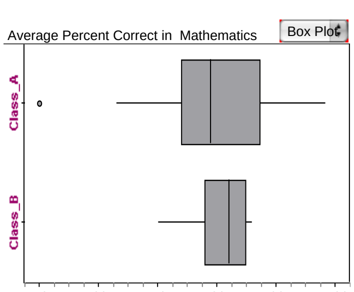


Figure 3: Bar Graphs of Student Scores Figure 4: Dot /Box Plot of Student Scores

They answered a series of questions about different box plots and discussed how the ideas related to interpreting scores and comparing across groups (Figures 5 and 6).



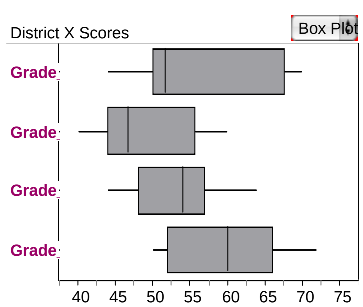


Figure 5: Comparing Classes Figure 6: Comparing Grades

After engaging in a variety of other activities focused on work with box plots, the participant evaluations at the end of the session seemed to indicate they had enough understanding to begin working with their own school data (see Figure 7 for example).

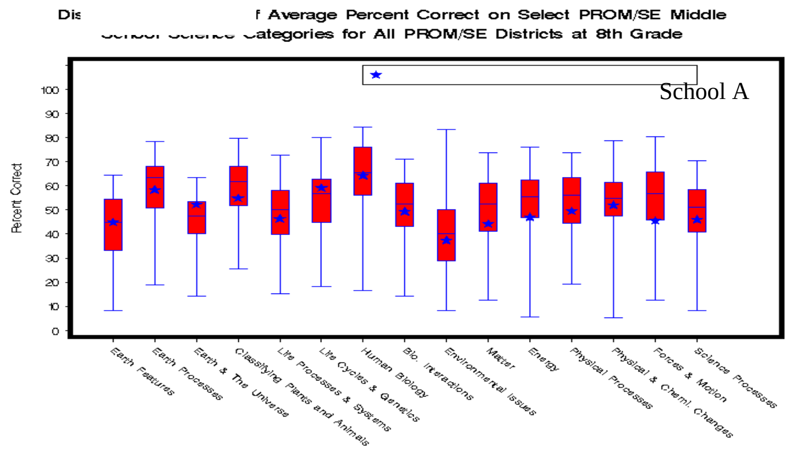


Figure 7: Box Plots of Science Achievement

This turned out to be an erroneous assumption, however, and we had to rethink how to approach the problem. We designed a second set of activities explicitly focused on helping teachers recognize the need for scale in a box plot; interpret an individual case displayed in a box plot; recognize that box plots show both magnitude and order; compare features across box plots, and see that comparing groups is different than comparing individuals. One of these activities was designed to help participants internalize key elements of box plots: the combination of order and magnitude; visible clusters and their relation to the whole distribution, and what you can learn from a box plot and what you cannot.

A second activity used the number of years of participants' teaching experience to illustrate the role variability plays in moving from a set of individual values, summarizing these values by finding the mean, repeating the process for many groups to build a distribution of means and exploring how these distributions are alike and how they are different. Box plots were used to represent the distributions (Figures 8 and 9). The discussion explored how these ideas related to data on their own schools and districts.

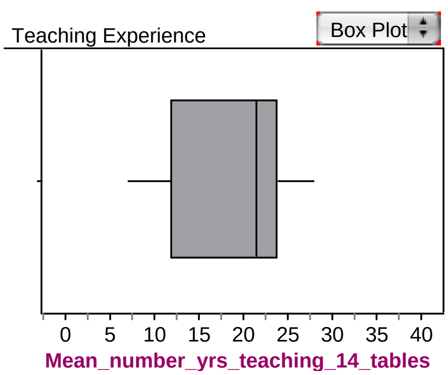


Figure 8: Mean years teaching by table

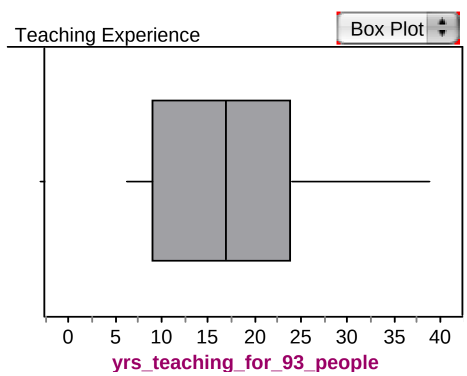


Figure 9: Individual years teaching

Although the participants were engaged in these and other related activities and could respond as a group, once again they were unable to use them to productively examine their own data. Many were unable to come to grips with the variability, relate medians and quartiles to a set of means, adapt to the shift in population from a distribution of schools in a district to a distribution of all of the districts to a distribution of all of the schools; and unable to make connections across box plots. They found it difficult to examine how student performance in a grade varied over a set of content expectations and how the same plot could be used to show how students in consecutive grades performed. The teachers had problems seeing relationships within and across grades.

The final session retreated to offering a table of mean achievement scores, and the discussion focused on what you could learn just from analyzing these data, without any regard to the variability. The next steps for the project are to provide each district with an individual analysis including a discussion of the observations about patterns and trends made by the project staff and to offer to work with district personnel on further interpretations.

The Dilemma

In most US curricula, making and interpreting box plots occurs in the middle grades, for students ages 12 to 14. The ideas are revisited in some curricula in high school, but the level of cognitive complexity is rarely deepened, and the applications are typically procedural. At the tertiary level, students intending to be teachers do not always have to take a statistics course; if they do, the course is often a mathematical statistics course; and as reported above, even with courses that stress reasoning about and with data, the evidence suggests that student learning is still ill formed. The dilemma is that as more data about teaching and learning becomes more readily available and the tools for analyzing the data become more sophisticated, the ability to produce useful information from the analyses is outpacing the capacity of the field to use the knowledge productively. This poses serious questions: what degree of statistical literacy should we expect for every one? How can we structure opportunities to enable those in our educational systems to reach this level and to ensure that a significant cadre in any school system is able to use statistics to improve their system? Some research has shown that carefully designed activities over a longer period of time have been successful in generating understanding and the ability to use statistical ideas such as box plots in a variety of situations. The challenge that faces the statistical education community is how to take advantage of this research and find ways to apply it so that we can increase the capacity of the field to improve educational practices by reasoning and thinking with data.

References

Bakker, A., Biehler, R., & Konald, C. (2004). Should young students learn about boxplots? In G. Burrill and M. Camden (Eds.) *Curricular development in statistics education: IASE 2004 Roundtable on Curricular Issues in Statistics Education.* Voorburg, the Netherlands: International Statistics Institute.

Bakker, A., & Gravemeijer, K. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.) *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 147-167). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Breaux, G. & Burrill, G. (2007). The relationship between the cognitive demand of mathematical tasks and assessment performance in first year algebra students. Paper given at AERA in Chicago IL.

Clark, J., Karuat, G., Mathews, D., & Wimbish, J. (2003). The fundamental theorem of statistics: Classifying student understanding of Basic statistical concepts. Unpublished paper.

Franklin, C. & Garfield, J. (2006). The GAISE project: Developing guidelines for K-12 and college courses. In G. Burrill (Ed.) *Thinking and reasoning with data and chance*: National Council of Teachers of Mathematics 2006 Yearbook. Reston VA: The Council

Garfield, J. & Ben-Zvi, D.(in preparation). Developing students' statistical reasoning: Connecting research and teaching practice.

Garfield J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education*, 10(3). http://www.amstat.org/publications/jse/v10n3/garfield.html

Kader, G. & Perry, M. (2006). A framework for teaching statistics within the K-12 curriculum., In A. Rossman & B. Chance (Eds.). Proceedings of ICOTS 7. Voorborg, the Netherlands: International Association for Statistical Education http://www.stat.auckland.ac.nz/~iase/publications/17/2B3_KADE.pdf

Makar, K. & Confrey, J. (2005). Using distributions as statistical evidence in well-structured and ill-structured problems. In K. Makar (Ed.) *Reasoning about distribution: A collection of current research studies. Proceedings of the Fourth International Resea4rch Forum on Statistical Reasoning, Thinking and Literacy* (SRTL-4), University of Auckland, New Zealand, 2-7 July. Brisbane: University of Queensland.

PROM/SE. Promoting Rigorous Outcomes in Mathematics and Science Education Michigan State University Mathematics Science Partnership. W. Schmidt & J. Ferrini-Mundy Principle Investigators. Supported by NSF Cooperative Agreement EHR – 0314866

Rumsey, Deborah J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education*. Vol 10, No. 3 (2002)

Shaughnessy, M. (1997). Missed opportunities in research on the teaching and learning of data and chance. In F. Biddulph & K. Carr (Eds.) *People in mathematics education*. (Proceedings of the 20th annual meetings of the Mathematics Education Research Group of Australasia) (pp. 6-22). Rotorua, New Zealand: MERGA.

Stein, M.K., Smith, M.S., Henningsen, M.A., and Silver, E.A., (2000). *Implementing Standards-Based Mathematics Instruction: A Casebook for Professional Development*, NCTM, Teachers College Press, New York.