From **Gal, I. & Garfield, J. B. (editors).** **The Assessment Challenge in Statistics Education**. IOS Press, 1997 (on behalf of the ISI). Pages 179-190. ISBN 90 5199 333 1. Copyright holder: International Statistical Institute. Permission granted by ISI to the IASE to make the chapters from this book freely available on the Internet. The References section is a separate chapter and is available on the Internet at http://www.stat.auckland.ac.nz/~iase/publications/assessbkref.

14. Technologies for Assessing and Extending Statistical Learning

Susanne P. Lajoie

Purpose

This chapter does not provide an extensive overview of the use of technology in statistics. Rather, a detailed summary is provided of how one researcher has used technology to teach and assess statistics in grade eight. Both traditional modes of assessment as well as technology-driven methods will be described in an attempt to demonstrate how multiple mediums of assessment can be used to provide a profile of students' statistical knowledge. The research reported here is grounded in cognitive theory with an emphasis on theories of learning that emphasize learning situations that are concrete rather than abstract. In other words, students learn better by "doing" statistics rather than just computing or reciting statistical equations or definitions. This research program is designed to provide authentic learning and assessment situations (Lajoie, 1995). The term authentic refers to meaningful, realistic tasks and assessments that validly assess what the learner understands.

INTRODUCTION

One of the lessons of the reform movement in mathematics has been that instruction and assessment can be thought of as a dynamic (as opposed to static) process whereby new forms of instruction lead to new forms of assessment, and that the results of student assessment can inform teachers in making instructional decisions to improve student learning (National Council for Teachers of Mathematics (NCTM), 1989; 1995). This chapter will provide a profile of what assessment methods can be used to monitor student progress in statistical comprehension and skill. In this regard, examples with the use of technology are provided so that teachers can see ways in which assessment methods can be developed given their resources.

There are both benefits and obstacles to the use of technological resources for teaching and assessing statistics. The author is not suggesting that technology is a preferential form of instruction in this domain but rather, that it is a useful alternative. Computer software, computer simulations, multi-media technologies, computer-based learning environments, and the internet are becoming more affordable and are being used for educational purposes. The real question is whether or not technology can improve statistics education. This question is addressed by reviewing how technology can increase understanding and facilitate assessment. Obstacles to incorporating technology are also reviewed.

Technology and learning with understanding

The essence of technology, as used in this paper, is that it is a tool or set of tools that can facilitate teaching, learning and assessment when designed to do so. Technology can be an ambiguous term when applied to education since it can refer to such things as the use of chalk, paper, pencil, TV, video, calculator, and computers (Hawkins, 1996). Technology, as used in this paper, refers to the use of computer software that is either commercially produced to assist students in graphing or analyzing data, or software that has been designed to make instructional and assessment goals more visible to students through multi-media technologies, or software that can be used to record students while they use computers to solve statistics problems. Technology, as described here, goes beyond simple game-like drill and practice activities and instead is used to teach understanding when pedagogical principles guide its use in classrooms (Perkins, Crismond, Simmons, & Unger, 1995).

Understanding emerges from mental activities that involve constructing relationships, extending and applying knowledge, reflection, articulation, and making knowledge one's own (Carpenter & Lehrer, in preparation). Technology can assist students' understanding of statistics by serving as a partner to students while they engage in such mental activities, helping them accomplish things they might not be capable of doing without such assistance (Lajoie, in preparation; Lajoie & Derry, 1993; Salomon, Perkins, & Globerson, 1991). Every cognitive task consists of many processes that compete for attentional resources from a learner. Computers can share some of the attentional burden by performing lower order skills, such as drawing a graph, while students' resources are free to think about the meaning of the graph. Cognitive overload can be avoided if the technology is user-friendly (easy to use) as opposed to user-unfriendly (difficult to use whereby a learner spends more time learning how to use the computer than how to reason with statistical data).

One of the most basic ways that technology can be used to facilitate understanding is by acting as an intellectual mirror for students to peer into while problem solving (Schwartz, 1989). The more interactive the computer technology the more it can help students reflect on their problem solving processes. Within the context of statistical investigations students can get instant feedback on their hypotheses when they enter their data into the computer for graphing or analysis. Positive feedback would result in the graphs or analyses of interest, whereas negative feedback would results in graphs and analyses that were not expected. Technology can also facilitate reflection though "procedure capturing" (Kaput, 1992) whereby student computer actions in the context of solving problems could be replayed at the completion of problem solving. Such "replays" help learners focus on the processes they used to solve problems. Computers can capture an entire sequence of student actions, some of which are plans, and some of which are actions, and replay these actions for the student to reflect on (Lajoie & Lesgold, 1992). Some computer-based learning environments replay student actions and let students compare their actions with expert solutions so that better models of performance are available for inspection. These same types of environments can be designed to be adaptive to individual differences by providing assistance or coaching to those learners who need help. Procedure capturing, tracing, and replays can be an effective instructional tool since externalizing thought processes and actions helps students monitor their own performance. In essence, this monitoring is a form of self-assessment where students learn to assess themselves and become more independent. Teachers benefit from such tools in that they can review student traces to assess student difficulties and make important instructional changes that can remedy these difficulties.

The dynamic nature of technologies makes instruction an active and often student-driven activity where multiple examples are created dynamically by the learner's own actions rather than instruction that is static whereby pre-established sets of examples are selected by the teacher (Kaput, 1992). Kaput suggests that technologies provide a mechanism for dynamically externalizing mathematical notations. He describes many types of external notations used in mathematics (algebraic, graphs, tables) and discusses ways in which technologies can be used to externalize mathematical notations in a dynamic manner so that students can focus on how their own actions affect changes in such representations. Kaput (1992) describes the impact that dynamic vs. static notations can have when students are learning about statistical variance. In static mediums, variance can only be described through multiple examples, whereas in dynamic situations, students can interact with the data and see the immediate impact that changing one data point could have on, say, a histogram or scatterplot (Hancock, Kaput, & Goldsmith, 1992; Konold, Pollatsek, Well, Lohmeier, & Lipson, 1993; Rosebery, & Rubin, 1989). Kaput (1992) suggests that manipulating objects or mathematical entities on the computer can be more effective than physical manipulatives when the computer makes the links between such notations and mathematical models. Simply observing something happen on a screen will not guarantee that the learner has made the connection between, say, a graphical representation and the meaning behind the representation. It is often necessary to make the direct connection for the learner. Still it seems apparent that technology can empower students in the use of statistics by letting them actively engage in "doing" statistics (Hancock, et al., 1992; Konold, et al., 1993; Lajoie, Jacobs, & Lavigne, 1995; Lehrer & Romberg, 1996; Rosebery & Rubin, 1989; Scheaffer, 1988). Technology can be adaptive to individual differences in learning specific content by providing multiple types of representations; i.e., hypertext, graphics, animations, and videoclips and can remedy student misconceptions by providing alternative views of a phenomenon or through computer simulations (Nickerson, 1995).

Technology and assessment

From an instructional perspective, teachers sometimes struggle to find ways to clearly demonstrate what their expectations are of students. It is often difficult to demonstrate abstract concepts and consequently to tell students how they will be assessed in such contexts. Teachers' expectations can be made clear when technology is used to post concrete examples of good performance. When such examples are made available to students, students can better attempt to meet the teachers' expectations, thus facilitating both instructional and assessment goals. As new assessment techniques are being developed for statistics, it is necessary to consider how technology can help teachers manage the assessment process. Just as students will need time to adapt to new technologies in the context of instruction, teachers will need assistance in learning how such technologies can be used to facilitate instruction and assessment.

Obstacles to practice

Although there are benefits of technologies, there are also obstacles to incorporating them in the classroom. Kaput (1992) reviewed many of these obstacles in mathematics classrooms; i.e., lack of physical resources, such as phone lines for internet hook-ups, or up-to-date hardware and software, and most importantly lack of teacher training in technological resources. The pedagogical obstacles are the most severe. Teachers often do not have the necessary training in teaching statistics or in the use of technology in this context. As Hawkins (1996) points out, there are too many choices that teachers must make with little guidance; i.e., which are the more appropriate statistical and graphical packages. She suggests that although these new technologies provide speed and access to larger data sets and more complete access to the entire statistical investigation process, attention must be paid to subtle problems that can occur with these tools. For instance, automatic scaling procedures may reduce the learners' awareness of how scaling affects the graphical representation and subsequent interpretation of data. Teachers must draw students' attention to these features. Further, there are times when the speed of access that technologies provide may lead to what Ben-Zvi and Friedlander (1996) referred to as uncritical statistical thinking, where students generate a large number of statistical analyses and graphs but their focus is on extrinsic or aesthetic features of these activities rather than the underlying statistical meaning. On the other hand, Ben-Zvi and Friedlander (1996) found that these same technologies could lead to the meaningful handling of data and metacognitive skills related to statistical reasoning.

In summary, technology can be used to improve statistical instruction and facilitate student understanding and assessment by providing active interactive learning opportunities where multiple representations, and dynamic statistical notations, can be used to extend one's comprehension of statistical data, graphs and analysis, and by confronting misconceptions head on with appropriate simulations that shed light on statistical concepts. Technology can also externalize the learners' problem solving processes in a manner that could help individuals monitor or assess their own performance. As an instructional aid, teachers benefit from technology, since some instruction can be dynamically created through student interactions with technology. For example, student input can result in multiple examples that are constructed dynamically based on their input rather than instruction that is pre-established based on a contained set of statistic examples. Teachers may have more time to interact with individual groups of students and observe their statistical investigations. Through these observations and by reviewing student problem solving traces, collected and stored by the computer, teachers can change their instructional strategies to fit the needs of their students. The Authentic Statistics Project (ASP) described below, uses technology for both instructional and assessment purposes.

THE AUTHENTIC STATISTICS PROJECT

This chapter describes a technology implementation conducted as part of a research project rather than as part of a pre-existing statistics curriculum. A pilot of this project was conducted based on a partnership that was formed with a mathematics teacher who taught two grade eight classes. The teacher reviewed his curriculum content and placed our statistics instruction following a unit on graphing. In so doing he was able to build on students' prior knowledge of graphing by relating it to data presentation. Once this link was made, students were introduced to computer tools that would assist them in constructing graphs and performing analyses in the context of their statistical investigations. The use of technology alone will not ensure the development of statistical reasoning. Just as in other instructional settings a theory of the learning and instructional process must guide the use of technology. The pedagogical model that guided the ASP was the cognitive apprenticeship model which includes six methods for developing an optimal learning environment: modeling, coaching, fading, articulation, reflection, and exploration (Collins, Brown, & Newman, 1989). What follows is a description of how this model guided the design of ASP, and how the use of technology facilitated this approach to statistical instruction and assessment. This description will be followed by a detailed look at how the ASP project could be changed to better make use of the strengths of technology, and to better help teachers in the classroom.

Modeling, coaching, and fading

Modeling, coaching, and fading represent an interconnected view of instruction and assessment. Instruction may consist of modeling the knowledge and skills you want students to learn. However, in order to coach or assist learners in the context of such models, one must assess the learner’s level of skill acquistion and provide the right level of assistance for that learner. Fading of such assistance occurs when teachers’ assess that these same learners no longer need assistance.

ASP was conducted over a two-three week period during the regular mathematics class. Eight Macintosh computers were brought into the classroom for the duration of the project and three to four students worked at each computer workstation. Each group had access to either the teacher, or graduate student mentors, who were available to facilitate the statistical investigation process using computers. For the sake of simplicity, teachers and mentors will be referred to herein as teachers. Each teacher was responsible for modeling the skills and knowledge needed to carry out the statistical investigation processes. Modeling refers to demonstrating skills or knowledge. There were two phases to our project, the knowledge acquisition and the producer phases, and hence depending on the sequence of instruction different things were modeled. During the knowledge acquisition phase students were taught basic statistics concepts hand-in-hand with procedures for “doing” statistics. Students were introduced to statistics through a tutorial that taught them factual knowledge in the context of solving statistical problems with computer tools. Since this was a pilot study and we were still trying to identify what aspects of ASP would be effective, the tutorial was in paper form. Each group was led through a standardized set of instructions that would help students understand statistics as well as how to use the statistical package MyStat™ and a graphical software package Cricket Graph™ to solve statistics problems. The teachers were available to coach students whenever difficulties occurred while carrying out these tasks. The factual portion of the tutorial introduced students to a new vocabulary of statistical concepts that dealt with a basic understanding of measures of central tendency. Statistical concepts and graphic representations were modeled by first providing concrete examples of such things from local newspapers. For example, histograms, pie charts, percentages were demonstrated in the form of voter polls, sports statistics, and box office grosses for popular movies. Next, the types of procedures, or tools, that were available on the computer, i.e., the MyStat and CricketGraph applications, were modeled. These applications provide opportunities for representing the data in a multitude of ways, statistically, numerically, and visually.

Mini-experiments served as instructional activities to teach the students concepts of central tendency and variability. These mini-experiments serve to situate and extend the learning through what Bransford, Hasselbring, Barron, Kulewicz, Littlefield, & Goin (1989) referred to as a macrocontext; complex situations that require students to formulate and solve a set of interconnected subproblems. It was through these activities that the various software tools were modeled. The first mini-experiment included data gathering activities on Pulse Rates. Each group collected pulse rates in their group and entered their data into the computer and analyzed it using MyStat. Individual group data was then compared to data from the whole class. Discussions about the mean, mode, and median, data, sample, randomization, population, and range were generated in the context of these experiments. Following this exercise another group of students "ran-on-the-spot" for a minute and students collected, analyzed and interpreted their data by comparing it with the "at-rest" data. By comparing two group means predictions were made, hypotheses were tested, and results were interpreted. These experiments required whole class participation. Students were coached by instructions in the tutorial and by the teachers in how to use the computer to enter, analyze, and graph data. After completing the tutorial (knowledge acquisition phase), the producer phase began where each group was asked to construct its own statistics project.

The projects included several components of statistical problem solving: designing a research question, collecting data to answer the question, representing the variables and data graphically, analyzing data through statistical methods, interpreting data, and communicating understanding through class presentations. Since students worked in small groups to design their projects, much of the modeling and coaching would take place within the group when students shared different perspectives on the learning task. The type of modeling and coaching by teachers changed at this point in time. Teachers faded the type of procedural help (i.e., how to use the computer software) since students were adept at this after the tutorial, and increased the amount of modeling and coaching regarding such things as data organization needed for entering data into the computer in a manner that was conducive to the research question at hand. However, the most valuable type of modeling that occurred in the production phase was provided through the use of technology, in our library of exemplars.

Library of exemplars as a model for assessment

A computerized library of exemplars was developed (using HyperCard and Quicktime) as a means to demonstrate teacher expectations regarding how students might perform when creating a statistics project. The library of exemplars served to model these expectations by making assessment criteria visible to learners prior to their engagement in the statistical design phase. Examples were presented of both average and above average performance on each statistical component that students would be assessed on when they developed their own statistics projects (Lavigne, 1994; Lavigne & Lajoie, 1995). Technology was used to make assessment criteria more concrete in an effort to make it is easier for students to achieve these goals. Several components of the statistical investigation were modeled: what a research question looked like, how to collect data to answer a question, how to represent variables and data graphically, how to analyze data through statistical methods, how to interpret such data, and how to communicate about such projects in the context of class presentations. Each component in the library had concrete examples of what such criteria mean and how they are assessed. Although technology is the medium that is used for modeling the performance criteria, real students are actually modeling the performance criteria .

The exemplars were constructed by selecting video-tapes of students who had participated in a prior study and demonstrated different levels of performance on the above statistical components. Weaker and stronger examples of student performance were selected for each component so that new learners could compare these examples as they pertained to assessment criteria. Figure 1 illustrates the data analysis component. A textual description of the data analysis criterion appears on the left side of the screen, explaining what data analysis means in the context of a statistical investigation, and the maximum number of assessment points that students can get for this criterion. For each statistical component two video examples are available on the computer. To select the average or above average performance example students click on the respective video camera icon to see a video clip of how other students performed on that component. The average example illustrates a group that lists the minimum and maximum scores and the range. The above average example focused on the interpretation of their data. For instance, when asked whether they thought they would get the same results if they conducted their survey in Toronto as opposed to Montreal, the group responded that they were sure that different results would be obtained in Toronto, because they might market the products differently in another city. These examples help students engage in dialogues about what data analysis means in the context of previous student work and how such work was assessed. By making the criteria clear to students, all students could benefit by sharing a common understanding of the problem.

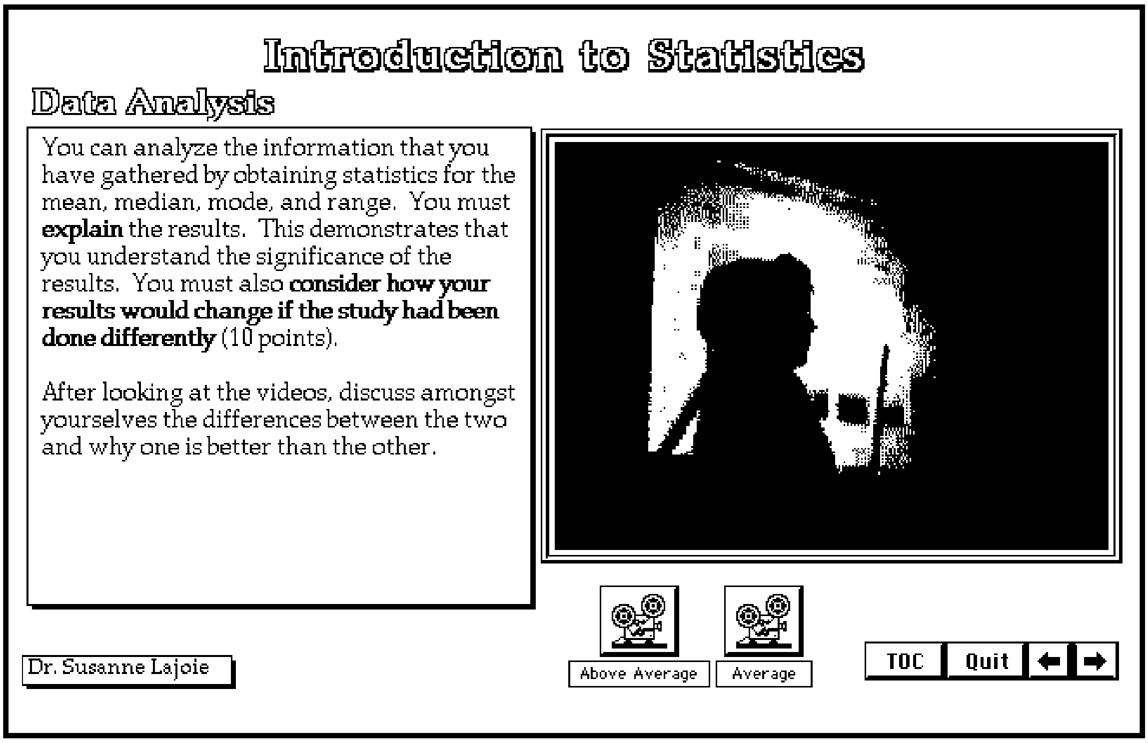

Figure 1. Data analysis exemplar

Reflection and self-assessment

The library of exemplars demonstrates how teachers can use technology to instruct students as well as to make their assessment criteria clear. Once students are given models of performance they can internalize the teacher's instructional goals and assessment criteria prior to constructing their own statistics projects. Furthermore, they can return to the computer at any time to refresh their memories of what a particular statistical concept means. Students can use such criteria to assess their own progress and compare their work with others. The use of technology in this case promotes students self-assessment and reflection on the cognitive components of the statistical investigation process. But how do we assess if students have internalized such criteria? One way that self-assessment was promoted was to ask each group to assess themselves on the same criteria they saw in the library of exemplars and to assess every other group's presentation on the same criteria. Assessment sheets were distributed to the groups with the maximum number of points available for each statistical component. Teachers independently assessed each group in order to monitor whether or not students' self-assessments were in accord with their own assessments of student work. After these assessment activities were completed, student self-assessment ratings were compared with teacher ratings, and ratings by other groups, and high correlations were found between the three types of assessments. These correlations tell us three things. First, students are not over-inflating their assessments in order to get high grades, rather their scores are in alignment with independent raters. Second, students are evaluating other students' performance fairly as opposed to under-rating other group's projects, which would indicate the negative effects of competition. Third, there is alignment between what is modeled and what is learned, or what is instructed and assessed. Technology, then, is useful in making teacher expectations clear and concrete enough for students to reflect on, internalize, and meet such expectations to the best of their ability. This exercise verifies the assumption that when teachers communicate their goals clearly learning will be facilitated.

Articulation and assessment

Students in ASP articulate their understanding of statistics in several ways. The most common form of articulation is a verbal one where individuals explain their reasoning with words, either orally or textually. In ASP, students verbally articulate their knowledge during the construction of their project where they must communicate their reasoning to their peers about a certain research hypothesis or data collection procedure. Students share their understanding amongst themselves and critique each other's ideas during the production phase. Groups also keep written journals documenting their statistical projects. Students also orally articulate their statistical comprehension during their presentation of their project, where they defend their reasoning about their hypothesis, their data collection procedures, choice of graphs, statistical analyses and interpretations of their findings. Articulation with technology is necessary in both these phases since the context of their statistical problem solving revolves around the computer as a tool used in their investigations. In fact, the liquid crystal display panel was used during student presentations to demonstrate different aspects of their statistical projects.

A less common form of articulation is what Kaput (1992) referred to as dynamic external notations. For example, the computer kept track of each group's use of the computer for graphing and analyzing their project data. All computer work was saved in the form of Screen Recordings (a Farallon product) which were exact computer-films of the sequence of activities that the students performed each day. An entire session on the computer could then be replayed like a videotape, only the video captures the procedures students used in the context of developing their statistical investigations. Teachers could view these recordings to examine whether or not these external representations were documenting students' progress in statistical reasoning.The dynamic external notations could be later used for assessment purposes to document changes in statistical reasoning.

The types of actions teachers considered important in terms of assessment of these external representations were: the number and types of graphic manipulations (i.e., pie charts), number and types of statistical manipulations (i.e., frequencies), the types of files that were accessed to develop projects, the creativity with which they personalized their presentations, and the types of difficulties students encountered in using the computer (i.e., attempts to run a statistics operation but not completing the operation for lack of knowledge). Off-task activities refer to activities such as playing with computer games or files that are not associated with the project.

Stiggins (1987) cautioned against the subjective interpretation of students' performance ratings and Linn, Baker, and Dunbar (1991) suggest that teachers be trained in a manner that would allow for some calibration in scoring performance data. We followed these suggestions by preparing a computerized demonstration of each of the types of actions students performed in the context of developing their statistics projects, as a way of assessing these external notations. This demonstration helped teachers be consistent in their assessments of such notations. Figure 2 describes the perform statistics criterion and what sorts of notations to expect. Text is used to describe the criterion and then teachers select PlayMT to get a screen recorded example of what performing statistics looks like in the context of ASP.

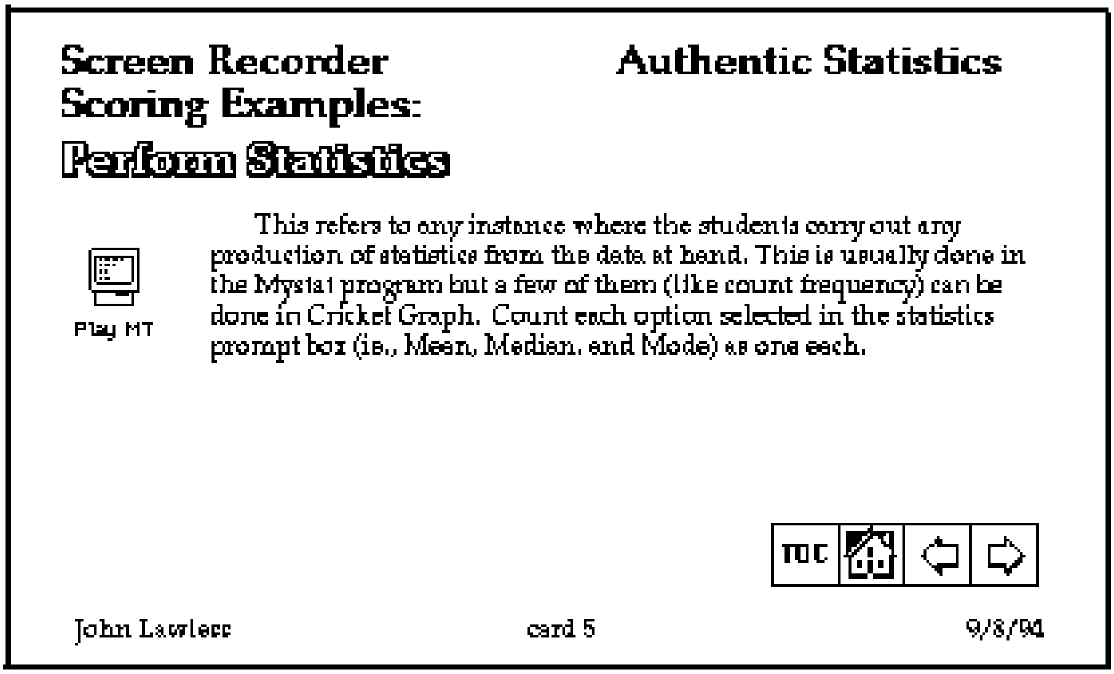

Figure 2. "Perform Statistics" criterion for scoring screen recordings

Along with the video clip are sound clips that describe the student activity and textual clips further elaborating the student actions. Together, the oral, textual and animated displays of each criterion help teachers score student notations in a more consistent manner.

The computerized demonstration serves to assist teachers in making decisions about what to assess and how to assess the students' problem solving capabilities. Once again, such assessment criteria for considering students' progress throughout their projects could be negotiated in that certain components may be more important than others, some added or deleted. The advantage of making the criteria visible through technology is that teachers can see contextualized video-screen segments of statistics work rather than single snapshots of performance. The screen recordings of students' progress in statistical investigations can tell us whether there is a relationship between types of notations students make and their actual understanding of statistics. Multiple forms of student data must be examined together in order to see the relationship between the type and number of graphs constructed and the final statistical presentation where justifications of student projects are communicated. In other words, we need to compare what students do with how they justify or reason about what they do in the context of statistical problem solving. Without such information it is impossible to say whether or not the different number of graphical representations students construct with the ease of technology demonstrates understanding or confusion.

Exploration

Students were given freedom to explore different aspects of the statistical investigation process and in so doing developed their own research question, collected their data, graphed, analyzed and interpreted such data. Computer tools were useful, in this regard, since students could generate and test their hypotheses as to what graph or what type of analysis might be appropriate for their question. The flexibility of the computer allows students to quickly find answers to the multiple hypotheses they may have at any particular time. One caveat, of course, is that students must understand how to pose such questions in a way that can be answered by the computer analysis. In other words, students must understand how to organize their data according to the type of analysis they choose to answer their questions. Teachers served as coaches, helping students when difficulties arose in the context of their statistical investigations. Given proper instructions, students can use computer software to graph their data in much less time then it would to draw a graph freehand. The speed of computer frees up time for students to reflect on the significance and interpretation of the data in the graph rather than on the construction of the graph. However, as mentioned above, student assessment must include evidence that students are indeed reflecting on the meaning of graphs and not just picking a number of graphs at random. Exploration as we see it should still be guided and monitored by teachers so that dead ends are not pursued and misconceptions are not reinforced.

SUMMARY AND IMPLICATIONS

This chapter provided an example of how technology promoted understanding of the statistical investigation process and improved assessment of statistical reasoning. The use of technology does not guarantee the coupling of instruction and assessment in statistical contexts; a theoretical framework should underlie the design of the use of technology in the classroom. The cognitive apprenticeship model is just one model that can be useful in this regard.

The six methods of the cognitive apprenticeship model were described to provide a framework for understanding the instruction and assessment process in the ASP project. Technology was not the sole instructional dispenser or assessment device in ASP. Rather, technology provided a context for facilitating student exploration of statistical problem solving through the pursuit of the multiple types of representations available to students for graphing or analyzing their data. Exploration in both the knowledge acquisition and producer phases of ASP was guided by teacher interventions or coaching when needed, and the fading of such assistance when no longer needed. Coaching and fading of assistance demonstrate the connection between instruction and assessment. Technology was successful at making the teachers' instructional expectations and assessment criteria clear through the library of exemplars. The computer was used to model the statistical components of problem solving by demonstrating how other students performed on such criteria. This library promoted students' reflection and critique of their or other student projects, and facilitated subsequent self assessments of their own projects. Furthermore, this use of technology helped align the teachers' assessment criteria with students' self-assessments criteria.

Technology, as used in ASP, was highly interactive, responding to students' requests by performing graphical and statistical analyses on the student data. All of the students' computer work and activities were recorded as screen recordings that could be played back and evaluated by teachers. The consistent or calibrated evaluation of this computer work was facilitated by the design of scoring templates. These provided teachers with multi-indexed examples of the types of activities and processes, including graphical and statistical manipulations, that could be assessed. There is much work to be done before technology-based projects such as ASP become fully implemented in the classroom. From a practical perspective, teachers need full access to technology in order to replicate or creatively add to the ASP work. Also, statistics must become a standard unit in the mathematics curriculum in order for teachers to allocate time and resources to this form of instruction.

Given our pilot data, improvements can be made to both the instruction and assessments of statistical problem solving. One improvement would be to use computers to standardize the knowledge acquisition phase so that the computer could take a key part in administering instructions and coaching, and in giving feedback. The computer could fade assistance when students have mastered the skills needed to manage their statistical investigations. Teachers can be most effective in coaching students in the production phase, helping students brain-storm about research questions and about how to collect, represent and analyze their data.

We know that technology can be used to effectively model components of the statistical problem solving process, and facilitate independent explorations and self-assessment. We need to spend more time thinking about the types of statistical knowledge that can best be exemplified using computers. For instance, Bright and Friel (in press) have examined how students develop graphical understanding. Results from such work are critical for designing more effective uses of technology in teaching and assessing statistical reasoning. If we understand how transitions in statistics learning occur, simulations can be designed along with direct instruction where key concepts and processes are modeled for the learner.

A single form or type of assessment is not adequate for representing statistical problem solving. Rather, multiple types of assessment must be combined. In ASP, oral forms of articulation were examined along with the recordings of students' computer actions during problem solving. Note that the computer can collect observations but it is the teacher who must interpret such evidence in order to understand and document student progress over time.

Perhaps as our understanding of how to assess these recordings or notations evolves, part of the assessment can be done dynamically by specialized computer programs. Methods must be found to ease the assessment burden of the teacher. Teachers need workshops that address both the cognitive aspects of statistical reasoning and how to assess changes in such reasoning, along with workshops that help them use technology in the classroom. At the same time, teachers must learn to use tools for managing complex forms of evidence of student learning and become adept at reliably interpreting such data.

Acknowledgments

I would like to acknowledge Nancy Lavigne for her assistance in all phases of this work, and Steve Munsie, and Tara Wilkie who have played a large role in this research. I acknowledge André Renaud who programmed the library of exemplars, and John Lawless who programmed the screen recording stack. I would like to thank Brent Blakely, Marlene Desjardins, Bryn Holmes, Cindy Finn, Steve Riesler, Jody Markow, Mike Thibault, Tina Newman, Litsa Papathanasopoulou for their assistance in various phases of data collection. I gratefully acknowledge the assistance of Phil Knox, a gifted mathematics teacher, who generously provided his classroom as a testing ground for this research. Special thanks to Thomas Romberg for his support of this research. Preparation of this document was made possible through funding from the Office of Educational Research and Improvement, National Center for Research in Mathematical Sciences Education (NCRMSE).