THE ROLE OF FORMATIVE ASSESSMENT IN TEACHING AND LEARNING STATISTICS

BURRILL, Gail
Michigan State University
USA

Providing tasks that enable teachers to understand how students are processing concepts allows teachers to shape instruction, plan, adapt, and differentiate, depending on what students need to learn. What does this look like when teaching statistics? This paper presents background on formative assessment and describes a framework for thinking about how it can be enacted in practice. The framework is illustrated by focusing on the nature of statistical tasks that can elicit information about student thinking and on instructional strategies that deliberately provoke such information. The discussion is grounded in work with students in middle school, preservice mathematics education, and inservice elementary teachers, describing the challenges and dilemmas that arose and the strategies employed to overcome these challenges.

Teaching is about enabling learning - what teachers do should cause learning to take place. Researchers suggest that when students are actively involved in choosing and evaluating strategies, considering assumptions, and receiving feedback their performance improves (NRC, 1999, p.57). *Adding It Up* reports that "…when teachers learn to see and hear students' work during a lesson and to use that information to shape their instruction, instruction becomes clearer, more focused, and more effective." (NRC, 2001, p.350). In recent years, this process of probing into student work and thinking, collecting and analyzing information about what students have learned, has become part of a broader view of assessment, moving beyond measuring students' knowledge for the purposes of reporting and sorting. In formative assessment the teacher uses this evidence to shape teaching, adapting teaching to meet student needs based on information about student reasoning and understanding (Black & Wiliam, 1998). In their literature review, Black and Wiliam found practices that strengthened the use of formative assessment produced significant and often substantial learning gains. The studies they examined ranged over age groups from 5-year-olds to university undergraduates, across several school subjects, and over several countries. The question this paper addresses is what formative assessment looks like in teaching statistics.

Providing feedback and guiding conversations among students that move learning forward should be more than providing correct answers and verifying procedures. The knowledge within and across statistical content areas can be broken down using a variety of classifications. Two of them seem particularly useful to help teachers frame questions and provide feedback. The first is the categorization of knowledge types used by Shavelson and colleagues (Shavelson, Ruiz-Primo & ayala, 2003) from their research aimed at designing better assessment instruments for science learning. They categorized the knowledge students draw on in four distinct types:

Knowledge Type	*Use of Knowledge*
Declarative	Defining, providing an example
Procedural	Executing and performing procedures
Schematic	Explaining, justifying, predicting, hypothesizing
Strategic	Choosing knowledge to use, formulating strategies raising questions, defining problems

A second approach might be to look at the nature of what is to be learned: 1) definitions; 2) how concepts are represented and the language used in describing them; 3) rules, theorems and properties that follow; and 4) connections or applications. Taken together these two approaches suggest the framework in table 1 as a guide for the process of probing student thinking and using the results to inform the design of lessons.

Table I
Content Analysis Framework

	Declarative	Procedural	Schematic	Strategic
Definition				
Representation/language				
Theorems, rules (or assumptions), properties				
Connections to other statistical concepts				
Applications contexts/ used to ground other statistical concepts				

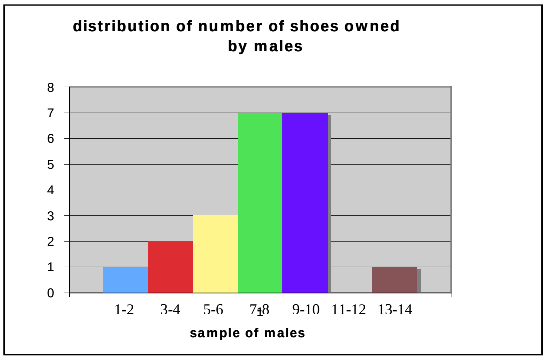
(Note that the use of strategic knowledge implies that students have command of the three lower-level knowledge types. For example, to choose which test of significance to use in a particular context, students have to know definitions of the different tests, understand the language, know and check the assumptions for each, understand how the tests are similar and how they differ, and recognize the need to apply the test for a given context. The following discussion gives several examples of how the framework might be used to promote formative assessment in two aspects of the teaching/learning process: choosing tasks and carrying out instruction during class.

TASKS

The heart of teaching is what students are asked to do – the tasks in which students engage as they explore concepts, learn procedures, and try to make sense of the content. Tasks need to be chosen carefully to build a coherent learning trajectory and justified in terms of the learning aims they serve. To serve the aim of formative assessment, tasks only work well if opportunities for pupils to communicate their evolving understanding are built into the planning (Black & Wiliam, 1998). For example, a discussion of "contrasting cases" can help students notice new features they missed and identify important ones that can shape their understanding of the major concepts (NRC, 1999).

Some tasks are designed to ensure that students master certain procedures or skills. Such tasks are very limited in what they reveal about students' thinking. Other tasks, much more open-ended, have the potential to reveal a great deal about students' conceptions and misconceptions. Such tasks have several of the following characteristics: multiple access points for students with different levels of understanding, models that can be used to develop concepts, visualization, multiple solution strategies, critical thinking, opportunities to make connections to other concepts, and progressive formalization. For example, a task for secondary preservice students who had covered some basic data analysis techniques was to specify a question around a variable related to something of interest to them, collect data that might shed some insight into an answer, graph the data, and write a short summary of what they found. The summary was to include a description of the question, the data collection method, and a reflection on what they would change, if anything, if they were to do the assignment over again. The task has multiple access levels, multiple solutions, requires critical thinking, and involves visualization. Some typical student responses are displayed in figures 1- 4.

Graphs such as those in figures 1 and 2 revealed students did not fully appreciate that graphs were about communicating information and that accurate scales and labels were a critical part of conveying the message. In terms of the framework, the knowledge they needed was schematic and the content stage was representing; the students were not able to represent their data in ways that would communicate and struggled with using a representation to explain the story. This suggested the teacher needed to revisit graphs as a tool for presenting a message – to help students think more deeply about what it means to represent a situation and what is required for the representation to be sense making for the data. In response, the instructor created a mini-lesson where several graphs were presented without labels or scales, and students were asked to tell a story from the graph putting in the labels and the scales to match their story. They shared their stories and in each case pondered whether the labels were appropriate and useful; subsequent work showed considerable improvement in this respect.



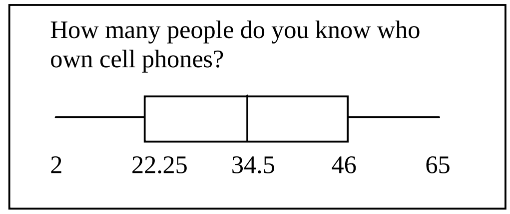
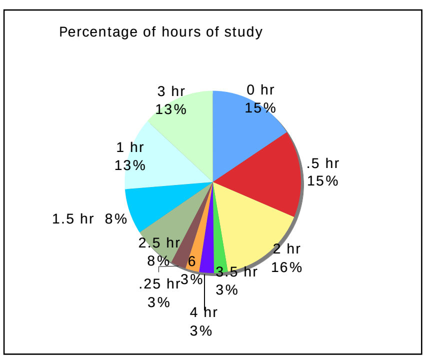


Figure 1: Student Work-Bar Graphs

Figure 2: Student Work- Box Plots

Figures 3 and 4 seem to indicate that students are struggling to distinguish categorical data from measurement data and to choose graphs that will make sense for the larger story in the data. (The tendency to preserve the identity of the source for the individual data point seems to be prevalent and occurs with both univariate and bivariate data, as a later example will show.)



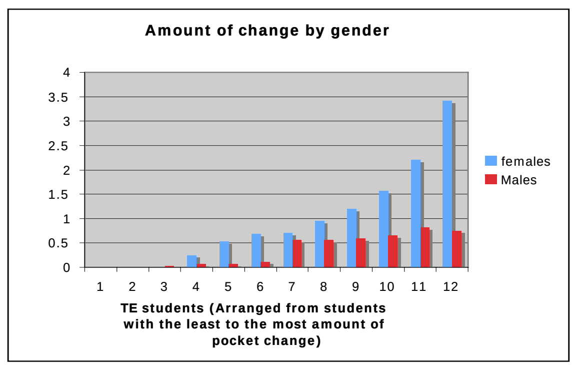


Figure 3: Student Work-Circle Graph

Figure 4: Student Work–Side-by-Side Bar Graph

These issues are again in the content domain of representation. While on the surface, it may seem that students are displaying lack of declarative knowledge or even procedural knowledge, the problem might actually lie in the domain of strategic knowledge, knowing enough to choose representations that are appropriate. Although this work was done with preservice college students, DiSessa and colleagues (DiSessa, Hammer, Sherin & Kolpakowski, 1991) made an observation about a sixth grade class that seems applicable: "One of the difficulties with conventional instruction…is that students' meta-knowledge is often not engaged, and so they come to know "how to graph" without understanding what graphs are for or why the conventions make sense…Particular representations may not be at the core of what we should teach so much as the uses they serve, criteria they meet, and resources they build on" (p.157). The preservice students had demonstrated in earlier work that they could make conventional graphs when instructed to do so and that they were fairly competent in interpreting them. Given the freedom to construct a graph of their own choosing, some struggled to find ones that made sense for their data. To address this, the instructor provided more opportunities for students to design their own graphs and took care to discuss whether a particular graph really conveyed important information or if the graph was in fact, accurate but did not reveal as much information as another representation might.

Lehrer and Romberg (1996) claim that textbook examples of graphs often are too preprocessed, which could be interpreted as saying that students are not given the opportunity to develop strategic knowledge. Asking students to create their own graphs can lead to insightful representations that are not typically taught. For example, middle school students were asked to analyze the growth of mung beans, a rapidly growing sprout, in four different solutions (salt, lemon-lime, and cola each mixed with water and tap water). The ultimate question was in which solution did the mung beans grow best. Each group was assigned a different solution and instructed to write a short summary of the results of the experiment, including a graph of the final lengths of the sprouts, a description of what the graph suggests about the growth of the sprouts over the 7 days, statements about the lengths using some measure of center and how the growth varied. Figure 5 shows a representation for the growth of the sprouts in a cola solution. The lemon-lime solution group chose to plot the average of the growth of their plants each day (Figure 6), and then during the discussion became convinced that their approach would help answer the larger question and added the other solutions (Figure 7).

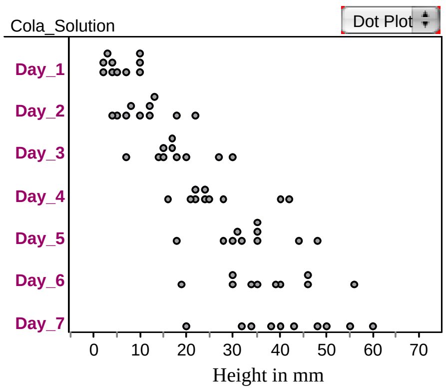


Figure 5: Student Work-Multiple Dot Plots

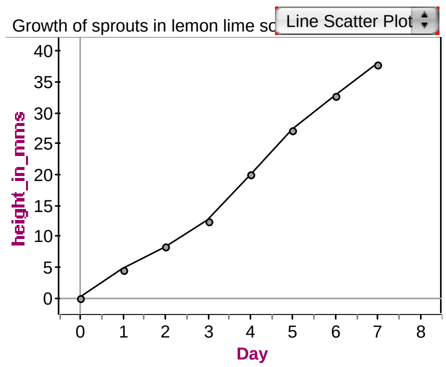


Figure 6: Student Work-Plot over Time

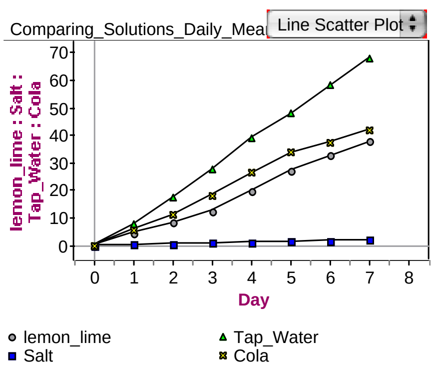


Figure 7: Student Work-Multiple Plots over Time

Sharing the work allowed others to see an application of rate of change and to look at a representation that clearly illustrated the contrast in growth rates among the solutions. Giving students the opportunity to share interesting and innovative solutions (for lower secondary students) is important in enabling students to have a vision of where the analysis can take them and to help them see the advantages and disadvantages of different representations.

INSTRUCTION

A teaching challenge is how to process discussion of the issues raised by student work in ways that honor students' efforts yet move the class ahead in a productive manner. In the best of all worlds, class norms have been established that celebrate errors and mistakes as opportunities for learning. Such norms apply to both the instructor and students, and discussions of different and incorrect solutions or strategies are part of the class routine. Another challenge is how to organize interventions that will help students move ahead in real time. One strategy is to plan ahead and build the opportunity to monitor student understanding into the lesson itself.

The Japanese have a vocabulary for describing parts of a mathematics lesson: while students work on the *shu hatsumon* (a key question), teachers engage in *kikan-shido* (walk among the desks). The central part of the lesson is *noriage* ("kneading" student solutions), where the teacher orchestrates the discussion, comparing student ideas and integrating solutions, to make the mathematics visible. The lesson ends with *matome* (summing up), where the teacher brings the ideas together (Bass, Usiskin & Burrill, 2002). This approach is rooted in formative assessment and can shape the classroom in statistics just as it does in mathematics. The work of the teacher is to use the work of the students to make the statistics visible.

For example, a newspaper headline stated "Women's income catching up to men's." Using data for the median incomes of men and women in the United States since 1970 students investigated whether the data supported the claim in the headlines. The Navigator™system, a wireless network that allows teachers and students to exchange data through graphing calculators and allows the teacher to capture students' calculator screens, provides a rich resource for tapping into student thinking. Student errors are immediately visible to the teacher who can decide how to address them (Is the error common for many students? Is the error related to understanding or a lost negative sign?). The teacher can also choose to display a set of student work for the entire class to consider. This task would be situated in the framework in strategic knowledge and application.

Figure 8 shows some of the approaches students took including a box plot of the differences, a plot of the differences over time, a plot of (men's income, women's income) for each year, a plot of (year, men's income) and (year, women's income), a plot of (year, difference). The teacher's task is to orchestrate the discussion of these strategies, getting at schematic knowledge: asking students to explain their plot and what they believe it shows. The teaching dilemma: Which should come first? Discussion of solution strategies should not be random; if students are to take something coherent from a discussion then the discussion needs to be organized around a line of reasoning that is recognizable and made explicit in the discussion. Some Japanese teachers tend to this by taking notes as they walk among the desks; the walk is purposeful, about how to manage a discussion of the work the teacher is observing, and the notes will help the teacher make choices. (Often in classrooms in the United States, the walk is to provide answers to individual or group questions or to maintain discipline).

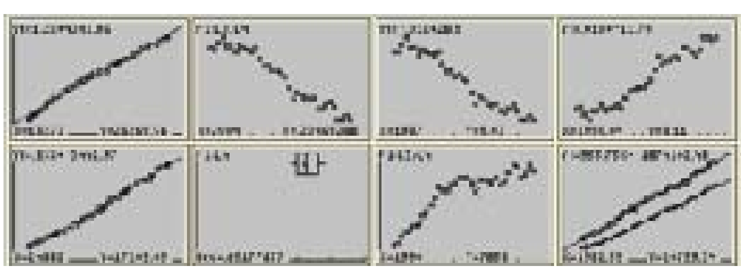


Figure 8: Comparing Income of Men and Women

In this case, are all of the graphs useful in interpreting the data? Should the box plot of the differences be considered first because it does not acknowledge the year and so will have little useful information for answering the question. Should a plot of (year, percent difference) follow the box plot discussion to show how the initial work with differences can be used to make a more meaningful plot? Or should the two graphs in the lower left corner be first, followed by the two graphs in the left column? The whole story seems to be contained in the graph (year, difference men-women) located in position (2,3), which shows the difference was increasing at a fairly constant rate until a given time (approximately 1984), when it leveled off and then rose again. Should this be last in the discussion? Students are actively involved in choosing and evaluating strategies, comparing and contrasting approaches, and receiving feedback; the teacher is questioning explanations, asking what if…., and why to probe for what students actually know and can do.

Another way to understand what students are thinking is to build in opportunities for predictions (schematic knowledge) into the lesson: “What do you think will happen?” “What will the graph look like if it supports your hypothesis?” “Predict the outcome before you calculate.” As students discuss their conjectures, the teacher can move among the groups listening to the discussions and taking notes. For example, in a lesson on experimental design, a class was conducting an experiment to determine which brand of chocolate chip melted faster in the mouth (Burrill, Franklin, Goldberg & Young, 2003). They decided to collect data in two different ways: 1) a matched pair design where each person tested both chips in a random order and 2) two randomly chosen groups, one to test the first chip and the second to test the second chip. After collecting the data, students were instructed to sketch a graph they might expect to see if there was a difference in the melting times for the two kinds of chips before they began analyzing the data. The situation illustrates the intersection of connections to other statistical concepts (analyzing data) and schematic knowledge. The results were instructive for the teacher; some in the matched pair design wanted to sketch side-by-side bar graphs for each individual, again preferring to retain the identity of the source for each set of data; a few choose to make box plots for each type of chip ignoring the fact that they had paired data; a very few initially choose to make a scatterplot of the paired data; some with data for each chip choose to make a plot over time of (trial number, melting time). Because the issues were connecting back to earlier work with analyzing data, the instructor sketched a series of different graphs on the board and worked with the class to consider what each might bring to understanding whether the evidence suggested that one brand of chips melted faster than the other.

A valuable strategy that can be built into lessons for finding out what seems to be working for students is to give them “minute papers”, a question or two they answer about what they understand and do not understand prior to leaving class (see, for example, Reading & Reid, 2005). The responses give the teacher the opportunity to design the next class period in response to issues raised in the papers. Consider a class that had covered sampling distributions for proportions and turned to sampling distributions for sample means. Responses on minute papers included comments such as “I do not know when to use the sample mean or the population mean in the equations” and “How does the standard deviation work?” It would seem their declarative knowledge is not very stable, and they are unsure of what the definition means and how to apply it. This prompted the instructor to begin the next class by revisiting the basic concept of variation in distributions using carefully drawn graphs of a variety of normal distributions, relating the formulas to the graphs, and providing students a structure for the concepts. Several responses such as “I understand better going from the normal distribution of the population to the sampling distribution and how they relate to each other; why you have to use z-scores” indicated to the instructor that some students grasped the key points, and the next class also included some deliberately grouped student-to-student conversations to help others move forward.

CONCLUSION

Formative assessment can be a powerful tool for enabling students to learn. Providing feedback – feedback that goes beyond right or wrong – can help students succeed. The feedback, however, to be most useful, should be provided prior to assigning grades (Black, Harrison, Lee et al, 2004). Choosing tasks that allow discussion enables the teacher to provide feedback; asking students to predict and anticipate outcomes, questioning what in the situation seems to support their conjectures allows teachers to hear what students are thinking. To be effective the feedback and interventions should focus on concepts central to understanding: definitions; representation/language; theorems, rules, properties; and applications where teachers look for the intersection of these with the type of knowledge students need in order to make sense of statistics.

This framework gives teachers guidance in designing interventions to address learning deficiencies and misconceptions they encounter as they gather information from their students about what they understand.

According to some Japanese educators, teachers must have “eyes to see the mathematics” and “eyes to see the students.” “The most important thing is for teachers to accept students’ ideas and to evaluate them flexibly. If a teacher couldn’t do so, the students might not come to life.” (Ikeda & Kuwahara, 2002) Formative assessment is about teachers being flexible, moving consistently towards statistical goals but using student thinking and obstacles to student understanding to shape their teaching. Such teaching is causing learning.

REFERENCES

Bass, H., Usiskin, Z, & Burrill, G. (2002). *Classroom practice as a medium for professional development.* Washington, DC: National Academy Press.

Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80,*. 139-148.

Black, P., Harrison, C., Lee, C., Marshall B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan, 86,* 8-21.

Burrill, G., Franklin, C., Godbold, L., & Young, L. (2003). *Navigating through data analysis in grades 9-12*. Reston VA: National Council of Teachers of Mathematics.

DiSessa, A., Hammer, D., Sherin, B., & Kolpakowski, T. (1991). Inventing graphing: Meta-representational expertise in children. *Journ*al *of Mathematical Behavior, 10*, 117-160.

Ikeda, T. & Kuwahara, Y. (2003). Presentation at Park City Mathematics Institute International Panel.

Lehrer, R., & Romberg, T. (1996). Exploring children’s data modeling. *Cognition and Instruction, 14*, 69-108.

National Research Council. (1999). *How people learn: Brain, mind, experience, and school*. Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). Washington, DC: National Academy Press.

National Research Council (2001). *Adding it up*. Kilpatrick, J., Swafford, J., and Findell, B. (Eds.) Washington DC: National Academy Press. Retrieved from: www.nap.edu.

Reading, C. & Reid, J. (2005). Consideration of variation: A model for curriculum development. In *Curricular Development in Statistics Education*. G. Burrill & M. Camden (Eds.). The Netherlands: International Statistical Institute.

Shavelson R., Ruiz-Primo, M.A., Li, M., and Ayala, C.C. (2003). *Evaluating new approaches to assessing learning.* (CSE Report 604, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Center for the Study of Evaluation (CSE), and the UCLA Graduate School of Education & Information Studies.