

## Evaluation of an Interactive Tutor on Sampling

Prof. Charles L. Cappell

Paper presented at the annual meetings of the American Educational Research Association,  
Chicago, IL., April, 2003

### Abstract

Through support from the National Science Foundation (DUE-9752619), I have been developing a computer-based, interactive tutor teaching the fundamentals of sampling. Several principles of cognitive learning and instructional design are implemented in the design: contextualized (social science) learning, cognitive mapping for relational learning, building of formal-operational understanding from concrete imagery, self-directed learning, and corrective, directive, and evaluative feedback. The tutor runs on PC platform and was developed using Macromedia Authorware/Director software, along with several auxiliary programs, such as MathCad.

The effectiveness of the first fully operational version of the program was evaluated by using an experimental design involving two sections (approximately 65 students) and two faculty. Within each section, half the students were assigned to study sampling concepts and procedures using the Sampling Tutor without attending lectures. Two different lecture formats were used: one instructor delivered his usual set of lectures based on his own material, a second instructor delivered lectures using the projected content from the Sampling Tutor.

Analysis of scores on a overall quiz assessing the student's understanding of sampling concepts were mixed and interacted with the section. A nested covariance analysis revealed that the Sampling Tutor group outperformed the Lecture group without Sampling Tutor content or examples, except for students with the lowest levels of GPA. The pattern of effects contrasting the Lecture group exposed to the Sampling Tutor content versus Sampling Tutor group revealed that the Sampling Tutor works as well or slightly better among the students with the lowest GPAs.

At the panel presentation of this conference paper, the latest version of the Sampling Tutor is demonstrated and the results of this initial evaluation are presented.

### *Introduction: Statistical and Quantitative Reasoning Learning Objectives and Associated Hurdles*

The general approach implemented in the development of this interactive tutor consisted of identifying the content and operational learning objectives and deciphering the cognitive based hurdles that impede that learning. Research on mathematical and statistical education has identified several cognitive transitions, stages, or 'reorganizations' of thinking needed to develop quantitative reasoning. In most general terms, in a review of its curricula, the mathematical profession identified four major cognitive transitions leading to "mathematical reasoning":

1. The transition from arithmetic to algebra: culminating in the ability to think symbolically and to manipulate symbols to solve problems.
2. The transition between algebra and geometry: culminating in the ability to think logically. Logical proof replaces calculation.
3. The transition from technique to theory: culminating in the ability to create proofs. Thinking shifts from a focus on solving a problem to creation of proofs of a more general nature.
4. The creative transition culminating in the ability to pose new problems needing solution. (American Association of Colleges, 1991: Ch. 5: Mathematics).

This outline of major cognitive transitions in mathematical reasoning focuses our attention on the formal stages of development and suggests possible hurdles that may hinder the transition from lower to higher forms of mathematical reasoning among students.

With broader objectives in mind, Jones et al. (1995) used the Delphi method to survey a panel of experts challenged to identify the dimensions of critical thinking. Their results revealed a separate dimension of critical thinking comprising "Statistical Argumentation Skills." These attributes, with their Delphi based factor loadings in parenthesis, are:

- “Employ graphs, diagrams, hierarchical trees, matrices, and models (.77).
- Apply appropriate statistical inference techniques (.69).
- Assess statistical information used as evidence in an argument (.65).
- Use multiple strategies in solving problems (.55).

Determine if a conclusion is based on an adequate/representative sample (.54).  
Determine how data might confirm or challenge a conclusion (.53).  
Assess how well an argument anticipates and responds to objections (.48)  
Determine and judge the strength of a causal reasoning argument (.41).”  
(Jones et al., 1995)

Statistical educators also have identified detailed cognitive skills, learning objectives, and some conceptual hurdles accompanying forms of statistical reasoning. Misleading intuitive notions about probability have been identified as a major stumbling blocks in people’s ability to correctly assess the probabilities of various events (Tversky and Khaneman, 1982; Konold, Clifford, Alexander Pollatsek, Arnold Well, Jill Hendrickson, and Abigail Lipson, 1991; Schloz, 1991).

Chance (2000) compares and contrasts statistical literacy, reasoning, and thinking in order to identify the skills that need to be taught and learned to achieve a “wider view” of statistics (Wild, 1994), namely a deeper understanding with elements of creativity, skepticism, and an ability to transfer knowledge to solve new problems. Garfield’s (1995) concise review of research in educational statistics highlights that understanding of statistical reasoning and related operations are positive functions of the levels of hands-on activity, of the amount of corrective feedback, use of computer simulations, and analysis of real data using statistical software.

Over the last 20 years I have been teaching courses on quantitative methods to undergraduate social science students. As my courses have evolved, I have tried to integrate pedagogical and cognitive research in their design and delivery. The broad learning objectives of the course include:

1. Statistical Critical Thinking Skills: the ability to think logically and to assess the quality of evidence, to identify logical fallacies and well-posed questions amenable to scientific inquiry, to assess the intuitive processes of collecting information by comparing it to the processes used in scientific sampling.
2. Formal, Symbolic Thinking Skills. Translate text and ideas into symbols, read algebra, create and manipulate symbolic models, move from concrete to formal-operational reasoning.
3. Conceptual/Variable/Hypothetical Language (CVHL). Think in terms of hypothetical implications, scientific method, causality, use concepts of counterfactuals, sample space.
4. Stochastic Reasoning Skills. Understand probability, variation, uncertainty, sample space.
5. Statistical Modeling Skills. Create and evaluate statistical models representing a current condition or process. Apply useful statistical summary measures to descriptions and explanations: correlation and regression coefficients, conditional probabilities, odds, summary measures of association.
6. Inferential Reasoning. Think inferentially, to recognize the limits of generalization. The basic gamut of inferential statistics, sampling distributions, confidence intervals, p-values, hypothesis testing. (see also: AAC, Ch. 5, 1991; Browne, N., and Keeley, S., 2001; Chance, 2000; Jones, 1995; Mallows, 1998).

### ***Cognitive Learning Principles***

Recent research in cognitive science has produced several models of how complex symbolic material is learned. The interactive tutors I am developing incorporate several learning strategies as part of the software design, including those discussed specifically in the context of statistical education (Cappell, 2001a, 2001b; Lovett and Greenhouse, 2000). The Sampling Tutor developed as the first module for my Quantitative Reasoning Project is designed to aid “conceptual learning”. The focus is on generating a ‘deep’ understanding of sampling concepts and procedures by an explication of the concepts and their relations (Fauconnier, 1997; Schau and Mattern, 1997). But the content also contains symbolic notation, so the tutors I’m developing should not be perceived as though ‘conceptual’ learning of sampling contains no symbolic notation and is antithetical to computation (Huberty, Dresden, and Bak, 1991). Inference is not included in the Sampling Tutor; a separate tutor is being designed for that topic.

The tutorial is only one of three delivery systems used in the course. It is meant to supplement, perhaps replace, traditional lecturing. It certainly can supplement, if not replace traditional textbooks. It is also interactive, so intermittent assessments of student understanding can be fed back to students. The second delivery system involves hands-on exercises with real data. The course is laboratory based, each lab session has 24 students with two instructors supervising the exercise. Each week a lab based exercise reinforces the learning objectives of the conceptual module being covered. The third delivery system is a

either a lecture or discussion section in a traditional classroom setting, but with computer displays so problems can be worked out or conceptual material presented.

The Sampling Tutor evaluated in this paper contained an architecture that was informed by some of the basic principles of cognitive learning. Earlier papers discuss in detail how some of these principles were incorporated, (Cappell et al., 2000; Cappell, 2001a, 2001b; Haapoja, 2001). Here I just briefly review the major cognitive learning mechanisms I tried to incorporate in the overall architecture:

1. Attention Mobilization: Goal-based learning motivates students by posing problems or arousing curiosity, providing a concrete substantive context to engage the learner. The Sampling Tutor tries to motivate learning by grounding the training in substantive material related to the sociology major (Context Learning).
2. Cognitive Mapping and Schema: Reasoning strategies are learned models built from linked concepts. A learned mode of reasoning invokes a cognitive map of related concepts. The Sampling Tutor aims for students to develop cognitive maps, literally grow neural networks, that relate concepts to one another generating a schema for a fuller understanding of material (Relational Learning, creating Schema).
3. Review and Repetition: Material must be moved from working memory to long-term memory. The Sampling Tutor has hypertext definitions of all key concepts as well as a glossary.
4. Transferred Learning: Context specific to formal-operational general problem solving. The ultimate learning objective is to have the material become part of the learner's active knowledge base, transferable to new problem settings. This involves a major mathematical-statistical transition from concrete to formal reasoning. Learning begins with the concrete, specific experiences and progresses to formal stages. Achieving generalized/transferred learning, i.e. an understanding that can be applied to novel problems and situations requires an understanding of more abstract concepts, stripped of the specific substantive meaning used to illustrate them. The Sampling Tutor incorporates this objective by explicitly linking recognizable, concrete images and building the formal definitions from them.
5. Chunking: Information needs to be presented in digestible amounts so it can be effectively processed and integrated in the learner's knowledge base. The Sampling Tutor is 'chunked', containing different modules on different topics;
6. Self-Paced Learning: the Sampling Tutor allows the student to move as slowly or as quickly as needed through the material.
7. Performance Learning: Linking conceptual learning to techniques and more formal reasoning skills. The Sampling Tutor aims to have the student see how the abstract idea emerges from one concrete realization of that idea. The recitation quiz questions used at the end of the current Sampling Tutorial include some harder, deeper understanding questions calling for students to apply their knowledge and not just recall it.
8. Self-Directed Learning: the Sampling Tutor does not force the student to follow any one path through the material, although it is set up so that one can follow a scaffolding path that facilitates hierarchical learning. The Tutor also contains simulations that the student must activate.
9. Corrective Feedback: near immediate feedback that increases the probability that the material presented is placed in long-term memory (Interactive Tutoring). The Sampling tutor has short review quizzes after a major conceptual module to test if the student has mastered understanding of the meaning of that concept, provides some corrective clue to correct their thinking, and gives them the opportunity to immediately review the material in the lesson about which they were just quizzed.
10. Metacognitive Feedback: Learning is enhanced when learners become conscience of the learning strategies and styles that they are using. Corrective, directive, and evaluative feedback can increase the student's awareness of whether their learning strategies are successful.

The evaluation of the Sampling Tutor discussed in this paper is a global evaluation of the Tutor, comparing the level of learning obtained by students after one week of access to the Sampling Tutor to those who were exposed to three hours of lecture covering the same material. Some specific aspects of the Sampling Tutor that were used in the assessment reported in this paper are discussed below. This tutor was written using MACROMEDIA'S AUTHORWARE/DIRECTOR platform providing multimedia presentations, interactive tutorial feedback and evaluation, and tracking of student learning patterns and performance. MATHCAD, various statistical software packages – primarily SAS, ADOBE PHOTOSHOP,

and MS OFFICE components are used to prepare content.

### Sampling Tutor Content

The table of contents illustrates the major concepts covered in the Sampling Tutor. While meant to be studied in a sequential order, the learner has full control and can navigate to any lesson in any order. The navigational menu used in the sampling tutorial emphasizes the relationship among the lessons. Each lesson is designed to be a short to moderate length coherent chunk of material. A short interactive quiz that corrects misconceptions and allows the user to review the material appears at the end of each sub-lesson.

Learning is intended to begin with the distinction between the population and sample, parameters and statistics. Within the Fundamental Sampling Concepts lesson, the key concepts of randomness, random selection, sample space, and selection with and without replacement are covered. Combinations and permutation counting formulas are introduced to demonstrate the size of various sample spaces for different designs. The lesson on sampling designs shows the nature and steps of creating systematic, cluster, stratified, complex, and non-probability based samples. And lastly, a short transition lesson on inference is included. The Sampling Tutor evaluated in this paper had only a set of short interactive quizzes at the end of each lesson where the student could navigate to the relevant content if responding incorrectly. The current version now has a Recitation, a comprehensive quiz with corrective and metacognitive feedback.

### Interactive Tutorial Architecture

The Sampling Tutor tried to implement several features of cognitive learning strategies as part of the tutorial's architecture, such as using a conceptual map as the main navigation menu. A glossary is always active, and key concepts are colored coded blue throughout a lesson to indicate that a quick definition is available as a pop-up.

The key concept of "sample space" is approached in several lessons and reinforced. Concrete experientially based material is used and then the formal definitions are introduced, including material on combinations and permutations.

The Sampling Tutor focuses on the major concepts of sample space, sampling variability, and sampling design.

The Sampling Tutor also contains animation to illustrate concepts as well as some simulations. Students can create multiple samples from a simulated population to demonstrate the concept of sampling

error or variability. The example used is based on the homicide rates of the 25 largest U.S. cities, since a large proportion of students taking this course are sociology/criminology majors. The same ‘population’ is used in other sections of the Sampling Tutor as well, although other contexts are introduced for variety. The interactive simulation built into the Sampling Tutor demonstrates to the student how a single sample is just one realization from the full set of possible samples that comprise the Sample Space. The section on complex sampling designs uses the National Criminal Victimization Survey to illustrate multi-stage sample designs.

**Fundamental Sampling Concepts** **Simple Random Sampling**

Sample Frame		
43.1 - Baltimore, MD	52.9 - Detroit, MI	15.5 - Jacksonville, FL
15.3 - Boston, MA	70.0 - Dist of Columbia	14.0 - Las Vegas, NV
16.5 - Charlotte, NC	7.0 - El Paso, TX	23.8 - Los Angeles, CA
33.1 - Chicago, IL	4.0 - Honolulu, HI	25.3 - Memphis, TN
15.4 - Columbus, OH	21.3 - Houston, TX	22.1 - Milwaukee, WI
27.8 - Dallas, TX	14.0 - Indianapolis, IN	21.3 - New York, NY
12.8 - Seattle, WA		25.9 - Philadelphia, PA
		21.5 - Phoenix, AZ
		19.4 - San Antonio, TX
		9.7 - San Diego, CA
		12.3 - San Francisco, CA
		4.0 - San Jose, CA

Sample 1	Sample 2	Sample 3
Memphis, TN - 25.3	Las Vegas, NV - 14	Los Angeles, CA - 23.8
El Paso, TX - 7.8	New York, NY - 21.3	Baltimore, MD - 43.1
Baltimore, MD - 43.1	Charlotte, NC - 16.5	Las Vegas, NV - 14
Boston, MA - 15.3	Boston, MA - 15.3	San Francisco, CA - 12.3
Seattle, WA - 12.8	Dallas, TX - 27.8	San Diego, CA - 9.7
The mean of this sample is 20.86	The mean of this sample is 18.98	The mean of this sample is 20.58

The mean of the three means is 20.14

The population mean is 21.96

Click on the notepad to take notes.

Help

Main Menu

Press the Help button for instructions

Glossary

Calculator

Exit Program

Sampling Concepts

Back

Forward

The current version of the Sampling Tutor demonstrated at the AERA 2003 conference contains some additional features not present in the version whose evaluation is reported in this paper: more interactivity, animation, simulations, audio, and an overall Recitation module - an interactive quiz giving feedback about the learner’s performance. Building more explicitly on the idea of “concept mapping” (Fauconnier, 1997; Schau and Mattern, 1997), at the beginning of each major lesson, audio and animated graphics highlight the learning objectives for that lesson. At the conclusion of the lesson, an animated conceptual map with audio is created to review, in relational images, the material which the student needs to integrate. The current version of the Sampling Tutor demonstrated as part of the presentation at the AERA conference includes a “Recitation” module, a randomized set of approximately 75 questions, allowing the student to repeat the recitation several times. As they take the Recitation, students are given substantive feedback on their answers, and in the easier questions, an opportunity to navigate to the page of the lesson where the material covered in the question is presented. At the finish of the Recitation, some metacognitive feedback is presented regarding the level of understanding the student has achieved. The questions embedded in the tutorial are classified according to the type of cognitive operation required to answer correctly: simple recall of conceptual definitions, relational learning that shows appropriate links have been made among concepts, and more operational learning which indicates students can transfer the knowledge learned to new problems. This is intended to increase the student’s consciousness of their own learning levels and to encourage them to reflect on their learning strategies. These improvements were added, in large part, from reviewing the evaluations of the Sampling Tutor received from students who responded to open-ended questions evaluating the Tutor’s design and from direct observations of how students used the Sampling Tutor in the laboratory setting (Haapoja, 2001).

### ***Evaluation of the Sampling Tutor***

The evaluation reported in this paper compares student performances on a sampling quiz after some students studied sampling concepts from an early version of the Sampling Tutor which included text enhanced with graphical imagery, interactive quizzes with feedback, a glossary, student directed navigation, and simulations, while other students attended two different lecture sections.

Two sections ( $n_s=25$ ,  $n_c=38$ ) of a three hundred level course teaching quantitative methods to sociology students were used to evaluate the Sampling Tutor. One section was taught by the author and the developer of Sampling Tutor, “C”. The second section was taught by an experienced professor who made use of his own lecture material, “S”. Both instructors had over 20 years of experience in teaching a variety of statistics and methodology courses. Half of the students in each section were randomly assigned to use the Sampling Tutor in lieu of attending lectures. These students attended computer laboratory sessions where they were given a brief introduction to the Tutor and told they could use the Tutor during

and outside of class time. The lecture section, “C”, taught by the developer, used some of the Sampling Tutor content projected onto a screen, much as one would use a power point presentation. The lecture format allowed interruptions and questions, although there were very few of these. Most students sat passively and listened to the material; few took notes. Three class periods were used to cover the material, about three hours total. The second instructor, “S”, taught a random half of his students in three lectures using his own lecture material enhanced by several visual aids often used in sampling lectures, but using none of the visual aids contained in the Sampling Tutor. Thus, the experimental design employed created different treatments nested within instructor: four groups, two each nested within each instructor, “C” versus “S”: Lecture mode versus Tutor mode.

Both instructors collaborated in designing a post-test sampling quiz (25 points) based upon a few verbatim sampling questions found in the instructional manuals of widely used social science research texts as well as questions more directly related to the material in the Sampling Tutor and presented in both lectures. We tried to create questions that were not content specific to any of the modes of presentation, but rather emphasized conceptual understanding that could be transferred to parallel, but different problems and questions than those covered in lectures and by the Tutor. The student’s previous cumulative GPA was used as a statistical control.

The analysis of the final sampling quiz scores is used to make an overall global evaluation of how the Sampling Tutor performed compared to the two different lecture sections. The contrast comparisons tested will reveal rather gross effects of the Tutor’s effectiveness or ineffectiveness. The comparison between “C” Lecture and Tutor sections assesses the difference between a “live” performance using the same content and a self-paced, learner-directed session. The lecture format also covered all of the topics presented in the Tutor; however students working independently with the Sampling Tutor may well have skipped sections or proceeded hastily. The contrast between “S” lecture and Tutor sections assesses the difference between a thoroughly thought-out presentation with clear connections between sub-lessons supported by relevant graphical illustrations delivered at the student’s pace to the traditional lecture format.

We first note the low overall performance level on this test, overall mean = 9.24. While one of the reasons for this low performance may be using parallel test items rather than items explicitly linked to the context and content of the review quiz questions in the Sampling Tutor, this result is not too far below the range of 50% mastery of material reported elsewhere (e.g., at a tier 1 university) when students are evaluated on their level of comprehension of statistical material using closed book and note formats (Lovett and Greenhouse, 2000).

The experimental design is nested within section (instructor): the comparison is between the two different lecture sections and the Sampling Tutor sessions. The student’s cumulative GPA is entered as a covariate to control for the student’s overall past academic performance. If statistically warranted, examining the interaction effect with student’s GPA can also reveal how the Sampling Tutor performs for students with varying levels of previous academic success.

Table 1 presents the results of a nested analysis of covariance, indicating that the effects of the Sampling Tutor have to be estimated separately for each instructor and at different levels of GPA. The presence of the covariance interaction term also means that no direct interpretation can be given to the differences across the experimental variable: the use of the Tutor v. Lecture, since this effect is complicated by the interaction within each section.

Table 1: Results of Nested Covariance Design for Tutor Experiment

Source	DF	Type I SS	Mean Square	F Value	Pr > F
CLASSID	1	2.350	2.350	0.35	0.559
TREATMNT(CLASSID)	2	13.921	6.961	1.02	0.366
GPA*TREATMNT(CLASSID)	4	115.225	28.806	4.24	0.005

The coefficients estimated by the nested covariance model are presented in Table 2. The ClassID, or section, effect shows that the students in section “C” performed slightly better overall, about 2 points higher than those in section “S”. The Treatment within ClassID effects can not be interpreted directly because of the interaction with GPA, the covariate. The effects of GPA can be interpreted to indicate how large of an effect previous academic performance has on mastering sampling concepts. We can see that all

of the effects are positive, but only one effect would be considered statistically significant, the impact of GPA within the lecture section “C”, that section taught by the developer of the Sampling Tutor using Sampling Tutor content projected in lecture. This estimate indicates that the steepest discrepancies in the learning of sampling concepts across GPA levels occurred in the section exposed to the lecturing format that used Sampling Tutor content. Students with lower levels of academic achievement apparently were not able to process oral information even when accompanied by visual aids. This may be attributed to poor ability at note taking, lack of concentration, and cognitive overload. It is notoriously difficult to master unfamiliar statistical or quantitative material in a fast-paced lecture mode. However, the beta coefficient underscores that as students’ academic abilities increase, they have less difficulty processing the material. The lowest coefficient for GPA was found under the lecture condition within section “S”. Here the effect of student ability in learning is nearly flat. Given that this section also has a lower level of performance, this group appears the least successful within any learning mode. We note that one indication of a successful educational intervention is to raise the overall level of performance and to attenuate the effect of some latent overall student ability; in other words, to make it possible for students of all levels of learning rates, with the proper motivation, to achieve nearly equal learning levels.

Table 2: Predictive Equations

TREAT-CLASSID	Overall Intercept	ClassID Intercept	Treatment Within ClassID	GPA Effect within Treatment (stan. error)
Lecture 1 “S”	6.90	-1.99	3.54	0.21 (1.08)
Tutor 1 “S”	6.90	-1.99	0	1.63 (1.36)
Lecture 2 “C”	6.90	0	-5.75	3.31 (0.85)
Tutor 2 “C”	6.90	0	0	0.79 (1.26)

Of primary interest is the effect of the Sampling Tutor compared to the lecture formats. Table 3 presents the estimated non-standardized effects for each of the experimental effects at 5 separate levels of student GPA. The Sampling Tutor did as well or better (for higher GPA students) within instructor “S”’s section. Instructor “C” outperformed the Sampling Tutor among higher GPA students, but the Sampling Tutor did as well or better among students at the lowest GPA level, 2.0.

Table 3: Estimation of Tutor-Lecture Effect at 5 GPA Levels

Parameter	Estimate	Standard Error	t-value	Pr >  t	Effect Sizes <sup>1</sup>
posttest “S” tutor-lecture at 2.0	-0.705	1.409	-0.50	0.62	-.31
posttest “S” tutor-lecture at 2.5	0.005	1.067	0.00	0.99	
posttest “S” tutor-lecture at 3.0	0.715	1.341	0.53	0.60	.32
posttest “S” tutor-lecture at 3.5	1.425	1.992	0.72	0.48	
posttest “S” tutor-lecture at 4.0	2.135	2.764	0.77	0.44	.94
posttest “C” tutor-lecture at 2.0	0.713	1.143	0.62	0.54	.22
posttest “C” tutor-lecture at 2.5	-0.546	0.868	-0.63	0.53	
posttest “C” tutor-lecture at 3.0	-1.806	1.164	-1.55	0.13	-.56
posttest “C” tutor-lecture at 3.5	-3.065	1.764	-1.74	0.09	
posttest “C” tutor-lecture at 4.0	-4.325	2.455	-1.76	0.08	-1.35

<sup>1</sup> Effect sizes are estimated as Cohen D statistics using pooled variances within each section: S and C.

Effect size scores are reported in Table 3 in order to summarize the magnitude of the effects observed. These effect scores are calculated as the Cohen D statistic, a pooled variance estimate within each section was used to calibrate the effect sizes. We note that the effect of using the Tutor among students in section “S” with lower GPAs, e.g. 2.0, was to lower performance by a third of a standard deviation. This is contrary to results needed to justify Tutor usage among this group. But among students with higher levels of GPA, the gains from the Tutor are nearly one standard deviation. Among section “C” students, the Tutor has the desired intervention effect, raising the performance of the lowest GPA students by nearly a quarter of standard deviation compared to their lecture counterparts. And as previously noted,

the slope of the GPA effect is smaller for section “C” students using the Tutor than that estimated for students exposed to the lecture.

The pattern of expected results can be seen clearly in Table 4, where the estimated performance on the post-test is calculated under varying treatments and at different levels of student GPA. Even as the Tutor did as well as or better than the instructor among students with GPA at 2.5 or lower, one must note the relatively small gains as well as the overall low performance. However, the Sampling Tutor produced more leveled results with section “C”. This seems to indicate that students with higher levels of academic performance are better acclimated to the lecture format and are able to carry away more knowledge from the “live” lecture than students with lower levels of academic achievement, where the Tutor operates more effectively.

Table 4: Least Squares Predicted Values of Posttest Score at 5 GPA Levels

TREAT-CLASSID	Least Square Means of Post-test Estimated at:				
	GPA=4.0	GPA=3.5	GPA=3.0	GPA=2.5	GPA=2.0
Lecture 1 “S”	9.30	9.20	9.09	8.98	8.88
Tutor 1 “S”	11.44	10.62	9.81	8.99	8.17
Lecture 2 “C”	14.37	12.72	11.06	9.41	7.76
Tutor 2 “C”	10.05	9.65	9.26	8.87	8.47

Reflect on the image of experimental results that would appear if a successful learning intervention was present. All students exposed to the intervention should perform better than those not exposed, and any interactions with some latent measure of student’s learning ability or rates should flatten out. If this were the case, the plots of the GPA effects and experimental interactions would show flatter linear relations within the Tutor sections than in the lecture sections, and overall higher levels of performance within the Tutor sections would be evident. The graph in Figure 1 shows the positive effects of the Tutor within section “C”, the major statistically significant finding. Note that within section “S”, the effects of the Tutor increase slightly as student academic ability (previous GPA) increases. While this is not in general a pedagogically desirable experimental effect, given the low level of performance among students in section “S”, this result shows that students with higher GPAs can increase their learning above their higher GPA counterparts who were not exposed to the Tutor.

The results for section “C” are more consistent with the type of findings consistent with successful educational interventions. The gap in the Tutor-Lecture performance is smaller at the lower GPA levels, and the tutor group actually outperforms the lecture group among students with GPAs at 2.0 and lower. By using the Tutor as a complement to lecture rather than as a substitute, these results suggest that the Tutor can help achieve higher levels of learning across students of varying GPA.

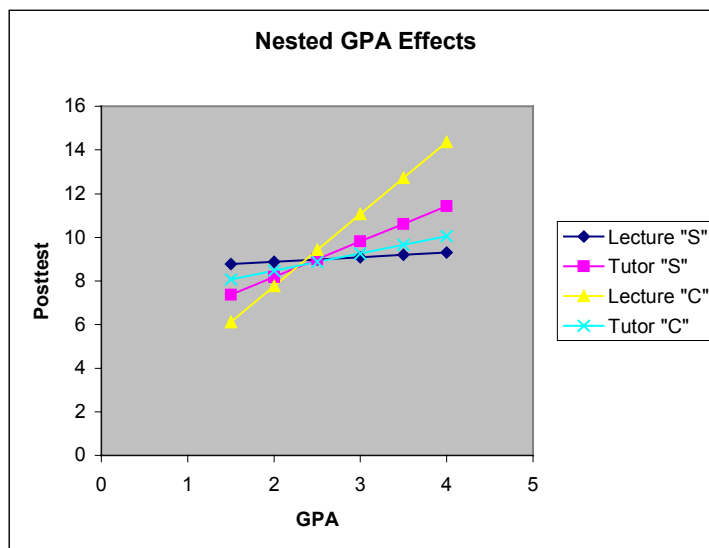


Figure 1: Plot of Expected Mean Post Test as a Function of GPA by Treatment Nested within Class

### Discussion

Considerable time and effort are involved in developing these types of interactive tutors. The modest gains in learning demonstrated by this initial evaluation of the Sampling Tutor may make these efforts highly inefficient at the outset. Results seem to indicate that this style of instructional architecture may be beneficial to students with lower levels of academic ability who are willing to spend sufficient time often not available in mass education general methodology courses. The findings also suggest that designs should incorporate more forced remediation to help these students. The interaction effects between GPA, experimental use of the Tutor, and instructor indicate the complex pattern of contextual factors that influence student learning. It may highlight the important non-cognitive aspect of motivating students to learn, regardless of the mode of delivery.

The Sampling Tutor was designed to serve as one of three prongs of instruction in this methods class, to supplement or replace only the traditional lecture mode of presenting information. While the results of this experiment were modest, one should note that the global evaluation conducted assessed the stand alone capacity of the Tutor. When used in conjunction with an active learning, laboratory based exercise and a problem review session, we speculate that the Tutor will provide a valuable component facilitating student learning.

Of course, student motivation for learning and an informed metacognitive orientation about how they are learning are needed ingredients as well. An evaluation of the current Sampling Tutor is being conducted that will assess whether some of the cognitive learning features available in the Tutor are being used and whether the achieved level of student understanding of sampling principles is a function of these tools. That evaluation currently underway focuses on student comprehension as a function of user-profiles that are tracked when students use the Tutor, including measures of time-on-task, use of multimedia presentations, and recitation results. A more fine-grained analysis will uncover which, if any, cognitive learning features of the Tutor are effective.

### Selected References

- Association of American Colleges. 1991. *Liberal Learning and the Arts and Sciences Major, Vol. 2. Reports from the Fields: Project on Liberal Learning, Study-In-Depth, and the Arts and Sciences Major*. Washington, D.C.: Association of American Colleges.
- Bradstreet, Thomas. 1996. "Teaching introductory statistics courses so that nonstatisticians experience statistical reasoning." *The American Statistician* 50:69-78.
- Cappell, Charles, L., Tom Haapoja, and Fang Gong. 2000. "Computer Based Training in Quantitative Reasoning and Analysis" Paper presented at the annual meeting of the American Sociology Association, Washington DC: August, 2000.
- \_\_\_\_\_. 2001a. "Implementing Cognitive Learning Strategies within an Interactive Tutor." Invited paper, NSF Grant Recipient, International conference on College Teaching and Learning, Jacksonville, FL. April 2001.
- \_\_\_\_\_. 2001b. "Using Interactive Tutors to Overcome Cognitive Learning Hurdles in Quantitative Reasoning." Paper presented at the Midwest Sociology Society meetings, St. Louis, MO. April, 2001.
- Chance, Beth. 2000. "Components of Statistical Thinking and Implications for Instruction and Assessment." Paper presented at the annual meeting of the American Educational Research Association, April, 2000.
- Cox, D.R. 1990. "Role of models in statistical analysis." *Statistical Science* 5:169-174.
- Derry, S., J.R. Levin, and L. Schauble. 1995. "Stimulating statistical thinking through situated simulations." *Teaching of Psychology* 22:51-57.
- Fauconnier, Giles. 1997. *Mappings in Thought and Language*. Cambridge, UKL: Cambridge University Press.
- Garfield, Joan. 1995. "How Students Learn Statistics." *International Statistical Review* 63:25-34.
- \_\_\_\_\_. 2000. "The Role of Statistical Reasoning in Learning Statistics." Paper presented at the annual meeting of the American Educational Research Association, April, 2000.
- Haapoja, Thomas. 2001. "Utilization of cognitive strategies within a computer-based tutorial for teaching statistical sampling." Dissertation completed at College of Education, Educational technology, Northern Illinois University. DeKalb, IL. 182 pages.
- Huberty, Carl J., Janna Dresden, and Byung-Gee Bak. 1991. "Relationships among dimensions of statistical knowledge." Pp. 456-460 in David Vere-Jones, (Ed.), *Proceedings of the Third International Conference on Teaching Statistics, Vol. 1*. Voorburg, Netherlands: International Statistical Institute.
- Kapadia, Ramesh. and Manfred Borovcnik (eds.). 1991. *Chance Encounters: Probability in Education*. Dordrecht, Netherlands: Kluwer Academic Publishers.
- Kvam, Paul. 2000. "The effect of active learning methods on student retention in engineering statistics." *The American Statistician* 54:136-140.
- Lovett, Marsha C., and Joel B. Greenhouse. 2000. "Applying Cognitive Theory to Statistics Instruction." *The American Statistician* 54: 196-206.
- Mathews, Colin. 1998. "The zeroth problem." *The American Statistician* 52:1-9.
- Moore, D.S. 1998 "Statistics Among the Liberal Arts." *Journal of the American Statistical Association* 93:1253-259.
- Schloz, R.W. 1991. "Psychological research in probabilistic understanding." Pp. 213-254 in R. Kapadia and M. Borovcnik (eds.). *Chance Encounters: Probability in Education*. Dordrecht, Netherlands: Kluwer Academic Publishers.
- Snee, Ronald. 1993. "What's missing in statistical education?" *The American Statistician* 47:149-54.
- Schau, Candace, and Mattern, Nancy. 1997 "Use of map techniques in teaching applied statistics courses." *The American Statistician* 51:171-175.
- Tversky, Amos and Daniel Kahneman. 1982. "Judgement under uncertainty: Heuristics and biases." Pp. 3-20 in Kahneman, D., P. P. Slovic, and A. Tversky, (Eds.). *Judgment Under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.
- Vellman, Paul, and David Moore. 1996. "Multimedia for teaching statistics: Promises and pitfalls." *The American Statistician* 50:217-225.
- Wild, C.J. 1994. "Embracing the 'Wider view' of Statistics." *The American Statistician* 48:163-171.