IMPLEMENTING A COMPUTERIZED TUTOR IN A STATISTICAL REASONING COURSE: GETTING THE BIG PICTURE ®

Oded Meyer and Marsha Lovett
Carnegie Mellon University, USA

Many schools, like Carnegie Mellon University, are now teaching introductory statistical reasoning courses in a way that emphasizes conceptual understanding of the basic ideas of data analysis. There are several challenges in teaching such a course; foremost among them is the difficulty of conveying a sense of the "Big Picture." This paper describes a computerized learning tool that we have developed to help overcome this obstacle. This tool is a cognitive tutor *in which students solve data-analysis problems and receive individually tailored feedback. We discuss our cognitive tutor's use in the course and its measured effectiveness in a controlled experiment.*

INTRODUCTION

Teaching Statistical Reasoning at Carnegie Mellon University

Introductory statistics courses are being taught to a large and broad audience of students (Gordon, 1995). In his 1998 presidential address, David Moore estimated that "hundreds of thousands" of students pass through a first course in statistics in the US each year (Moore, 1998) In Carnegie Mellon University (CMU) one such course is "Introduction to Statistical Reasoning" (36-201), which is a required course for all Humanities and Social Sciences students, as well as some other majors. There are roughly 450 students taking this course each year (Fall - 240, Spring - 200, Summer - 10), the vast majority being freshmen. For some of these students this course is "terminal" and is the only formal exposure to statistics. Others, depending on the their major, go on to take additional upper level courses. The course emphasizes conceptual and critical understanding of statistics and utilizes statistics software (MINITAB) to minimize computational mechanics. The main goal of the course is to help students "get" what statistics is all about, in other words, that students will see the "Big Picture" of statistics. By "Big Picture" we mean understanding the process of: (1) producing data (sampling data from a population and considering study design issues), (2) conducting exploratory analysis on the collected data, and (3) making inferences from the sample back to the population of interest. The course curriculum is organized into three roughly equal parts corresponding to the above.

Teaching Statistical Reasoning: Challenges

Many challenges arise when teaching an introductory statistical reasoning course. Freshmen tend to approach material with an authoritative view of knowledge recognizing only right or wrong answers. It is therefore very hard for the students to really "get" statistics, which embraces ideas such as variability and uncertainty. Another challenge is to overcome students' prior beliefs about statistics (Gal & Ginsburg, 1994). Many students' conception is that statistics is a boring, "plug-in numbers to a formula" kind of subject, and that it has no relevance to their life, studies, or future plans. Some students see statistics as another mathematics course, and project their math-phobia onto it. Our biggest challenge, though, is to convey to students a sense of the "Big Picture" as frequently as possible without them getting lost in the details.

How does our course currently try to convey the Big Picture?

Our statistical reasoning course has three components: lectures, labs and homework. The course meets twice a week for 50-minute lectures (200-240 students). Lectures typically cover a specific topic that can be viewed as a point along the path of the course, plus some "ε-neighborhood" around this point consisting of a short review of the previous lecture and a peek into what is coming ahead. There is little opportunity for conveying a bigger picture than that, except in the beginning of the course and in the transitions from one part of the course to the next.

The students split into smaller groups for weekly computer labs. Students work in pairs and go through a paper-version exercise in a guided environment in the sense that TAs are available to answer questions and are required to check students' answers to a pre-selected subset of questions. The labs allow for a somewhat bigger picture, but still mainly focus on providing practice on the current week's material. Moreover, in labs, students do not get to make choices regarding which analysis to choose, since the relevant MINITAB instructions are provided in the exercise itself. In addition, there is a weekly homework assignment, which again, is related almost exclusively to the current week's topics, and gives the needed practice but very little chance for synthesis.

HOW COGNITIVE TUTORS CAN HELP

The above discussion suggests that we need a "tool" that will engage students with statistical problems covering more than the current week's topic. Ideally, this tool would encourage students to use of variety of skills (e.g., from considering the study design to selecting appropriate analyses to drawing conclusions) and to apply these skills *in the context* of real-world data sets. We propose that cognitive tutors offer one way to do just that. The name "cognitive tutor" refers to a computerised learning environment whose design is based on *cognitive* principles and whose interaction with students is based on that of a (human) *tutor*—i.e., making comments when the student errs, answering questions about what to do next, and maintaining a low profile when the student is performing well.

What is a Cognitive Tutor?

A cognitive tutor is a computer system that has both (a) a problem-solving engine that gives it the capacity to generate step-by-step solutions and (b) an enriched interface that allows students to communicate their own step-by-step solutions. These two components enable the system to track students' problem-solving processes at a detailed level and offer individually tailored feedback and hints. That is, the student takes a step by interacting with the computer interface, and the problem-solving engine judges the appropriateness of that step in the current situation, responding (if necessary) based on its knowledge of what step *it* would have taken and why. In addition, by collecting a database of information on individual students' performance, the system can make inferences about students' states of knowledge and suggest additional exercises that could remediate any apparent gap(s).

Past research has shown that cognitive tutors in various domains, including algebra, geometry, and computer programming, have been effective. Both randomized experiments and field studies have shown that students working with a cognitive tutor learn more efficiently and/or show better scores on posttests, including standardized tests such as the SAT (e.g., Anderson, Conrad, & Corbett, 1989; Anderson, Corbett, Koedinger & Pelletier, 1995; Koedinger, Anderson, Hadley & Mark, 1997). There is a body of theory on which the cognitive tutor methodology is based. Because of space constraints here, however, suffice it to say that cognitive tutors work by allowing students to solve problems *on their own* and to receive help when needed to avoid getting lost or confused. This help, which comes from the structure of the tutor interface and the hints and feedback from the problem-solving engine, can be thought of as a mental *scaffolding* that supports students' knowledge as it is constructed through practice—just as a physical scaffolding supports a building as it is erected. This metaphor additionally suggests that, ideally, a cognitive tutor should include mechanisms for reducing the scaffolding when appropriate, so that students—just like the finished building—can stand on their own.

Building a Cognitive Tutor for Data Analysis

Before building a cognitive tutor for data analysis, it made sense to us to look for data that could help shape such an endeavor. There is much evidence that students have difficulty applying statistical concepts, in part because of competing prior conceptions (Shaughnessy, 1992; Garfield & Ahlgren, 1988). We added to this body of empirical work by further investigating where (and hopefully why) these difficulties arise in the context of solving data-analysis problems (see Lovett, 2001, for more details). We found that, even among students who had completed the above course with a grade of B or better:

- Students have difficulty choosing appropriate graphical displays and statistical tools (e.g., many of their analyses were inappropriate or at best not directly relevant to the question).

- Students often fail to interpret their results with respect to the question of interest (e.g., students would say they were finished with the problems almost immediately after producing a display or statistic).
- Generally described, students do not take a systematic approach to solving these problems (e.g., their behavior appeared instead to be driven by the menu options of the statistics package or by a random process-of-elimination strategy).

These results indicated that a cognitive tutor for data analysis could be quite beneficial if it helped students choose appropriate analyses, remember to draw conclusions from results (beyond just restating them), and use a conceptual structure to approach these problems.

Therefore, we designed a cognitive tutor called *SmartLab* (written in Java) to address these points. Most importantly, we tried to do so in a way that would facilitate students' learning of the "Big Picture". First, we designed *SmartLab* to highlight the structure that is common *across* data-analysis problems. This common structure consists of all the elements of the "Big Picture" and is repeated for each problem, regardless of its details. Second, many of the "headings" in this structure represent steps that were previously "hidden" from students' point of view in that they were covert, planning steps (e.g., identifying relevant variables and classifying them as to type). By revealing these steps, *SmartLab* both provides scaffolding to students in the steps that precede selecting an appropriate analysis, and it makes that planning process open to feedback (so students can see from *SmartLab*'s feedback where they are going wrong before they get too far down an erroneous path). Third, the provision of hints and feedback in *SmartLab* applies to all the steps of problem solving, so students can get quicker diagnosis of their errors than they would otherwise (e.g., using paper handouts in labs or on homeworks). Fourth, *SmartLab* was designed so that the scaffolding we offer to beginning students can gradually be faded away. We accomplish this in several stages, first by substituting fill-in-the-blanks for pull-down selectors and later by removing "hidden skill" headings from the structure. One additional point about *SmartLab* is that when students finish solving a problem, they have produced a well organized, printable report of their work.

USING AND TESTING OUR COGNITIVE TUTOR FOR DATA ANALYSIS

We have used *SmartLab* in two different venues to test its effectiveness, namely, in the classroom and in an experiment. We will discuss the results of each briefly below.

Use in the Classroom. We have used *SmartLab* in the past three semesters in several of the labs instead of the paper-version exercise, and asked students as well as TAs to provide technical and conceptual feedback on their experiences. Based on this valuable input we have made a series of technical and pedagogical refinements to our tool, and we are currently working on including the use of *SmartLab* in homework assignments (a web-based Javascript version). It should be noted that *SmartLab* has an extremely positive impact on the interaction between TAs and students in the lab. In the paper-version exercises TAs were required to check the students' answers to the more technical questions, making sure that students were on the right track. Using *SmartLab*, students get tailored feedback from the system for those questions, and thus the interaction between the TAs and the students can be shifted to the interpretational questions. In this new role, TAs are more engaged with the students and can get a better feel for their understanding of the material.

Use in an Experiment. To evaluate *SmartLab* in a more controlled environment, we designed an experiment in which people could get a fairly intensive statistics experience in a relatively short amount of time. Participants without any prior (formal) statistics training were recruited for pay. They were asked to attend five sessions, for two to three hours per session, and were assigned to work with *SmartLab* when solving problems. (This experiment also involved a comparison with a variant of *SmartLab*, which is outside the scope of this paper.) For sessions 1-4, the participants watched videotaped lectures and worked on sequences of problems. In addition, at the beginning of session 1 and during session 5, participants completed several paper-and-pencil tests. Moreover, during session 5, participants worked on additional, open-ended data-analysis problems on the computer with minimal scaffolding.

One paper and pencil assessment was a multiple-choice test covering the basic skills and concepts of exploratory data analysis, including questions on identifying study designs, selecting appropriate analyses, and drawing conclusions from the results. Participants' scores increased by 23%, a significant improvement, $t(19)=5.877$, $p < .001$. Another paper and pencil assessment asked students to read through short, data-analysis situations and classify them into groups (on whatever basis they felt reasonable). By analyzing participants' categories, we found that there was a significant pre-post shift in the way participants classified the problem: before the experiment, they tended to base their classifications on the subject matter of the problems and, after the experiment, they tended to base their classifications on the appropriate exploratory analysis, $t(19) = 4.11$, $p < .001$.

We also looked at participants' performance on the open-ended quiz problems. An interesting comparison here is that, after using *SmartLab*, participants in the experiment made only 0.73 errors on average per opportunity to select an appropriate analysis, whereas in Lovett (2001) students who had taken an entire semester's course made more than 9 errors per selection opportunity (with the appropriate analysis being, on average, the 6^{th} selected).

CONCLUSIONS AND FUTURE WORK

This paper discusses the design and implementation of a cognitive tutor for data analysis called *SmartLab.* Each time students use *SmartLab*, they get exposed to the Big Picture of data analysis and see that the same process applies across all problems. In particular, *SmartLab* puts an emphasis on helping students learn how to choose the appropriate analysis and requires them to draw conclusions in context. We found that it works well in the computer lab sessions of our course. By adding this tool, we find that TAs can be released from attending to the details of students' solutions and can focus then on the deeper issues. Moreover, a controlled experiment showed that, even over short period of use, this tool led to significant improvements in students' approach to exploratory data analysis. In particular, results suggested that after the experiment, when participants encountered a new problem, they were thinking about it in terms of the appropriate analysis instead of in terms of the subject matter.

Our future plans include extending the use of *SmartLab* in the classroom (to more labs throughout the semester and to homeworks). We also will conduct more experiments that compare *SmartLab* with control conditions representative of current typical lab experience (e.g., paper versions). Also, because our focus thus far has been to refine SmartLab pedagogically, we have plans to collect additional data to help assess the usability of this tool in terms of human-computer interaction and to make refinements accordingly.

REFERENCES

Anderson, J.R., Conrad, F.G., & Corbett, A.T. (1989). Skill acquisition and the LISP tutor. *Cognitive Science, 13,* 467-505.

Anderson, J.R., Corbett, A.T., Koedinger, K.R., & Pelletier, R. (1995). Cognitive tutors: lessons learned. *Journal of the Learning Sciences, 4,* 167-207.

Gal, I., & Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: Towards an assessment framework. *Journal of Statistics Education, 2*(2).

Garfield, J., & Ahlgren, A. (1988). Difficulties in learning basic concepts in statistics: Implications for research. *Journal for Research in Mathematics Education, 19,* 44-63.

Gordon, S. (1995). A theoretical approach to understanding learners of statistics. *Journal of Statistics Education, 3*(3).

Koedinger, K.R., Anderson, J.R., Hadley, W.H., & Mark, M.A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8,* 30-43.

Lovett, M.C. (2001). A collaborative convergence on studying reasoning processes: A case study in statistics. In S. Carver and D. Klahr (Eds.), *Cognition and instruction: twenty-five years of progress* (pp. 347-384). Mahwah, NJ: Erlbaum.

Moore, D.S. (1998). Statistics among the liberal arts. *Journal of the American Statistical Association, 93,* 1253-1259.

Shaughnessy, J.M. (1992). Research in probability and statistics: Reflections and directions. In D.A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465-494). NewYork: Macmillan.