5. THE USE OF TECHNOLOGY FOR MODELING PERFORMANCE STANDARDS IN STATISTICS

Susanne P. Lajoie
McGill University

OVERVIEW

The goal of the Authentic Statistics Project (ASP) is to make statistics meaningful to middle school students, particularly grade 8, and to assess the progress students make in learning statistics. One way of enhancing the value of statistics to middle school students is to demonstrate how statistics can be used to answer important questions and make everyday decisions. Within this context, students learn to perceive statistics as a valuable tool rather than a bother. This paper describes how ASP uses technology to facilitate both instruction and assessment by modeling performance standards for the statistical investigation process.

THE AUTHENTIC STATISTICS PROJECT (ASP): A DESCRIPTION

ASP is designed to provide instruction and assessment of descriptive statistics in computer-based learning environments. The goal of ASP is to provide learners with opportunities to solve a variety of statistical problems, to reason about statistical information, and to communicate their understanding of statistics. In the ASP, students used the statistical process of investigation to generate research questions that were meaningful enough to pursue (Lajoie, Lavigne, Munsie, & Wilkie, in press). The mathematics classroom was transformed into small groups of design teams, where each team constructed a research question, collected data to answer it, represented the data graphically, analyzed the data, and presented their projects to the classroom. Each group was provided with technological supports that were intended to facilitate the investigation process. New technologies are changing the ways in which problem situations and methods of representation are used in instruction (Romberg, in press). Technology, as used in ASP, served to facilitate both instruction and assessment.

In order to facilitate the learning process, ASP anchors statistical concepts and the statistical investigation process through examples that model the use of concepts for a particular process on a variety of real-world problems. Thus, ASP uses modeling as its primary instructional technique to teach students about descriptive statistics. Modeling involves an expert or more skilled individual carrying out a task so that students can observe and build a representation of how to solve a task (Collins, Brown, & Newman, 1989). Students are then given opportunities to apply their knowledge and skills to a variety of problems. The key is for students to learn fundamental concepts through modeling, to apply their knowledge and skills to a variety of contexts, and to reason about the statistical process. Providing multiple contexts can help students build and transfer their knowledge.

WHAT AND HOW TO MODEL?

Observation is a key component to learning, especially when one observes skilled individuals solve a complex task. After observation comes practice. Technology can be used to model complex skills and knowledge to students. ASP is designed to support such modeling.

The selection of what to model was guided by the *National Council of Teachers of Mathematics Curriculum and Evaluation Standards* (National Council of Teachers of Mathematics, 1989) for middle school statistics. The standards emphasize a problem-solving approach to statistics where students collect, graph, analyze, and interpret statistics. Thus, ASP decomposed the statistical investigation process into exemplars of four basic components: designing a research question, collecting, graphing, analyzing, and interpreting data based on the research question (Lajoie et al., in press; Lavigne & Lajoie, 1996). Since statistics was foreign to most of these students, exemplars were designed that would model expertise for each component so that students would be able to model these skills for their own statistical investigations. The teacher's expectations for students are made clear by these exemplars. Furthermore, the teachers can make their assessment standards clear to students in the same fashion. One of the *Assessment Standards for School Mathematics* (National Council of Teachers of Mathematics, 1995) is to make assessment criteria open to learners so that they will understand the teacher's expectations prior to engaging in a task. When students understand the learning goals and the ways in which they are expected to achieve them, it is easier for them to align their performances to the criteria. Moreover, when students understand the criteria, it is easier for them to set high expectations for themselves and consequently monitor their own progress (Diez & Moon, 1992; National Council of Teachers of Mathematics, 1995). One way of clearly communicating such expectations to students is by publicly posting these performance criteria through the use of technology.

Technology is used in this project both as a cognitive tool for modeling performance standards in statistics and as a tool for statistical analysis and graphing. This paper will describe how technology can situate students in authentic learning situations where students drive the curriculum by their own intrinsic interests. The performance standards that are modeled using the computer are standards for generating research questions, collecting data, analyzing data, graphing data, and presenting interpretations of the data. Statistical tasks were designed to provide students opportunities to reflect, organize, model, represent, and argue within and across mathematical domains. The instructional focus is on statistical problem solving, reasoning, and communication. Three computer-based learning environments have been developed for teaching grade 8 mathematics’ students about descriptive statistics. A computer-based learning environment is a place where learning occurs, such as a classroom, but the computer provides the instructional platform. The environments described here include *Discovering Statistics*, *Authentic Statistics*, and *Critiquing Statistics*. Different statistical skills are taught using these environments (described below). ASP uses each of these environments to provide learners with an opportunity to learn about descriptive statistics and statistical investigation, to apply their knowledge by designing and engaging in their own investigation, and to reason about statistics by critiquing investigations conducted by others. Research on the *Authentic Statistics* environment has been conducted. The *Discovering Statistics* and *Critiquing Statistics* environments are still under development and we plan to complete and test these new environments in the next few years. My goal in discussing these environments is to establish a context for how to model statistics using technology. These environments and the sequence (see Figure 1) in which they are introduced are described below.

A LOOK INSIDE THE CLASSROOM

ASP is designed according to a producer/critic model whereby students generate their own experiment (producer phase), critique research conducted by former students (critic phase), and create a second experiment (producer phase). Kelley and Thibaut (1954) suggested the role of critic and evaluator is first learned in a group situation and then becomes internalized as self-regulatory skills. If critiquing can be modeled through technology, self-regulation can occur through modeling the skills of predicting, verifying, monitoring, and reality testing in an effort to foster comprehension regarding the statistical processes of investigation. When students learn to evaluate others, they ultimately learn to assess their own statistical problem-solving processes. The producer-critic model has been used successfully in reading, where one student produces summaries and another student critiques the summary and then the roles are reversed (Palinscar & Brown, 1984). The notion is that the critic phase of ASP will enhance students' reasoning skills and will consequently, result in a stronger understanding of the investigation process that will be reflected in the second experiment. The sequence in which students are presented with the computer learning environments and their relationship to the producer/critic model is illustrated in Figure 1.

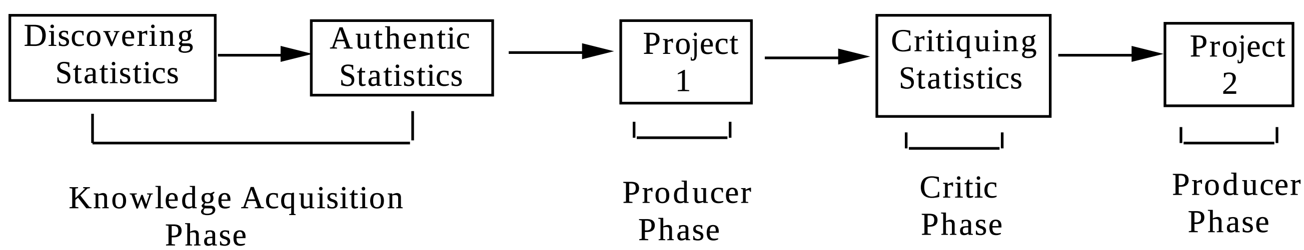

Figure 1: Procedural sequence of the Authentic Statistics Project

Discovering Statistics is a computer tutorial created to standardize the instruction presented to students prior to their producing their own statistics projects. *Discovering Statistics* is still under development and will be piloted in the near future. This environment integrates the teaching of statistical concepts with the teaching of appropriate computer skills students need to construct their projects. The content is taught within the context of the statistical investigation process. For instance, instruction of measures of central tendency and measures of variation are taught within the context of data analysis. Students are taught how to use EXCEL (Microsoft Corporation, 1992), a computer software program that allows students to both graph and analyze their data. Once students have acquired the relevant knowledge and skills, they are required to demonstrate their understanding of statistics and the investigation process by designing and conducting their own statistics project. To ensure that students understand the task and how they will be evaluated, they are shown *Authentic Statistics*. This tool models the investigation process by providing concrete examples of performance on each aspect of the investigation process (i.e., designing a research question, collecting data, analyzing data, representing data, and interpreting findings). Each aspect is presented to students as a criterion for assessing performance on the task. Both *Discovering Statistics* and *Authentic Statistics* represent the knowledge acquisition phase.

The task of designing and conducting a research experiment represents the first producer phase. This task also requires that students communicate their understanding of statistics by presenting their research to peers. Thus, students must explain their reasoning behind the design and implementation of their research. To enhance students' reasoning abilities, *Critiquing Statistics* will be presented. This environment will provide students with opportunities to critically examine and assess two complete presentations given by former students (assessments will be based on the criteria presented in *Authentic Statistics*). This critic phase will allow students to further develop their statistical understanding and reasoning abilities. Finally, students will be required to perform a second statistics project that consists of the second producer phase. This second phase is included to examine whether or not the reasoning task facilitates students' understanding of statistics. Observations from previous research suggest that allowing students to present and question others about their research helps them to better understand statistics. Hence, the inclusion of a reasoning task after having presented one statistics project is expected to help students develop their second project.

ASP classrooms have been designed for grade 8 mathematics students, where students work cooperatively in groups of three at their own computer work station (generally eight per class) for the entire unit. Teachers involved in ASP classrooms include the mathematics teacher and graduate research assistants involved in the ASP research. The role of the students is to follow the sequence presented in Figure 1. The role of the teachers is to facilitate student inquiry and assist when difficulties arise. The classroom culture has already been pre-established prior to our introduction of ASP. This culture is one in which group problem solving is the norm. Hence, students are used to working in groups and sharing responsibilities. Team work is essential for all aspects of ASP. The mathematics teacher provides linkages between the graphing unit and unit on averages as a way to introduce students to the ASP focus on statistical investigations. The ASP unit generally takes two weeks to complete; however, with the introduction of the *Discovering Statistics* and *Critiquing Statistics* environments will add another week or two.

DESCRIPTIVE STATISTICS HYPERCARD™ STACK: *DISCOVERING STATISTICS*

Discovering Statistics is a computer-based learning environment that provides standardized instruction of concepts and procedures in descriptive statistics for grade 8 mathematics students. The instruction is provided in the context of statistical problem solving and scientific investigation. *Discovering Statistics* is a HyperCard™ (Claris Corporation, 1991) stack that serves several functions. First, it models the types of knowledge students need to produce their own projects: that is, declarative knowledge (knowledge of facts); procedural knowledge (knowledge of how to compute statistics and graphs statistics using a computer); and conceptual knowledge (knowing how to apply their knowledge to their own projects that require problem solving). The tutorial briefly introduces learners to the notion of statistics, namely, that statistics is used to describe and make predictions about data that are collected to answer a particular question. An example research question is provided to help learners situate the process of doing statistics in the context of a problem. The *Discovering Statistics* tutorial provides instruction and models the use of concepts in the area of descriptive statistics in terms of the components of the scientific investigation process. Data collection, for example, requires a basic comprehension of concepts such as randomization, sample size, and representative sampling. *Discovering Statistics* models this knowledge for students using demonstrations and then provides practice opportunities where students apply their knowledge to new situations. Instruction consists of providing definitions of concepts as well as illustrating examples, in the form of "demos," which model the use of such concepts in a variety of contexts. Providing learners with concrete examples that clearly illustrate the meaning of concepts is an important feature of the tutorial, particularly given student difficulties with the abstract nature of statistical concepts. Learners have access to a flowchart or "map" illustrating the concepts that will be taught in the tutorial and how these are linked to the processes of data collection, data analysis, and data representation (i.e., data and graphs), which are also emphasized in the *Authentic Statistics* and *Critiquing Statistics* environments. Figure 2 provides an overview of the statistical concepts and procedures illustrated in this map. Students can access information about any of these processes directly by clicking on the appropriate button. Once a student clicks on data analysis the student is brought to this section of the tutorial. This map is available at every point in the tutorial because it serves as a navigation tool.

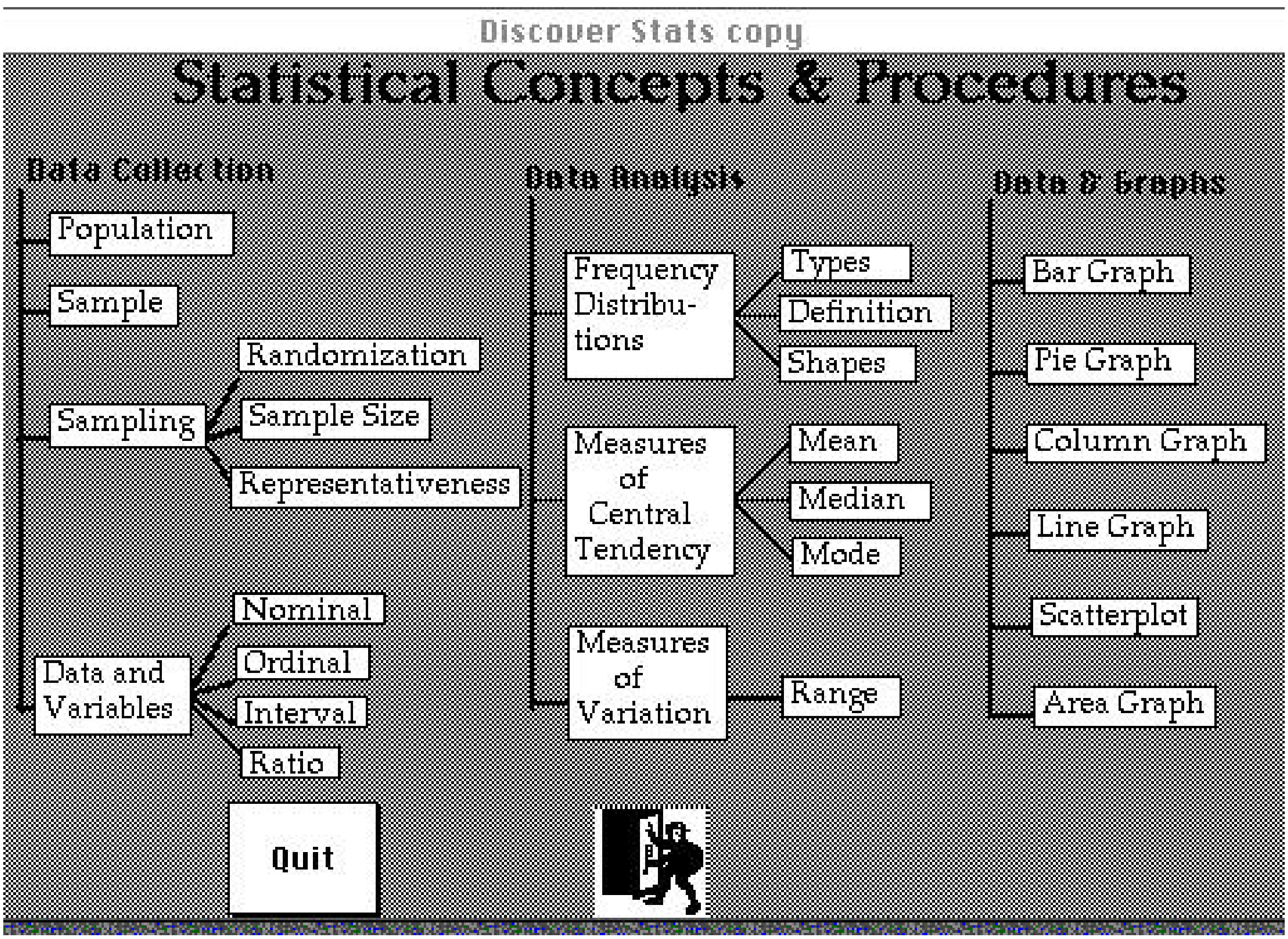

Figure 2: Table of contents for *Discovering Statistics*

Information about each concept is presented to learners in four ways: (1) as a lay definition of the concept in question; (2) as an example (written in text) illustrating how the concept is used in the context of an experiment (the same example is used throughout the tutorial so that students can build their understanding of statistics in a particular context); (3) as a statistical definition of the concept to help learners develop and become familiar with the language of statistics as well as to immerse themselves within the culture of statistics; and (4) as an example using text, sound, screen recordings (computer recordings of actions performed on the computer), or animations that illustrate how a particular concept is used in a problem (this is different from that presented earlier to facilitate students' understanding of statistics across contexts) and how procedures can be applied using the appropriate software (such as EXCEL for the entry, analysis, and representation of data). Any unfamiliar terms are underlined in italicized bold and referred to as "hot buttons." To obtain a definition of a term, learners merely have to click on the word and a pop-up window appears with the relevant information.

In addition to learning about statistics through definitions and examples, students are given an opportunity to apply their knowledge to problems provided in the tutorial. A mechanism has been created so that student activities (e.g., problem solving in "Practice" sessions as well as the type of information and frequency of review sought by students within the tutorial) can be traced by the computer. The extensive use of student user files will provide researchers and potential teachers with the means of assessing performance directly. In *Discovering Statistics*, assessments are embedded within the instruction through practice sessions that are provided both within subsections, to ensure an understanding of particular concepts, and at the end of each major section (e.g., data collection) to assess overall comprehension of the relationship among specific concepts and the various components involved in the process of statistical investigation. Assessment of learning arising from the tutorial is thus inextricably linked to the instruction. *Discovering Statistics* provides the background or prerequisite knowledge that students build on in subsequent phases of ASP. Consequently, there is overlap in the types of knowledge that students acquire in the *Discovering Statistics* and the *Authentic Statistics* environments. The most obvious overlap in content is that both environments model knowledge that is required to be successful in the overall statistical investigation process. However, *Discovering Statistics* is a precursor to the *Authentic Statistics* environment in that the skills acquired in the former lead to more informed decisions about how to design a research question, which is modeled in the latter. A description of *Authentic Statistics* is provided below.

THE *AUTHENTIC STATISTICS* ENVIRONMENT: A LIBRARY OF EXEMPLARS

Our assumption was that "making assessment criteria transparent" by demonstrating exemplars of student performance would facilitate learning. When students are made aware of what is expected of them, it is easier for them to meet those expectations, to assess their own progress, and to compare their work with others. Technology was used to demonstrate performance standards for specific aspects of the statistical investigation process. The *Authentic Statistics* environment provides a library of exemplars (Frederiksen & Collins, 1989) of these performance standards for grade 8 students (Lajoie, Lavigne, & Lawless, 1993; Lajoie et al., in press; Lavigne, 1994). HyperCard™ drove the interactions in *Authentic Statistics* in which students were shown concrete examples of statistical components by using QuickTime™ (Apple Computer Inc., 1989) to display digitized video tapes of student performance. Media Tracks™ (Farrallon, 1991) software was incorporated in the environment to display computer screen recordings of graphs and data used in student projects. Textual descriptions of our scoring criteria accompany the video clips and screen recordings.

Students are provided with an overview of what the library is used for and how it can help them develop their own projects. The total project was worth a maximum of 50 points, and assessment values for each component are provided. Figure 3 presents the Table of Contents for the *Authentic Statistics* environment.

There are six assessment components: (1) the quality of the research question, which is evaluated based on how clear the question was, whether or not all the variables were specified, and whether or not all levels of each variable were discussed (5 points); (2) the data collection, which is evaluated based on how students gather information pertaining to their question and whether or not they avoided bias in the data collection process (10 points); (3) the data presentation, which is evaluated based on how the data is summarized and presented (i.e., whether or not the types of tables, charts, and/or graphs that the students constructed were appropriate) (10 points); (4) the data analysis and interpretation, which is evaluated based on the choice of statistics selected to analyze a dataset, as well as whether or not an understanding and interpretation of the data analysis was demonstrated (10 points); (5) the presentation style, which is evaluated based on the thoroughness of the explanations regarding the project and on how well it was organized (10 points); and (g) the creativity, which is evaluated based on the originality of the statistics project (5 points).

A student could access information about each category at any time by "clicking on" the icon for a textual and visual demonstration of average and above average performance. To select the "research question" criterion, for instance, a student would click on the image corresponding to this standard (see Figure 3). The student would then receive information about that criterion (see Figure 4).

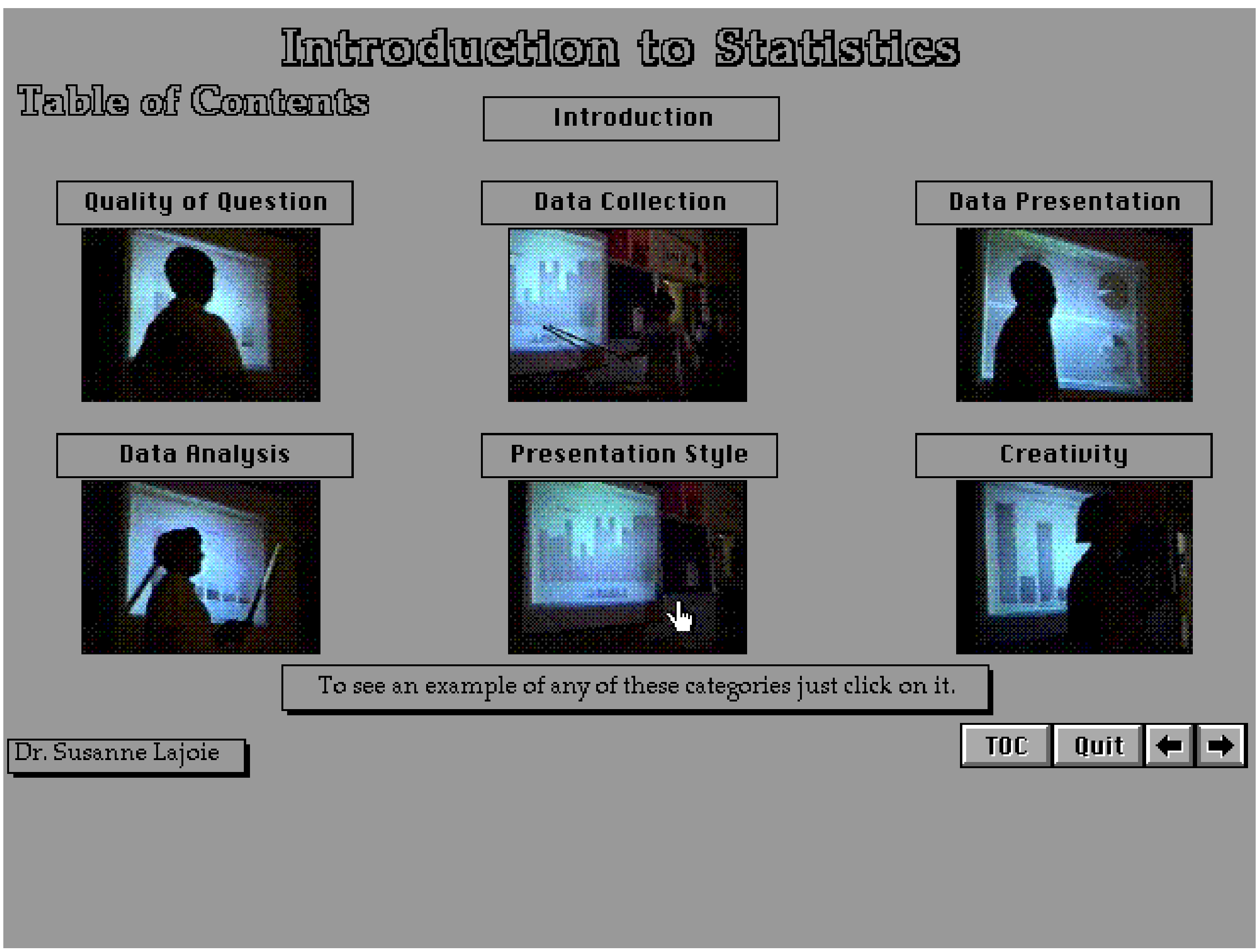

Figure 3: Table of contents for *Authentic Statistics* environment

Text describing how to develop a research question, for instance, would be read by the student. Examples of how former students performed on this criterion would also be viewed. The student could see what average or above average performance looked like by viewing videoclips. Clicking on the average performance button, for example, illustrates a videoclip of a student from another class presenting the following research question: "What is your favorite fast-food restaurant?" This videoclip conveyed to the new student the type of performance that was acceptable on the research question criterion. The above average performance example illustrates a videoclip restating the group's question as "What is your favorite fast-food restaurant between Harvey's, Burger King, McDonald's, and Lafleur's?" This videoclip demonstrated stronger performance on the criterion because the categories given to the sample were specified. New students view the examples in *Authentic Statistics* and respond to textual prompts by discussing and reasoning about performance differences with their group (see Figure 4). After viewing information about each criterion and discussing differences between performance levels, the group develops its own project and aligns its performance to criteria accordingly. Students have opportunities to internalize such criteria before they start their own statistics projects and can return to the computer at any time to refresh their memories of what a particular statistical concept means.

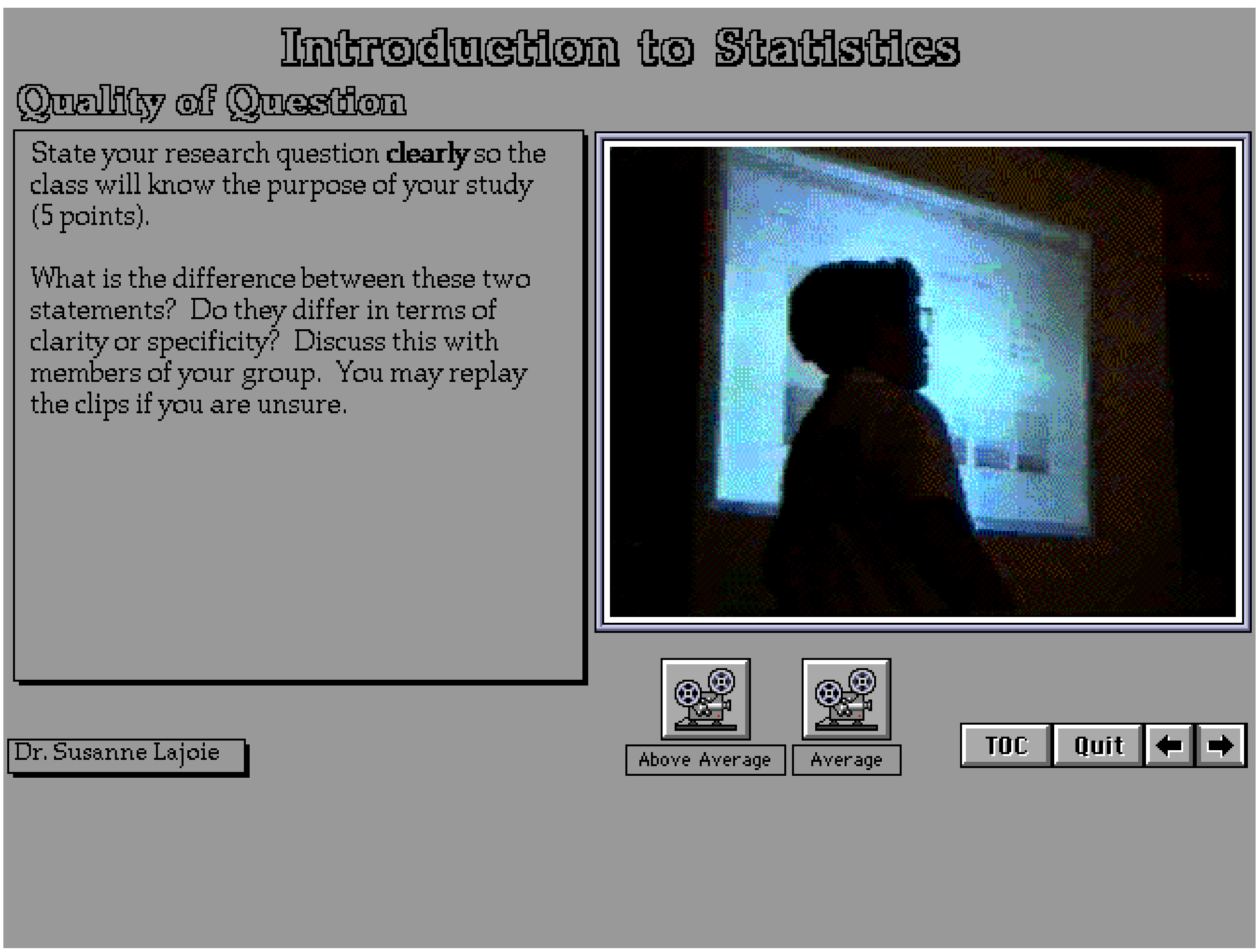

Figure 4: Quality of research question criterion

Method

In a pilot study conducted in the Montreal area, 21 multicultural students from one grade 8 mathematics class were divided into eight groups and each group worked at their own computer workstation. After receiving a tutorial, similar to that described in *Discovering Statistics*, students explored *Authentic Statistics*. Two conditions were developed for the exemplars: a video-based condition in which students saw video clips and textual descriptions of average and above-average performance; and a text condition in which students were only given textual descriptions. Groups were randomly assigned to the two conditions, which resulted in four groups in the video condition and four in the text condition. Group composition was matched on ability groupings. A pretest and posttest were administered and assessment data were collected for each group's project.

Results

A condition (video, text) by test (pre, post) split-plot analysis of variance was conducted (Lavigne, 1994). No condition effect was found [$F(1,16) = .01$, $p > .05$], however, there was a test effect [$F(1,16) = 50.31$, $p < .05$]. There was no interaction between condition and test [$F(1,16) = .92$, $p > .05$]. The means are presented in Table 1.

Table 1: Means for the condition by test ANOVA

Condition	n	Pretest		Posttest	
		Mean	SD	Mean	SD
Video	11	6.64	1.12	13.82	1.74
Text	7[a]	5.71	1.41	15.14	2.18

Note: Maximum score = 50

[a]Three subjects were excluded from the analysis due to attrition.

Both conditions produced significant changes in statistics performance from the pretest to the posttest. This finding is encouraging because it reaffirms the assumption that making criteria explicit can enhance learning. We anticipated that the video condition would have a greater effect than the text-only condition because it used multimedia. This prediction was not confirmed. Further exploration of the results indicated that students in the video condition outperformed the text condition for specific statistical concepts (i.e., sample representativeness and sample size). The use of multimedia for modeling statistical reasoning needs to be explored further.

Student alignment with the goals of instruction and assessment were examined by comparing student self-assessments with teacher assessments, and with assessments by other students. Doing this verifies that teachers are communicating their goals clearly and that the exemplars are clear, and provides a mechanism for both the teacher and student to evaluate whether learning is occurring. Three types of assessments were made. First, teachers assessed groups on each criterion. Second, each student group rated themselves on the criteria by discussing their performance and reaching a consensus in their assessment. Finally, each group assessed other group projects in the same manner. A *t*-test was performed to examine the mean assessment score differences between self-assessments and group assessments, and no significant differences were found [$t(1, 82) = .457$, $p < .05$]. The self-assessment means ($M = 13$) were slightly higher than the group ratings ($M = 12$), but there was not a significant difference between them. This indicated that competition to do well on such tasks did not over-inflate their self-assessment or their assessment of others. Six teacher/researchers rated the group projects as well. There was consensus between the teachers and the student groups regarding which group gave the best presentation. These findings suggest that making criteria open and visible to groups through technology helps them to align their self-assessments with the teachers' assessment. These findings are encouraging on two fronts. First, they tell us that modeling student performance can closely align teacher expectations with student performance. Secondly, they confirm that technology can be a useful tool for making abstract performance standards clear and open to learners. However, in reviewing our data we found that modeling performance standards can have an interesting drawback, or asset, depending on your viewpoint. Namely, what you model, you get. By examining student performance on the criterion of establishing a research question, we found that 93% of our groups constructed a research question that was similar to the question modeled in *Authentic Statistics*. Consequently, 93% of the research questions involved designing a survey. Modeling can have a major impact on student performance; consequently, we must consider what is an appropriate model to provide to students. If flexibility in research design is a goal of statistical problem solving then we must model the variety of designs we want students to learn and use.

The next step in the research with ASP considers this flexibility issue in the design of a critiquing environment in which technology is used to extend students' statistical understanding. The goal of *Critiquing Statistics* is to promote reasoning about the statistical investigation process.

CRITIQUING STATISTICS

A critiquing environment is under development that will go one step beyond *Authentic Statistics*. In the *Authentic Statistics* environment students observe and model performance standards and apply these standards to their own projects. The addition of a critiquing environment will provide these same students with a mechanism for facilitating discussion about what could be done better in the statistics projects that they view through technology. Furthermore, it provides these same students with multiple models of statistical performance rather than limiting their observations to a few examples. The critiquing environment uses digitizing videotapes of student projects to focus the dialogue on the components of statistical problem solving. The purpose of the environment is to promote small group discussions about the appropriateness of statistical methods, data collection procedures, graphical representations, analyses, and interpretation of data. The empirical question is whether or not this critiquing process can help students perform better in designing their subsequent statistics projects. In addition to observing other students' performance on the standards, active critiquing of such performance might further engage students in statistical reasoning. The intent is to build a community of scientific reasoners who share their knowledge, reasoning, and argumentation. These classroom dialogues can be used to document student reasoning in complex learning situations that might not be assessed within the computer learning environment. These dialogues will be recorded and analyzed to determine the nature of statistical reasoning and to document changes in student reasoning that are facilitated by the group discussions.

If critiquing can be modeled through technology, self-regulation can occur through modeling the skills of predicting, verifying, monitoring, and reality testing in an effort to foster comprehension regarding the statistical processes of investigation. When students learn to evaluate others they ultimately learn to assess their own statistical problem-solving processes. Differences between producing and critiquing statistical investigations will be further explored in ASP. The relationship between statistical reasoning and problem solving will be traced between the different stages of learning. For instance, does reasoning during a critiquing phase extend to visible changes in the research questions that are produced by students and the types of data collection, graphing and analysis procedures that are followed?

What will it look like?

Students using *Critiquing Statistics* will be required to critique two research projects presented by former students based on their understanding of the statistical investigation process (i.e., the quality of the research question, the data collection, the data analysis, and the data presentation). The projects will vary in terms of their strengths and weaknesses on the various statistical components, in that one group may be superior to another on generating a research question, but not on their methods of data collection. Students using the critiquing environment will have to determine these strengths and weaknesses themselves.

Critiquing exercise

Two presentations of research projects developed by groups of students participating in previous studies will form the basis of the critiquing environment. Both presentations will be digitized and linked to a HyperCard™ stack. The purpose in digitizing the videotaped presentations rather than simply showing the original videotapes is twofold. First, all relevant information pertaining to the presentation would be presented on one medium. Data, analyses, and graphs that are not clearly discernible on the videotape can be copied from the groups' computer files and presented on the computer screen at the same time the presentation is displayed. Similarly, the criteria by which students critique the performances could be displayed next to the presentation on the screen. Students' critiques would thus be guided by the criteria that would be readily accessible on the screen. Second, student actions can be documented better on the computer than on a videotape of students examining a videotape. Students' requests for information through rewind, fast-forward, and replay options could be documented, allowing researchers to isolate specific information that influenced the students' critiques of the performance on a particular criterion (or investigation process). Similarly, students' original, revised, and final assessments of the presentations, in terms of grades, can be documented easily by the computer. Changes in assessment along with an index of the information sought by students can provide some insights into the types of information that were most salient and, consequently, affected the assessment process. Finally, students can explain the reasoning behind their assessments and suggest ways the group presentations could be improved by typing their explanations and suggestions directly in the environment or document. To complement this information, students' verbal explanations will be audio-recorded.

Once students critique an entire presentation, they will be shown how an "expert" statistician would critique the research project. The purpose of providing students with an example of an expert critique is to provide feedback about the assessment process and bring students' evaluations in alignment with the assessment goals of this project. This feedback should facilitate their critique of the second presentation. Furthermore, making expert thinking visible through the expert's critique can help students' internalize a mental model of the statistical investigation process.

Tentative presentation format of *Critiquing Statistics*

The presentation format of *Critiquing Statistics* is still tentative. Students will view two research presentations, one at a time. The goal is to have students critique each presentation individually to ensure that they examine each digitized video presentation fully. Each group will discuss how to assess each component and then reach consensus about the assessment scores for each statistical component. The groups will discuss which of the two presentations is better and why. *Critiquing Statistics* is conceived of as containing four major components. The first component involves viewing the descriptions of each statistical criterion with specific questions directing students about what is important to address in statistics projects and how answering these questions leads to meeting the assessment criteria (see Figure 5). The second component involves playing the digitized presentation, discussing possible assessment ratings, and then allocating a score for that presentation on a particular criterion (see Figure 6). Close-up pictures of data files (in Data Collection), analyses (in Data Analysis), and/or graphs (in Data Presentation and in some cases Data Analysis) will be provided because these tend to be unclear in the videotape. The students discuss their reasoning for each assessment before they type their ratings for each question on the line next to the maximum value allotted for each question. The presentation can be replayed at any time. The third component consists of a debriefing session where students will be required to elaborate on their assessments.

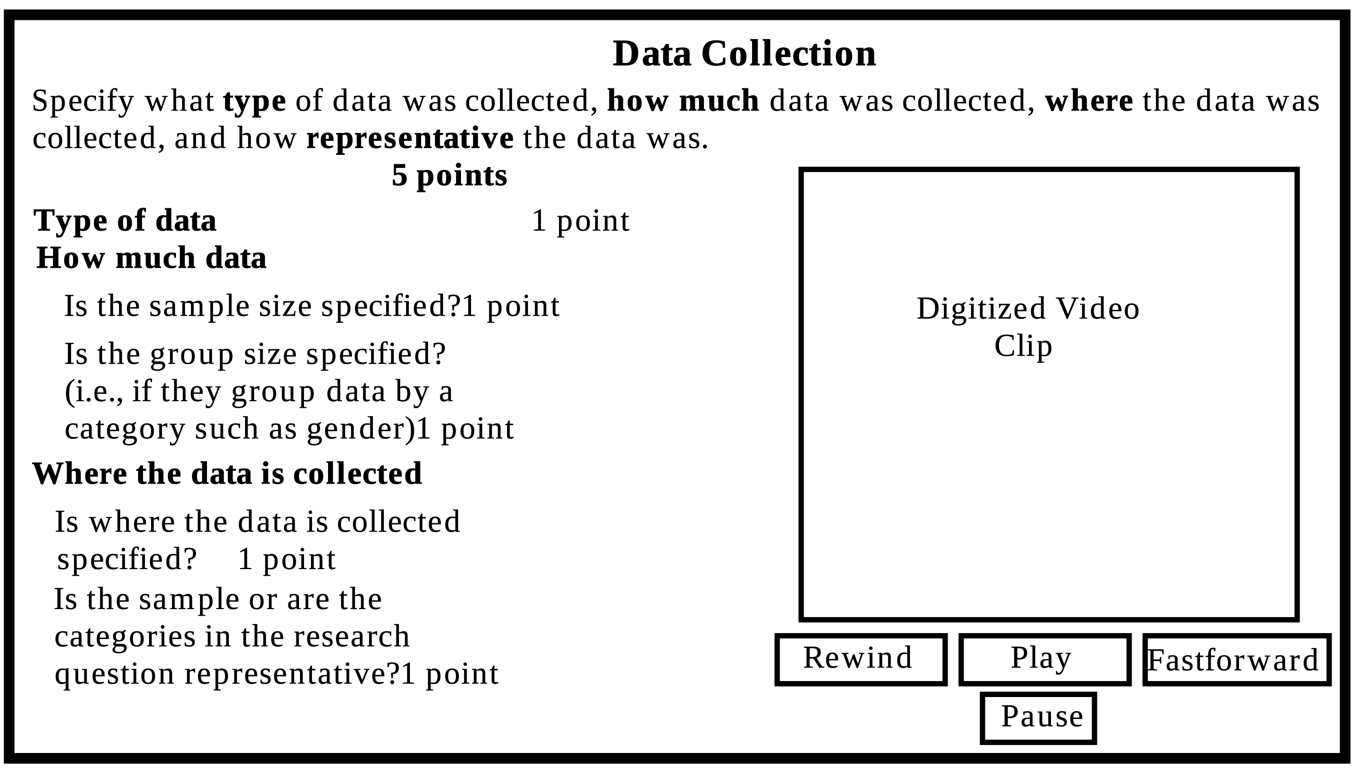

Figure 5: Example of general information for data collection component of *Critiquing Statistics*

This debriefing component will entail comparing the two presentations and indicating which was better and why. At this point, *Critiquing Statistics* will list the ratings that students assigned on each criterion for each presentation (a mechanism will be developed so that ratings are tracked by the computer and shown to students at this point). This list will allow students to review their ratings. The fourth component consists of an "expert" critique, so that students can compare their critique of the presentations with an "expert" who outlines his/her reasoning behind his/her critique. The overall process involved in performing this reasoning task is expected to enhance students' understanding of statistics. Students' understanding and their reasoning which was developed through the *Critique Statistics* task is expected to be reflected in the design, implementation, and presentation of a second research project.

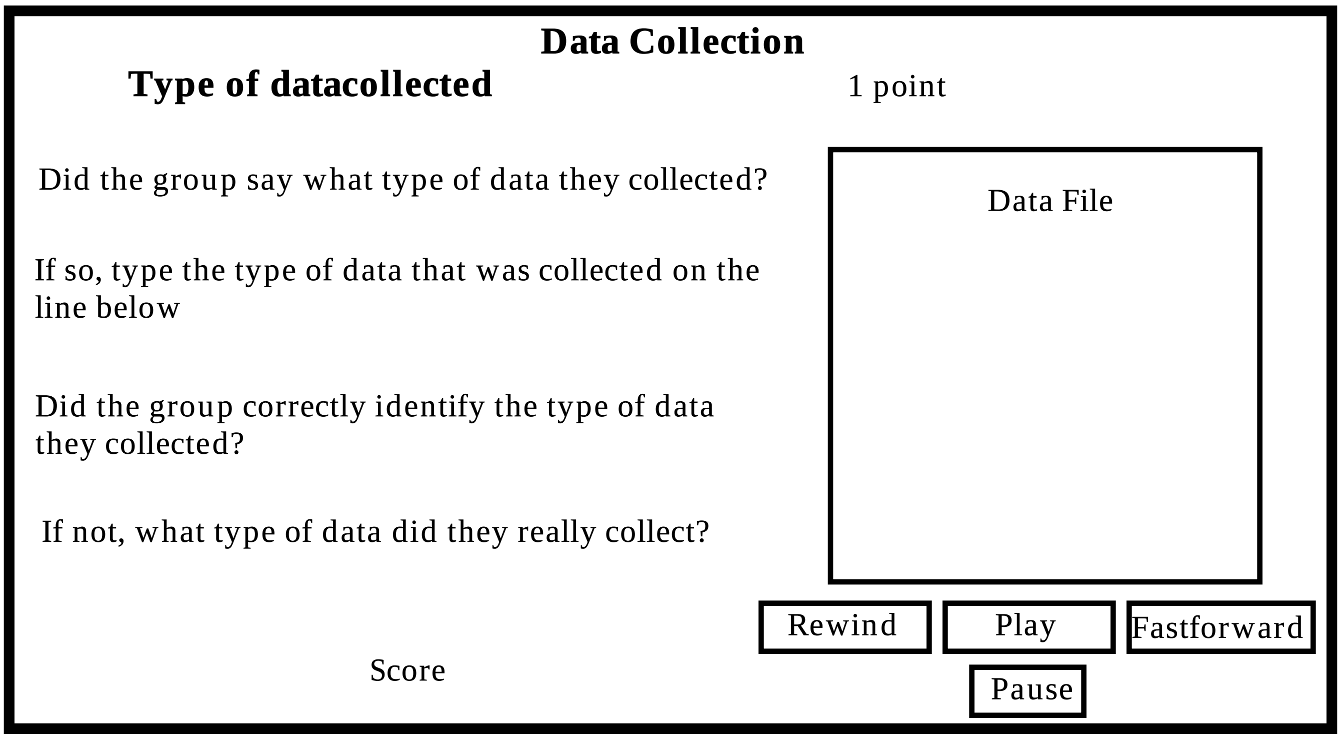

Figure 6: Example of critique component of *Critiquing Statistics*

CONCLUSION

The research described here provides some prototypes of statistical learning environments for middle school students. Three environments were described, and one of these was empirically evaluated. The *Authentic Statistics* environment proved successful in modeling performance standards for the statistical process of investigation. Furthermore, it helped students monitor and assess their own progress towards meeting these standards. Future studies should investigate the effectiveness of the other environments in terms of improving learning and assessment in statistics.

Acknowledgments

Preparation of this document was made possible through funding from the United States Office of Educational Research and Improvement, National Center for Research in Mathematical Sciences Education (NCRMSE), and the Social Sciences and Humanities Research Council, Canada. The author acknowledges the assistance of Nancy Lavigne for her contributions to this research, her continued Doctoral research in this area, and her editorial suggestions. I would also like to acknowledge the programming talents of Litsa Papathanasopoulou and André Renaud.

REFERENCES

Apple Computer Inc. (1989). *QuickTime™ version 1.5* [Computer program]. Cupertino, CA: Author.

Claris Corporation. (1991). *HyperCard™ version 2.1* [Computer program]. Santa Clara, CA: Author.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Erlbaum.

Diez, M. E., & Moon, C. J. (1992). What do we want students to know?... and other important questions. *Educational Leadership, 49,* 38-41.

Farallon Computing Inc. (1990). *Media Tracks: Version 1.0* [Computer program]. Emeryville, CA: Author.

Frederiksen, J. R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher, 18*, 27-32.

Kelley, H., & Thibaut, J. (1954). Experimental studies of group problem solving and process. In G. Lindzey (Ed.), *Handbook of social psychology* (pp. 735-785). Cambridge, MA: Addison-Wesley.

Lajoie, S. P., Lavigne, N. C., & Lawless, J. (1993, April). *The use of HyperCard™ for facilitating assessment: A library of exemplars for verifying statistical concepts.* Poster session at the annual meeting of the American Educational Research Association, Atlanta.

Lajoie, S. P., Lavigne, N. C., Munsie, S. D., & Wilkie, T. V. (in press). Monitoring student progress in statistical comprehension and skill. In S. P. Lajoie (Ed.), *Reflections on statistics: Agendas for learning, teaching, and assessment in K-12* . Mahwah, NJ: Erlbaum.

Lavigne, N. C. (1994). *Authentic assessment: A library Of exemplars for enhancing statistics performance.* Unpublished master's thesis, McGill University, Montreal, Quebec, Canada.

Lavigne, N. C., & Lajoie, S. P. (1996). Communicating performance standards to students through technology. *The Mathematics Teacher, 89*(1), 66-69.

Microsoft Corporation. (1992). *Microsoft EXCEL Version 4.0* [Computer program]. Redmond, WA: Author.

National Council of Teachers of Mathematics. (1995). *Assessment standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: Author.

Palinscar, A. S., & Brown, A. (1984). Reciprocal teaching of comprehension-fostering and comprehension monitoring activities. *Cognition and Instruction, 1*, 117-175.

Romberg, T. A. (in press). What is research in mathematics education and what are its results? *Proceedings of the ICMI Study at the University of Maryland.* College Park, MD.