Learning Abstract Statistics Concepts Using Simulation

Jamie D. Mills
University of Alabama

The teaching and learning of statistics has impacted the curriculum in elementary, secondary, and post-secondary education. Because of this growing movement to expand and include statistics into all levels of education, there is also a considerable interest in how *to teach statistics. For statistics concepts that tend to be very difficult or abstract, many researchers have recommended using computer simulation methods (CSMs), but there have been very few empirically and theoretically based studies related to student achievement using these methods. The purpose of this study was to determine whether using CSMs enhanced student understanding of abstract statistics concepts for students enrolled in an introductory course. Based on a theoretical framework of how students learn statistics, the preliminary results of this study indicate some evidence that these methods may improve student understanding of abstract statistics concepts.*

Statistics courses are increasingly becoming apart of the mainstream curriculum in all levels of education. In post secondary education, the teaching and learning of statistics continues to be an integral part of the curriculum. In almost every discipline, the ability to understand, interpret, and critically evaluate research findings are becoming important objectives (Giesbrecht, 1996). In recent years however, an appreciation of the importance of statistics in the elementary and secondary grades has also evolved. The implementation of the Quantitative Literacy Project (QLP) of the American Statistical Association (ASA), which provides instructional materials on probability and statistics used in the pre-college curriculum, is one indication of interest in this movement (Scheaffer, 1988). In addition, the release of the NCTM Principles and Standards for School Mathematics (NCTM, 2000), designed to improve mathematics education from pre-kindergarten to grade 12, includes a content standard that also emphasizes statistical reasoning (i.e., Data Analysis and Probability). Consequently, many states now include and emphasize statistical thinking in their statewide curriculum guidelines (i.e., Alabama State Department of Education, 1989). Because of this growing movement to expand and include statistics into all curricula, Becker (1996) stated that there is also considerable interest in how to teach statistics -- in a variety of fields (Richardson, 1991) and to a variety of age groups (Shulte, 1979).

Another important change that has had a major impact on the teaching and learning of statistics over the past few decades has been the integration of computers, particularly in the post secondary classrooms. Microcomputer development has led to increased accessibility for students and an increase in the development of more user-friendly statistics packages (i.e., SAS, SPSS, MINITAB). Many teachers encourage students to use these software programs to perform routine data analysis tasks, often in hopes of enhancing student learning. Unfortunately, this may enable students to only master the mechanics of data analysis (Marasinghe, Meeker, Cook, & Shin, 1996). Even when students use software programs to apply these techniques, abstract statistics concepts may still be difficult for students to comprehend.

One exciting advantage of the microcomputer, which has been suggested in the literature, lies in its capability of enhancing student understanding of abstract or difficult concepts (Hesterberg, 1998; Kalsbeek, 1996; Shibli, 1990). By using current computing technology, it is possible to supplement standard data analysis assignments by providing students with additional statistical experiences through the use of computer simulation methods (CSMs). CSMs allow students to experiment with random samples from a population with known parameters for the purpose of clarifying abstract and difficult concepts and/or theorems of statistics. For example, students can generate 50 random samples of size 30 from a non-normal distribution and compute the mean for each random sample. A histogram of the sample means can show the student that the sampling distribution of the sample mean is normally distributed. Computer simulations are invaluable in this regard because hard to understand concepts can be illustrated visually using many standard programs (i.e., EXCEL, MINITAB). This may enhance the learning experience, especially for students in introductory statistics courses.

Many researchers in statistics education recommend the use of CSMs to teach abstract concepts in statistics (Bradley, Hemstreet, & Ziergenhagen, 1992; delMas, Garfield, & Chance, 1999; Halley, 1991; Hesterberg, 1998; Karley, 1990; Mittag, 1992; Prybutok, Bajgier, & Atkinson, 1991; Shibli, 1990). For example, interactive simulation programs on the World Wide Web (WWW) are the latest internet resources many educators are now using to illustrate statistics concepts. Ng and Wong (1999) reported using simulation experiments on the internet to illustrate Central Limit Theorem (CLT) concepts. At URL www.ruf.rice.edu/~lane/stat_sim/sampling_dist/index.html, the CLT can be demonstrated graphically, either in large lectures or by the student with guidance from the instructor. Other statistics educators have used simulation exercises on the internet for the CLT (West & Ogden, 1998) and with other topics (Schwarz, 1997; Schwarz & Sutherland, 1997).

Hesterberg (1998) reported that simulation methods can offer students intuitive understanding of confidence intervals (and other topics) and recommends the program S-PLUS due to its flexibility. These methods appear to be especially helpful for illustrating the interpretation and the 'behavior' of confidence intervals (i.e. whether the interval encloses the true parameter or not) and the 'randomness' of the sample mean.

The binomial distribution (Shibli, 1991), regression analysis (Franklin, 1992), sampling distributions (delMas et al., 1999; Marasinghe et al., 1996; Weir, McManus, & Kiely, 1990), hypothesis testing (Flusser & Hanna, 1991) and concepts related to survey sampling (Chang, Lohr, & McLaren, 1992; Kalsbeek, 1996; Schwarz, 1997) have also been recommended by researchers as topics in which CSMs can potentially facilitate the learning of difficult concepts.

Purpose

Many teachers and researchers have recommended using CSMs to teach difficult and abstract statistics concepts but there have been very few empirical studies to support the recommendations. The primary purpose of this paper is to determine the effect CSMs have on student understanding of abstract statistics concepts related to the CLT for students enrolled in an introductory level statistics course.

Theoretical Framework

In order to determine if computer-assisted or other innovative teaching methods are effective, a link to a theory or theories of learning can be the instructor’s most powerful tool in understanding and changing practice. Therefore, it is important to think about how students learn in general before new teaching and/or learning methods can provide some insight.

One theoretical framework about how students learn statistics has been discussed in the literature (See Mills, 2003). The theory of constructivism suggests that students develop and construct their own understanding both internally --by transforming, organizing, and reorganizing previous knowledge (Cobb, 1994; Greeno, Collins, & Resnick, 1996) as well as externally -- through environmental and social factors that are influenced by culture, language, and interactions with others (Bruning, Schraw, & Ronning, 1999). By constructing their own ideas and knowledge about statistics concepts from the CSM experiences, students will assimilate this new information to any previous knowledge. The theory of conceptual change contends that as a result of interacting with this new knowledge, learners may eventually ‘change’ their ideas, or come to understand the scientific explanations (Posner, Strike, Hewson, & Gertzog, 1982). In this instance, learners must realize that any ‘old’ and faulty ideas are at least partially in conflict with the ‘new’ and scientific ideas and that the scientific explanation provides a more convincing and powerful alternative (Posner et al., 1982). If learners are able to change to the new and scientific conceptions, they may then be able to solve problems over time in a more well-defined, straightforward fashion. The theory of expertise (Ericsson & Smith, 1991) may be used to describe the development of a learner's problem-solving skills which may provide some evidence of a learner's progress toward statistical understanding and literacy.

The section that follows presents the methods and procedures used in the study. The results, discussion, and final thoughts conclude the paper.

Method

A study was conducted for students enrolled in an introductory level statistics course to investigate the effect of CSMs on concepts related to the Central Limit Theorem (CLT). The following research questions were considered:

1) What effect do CSMs have on student understanding of abstract concepts related to the CLT?
2) What are student attitudes toward learning using CSMs?
3) How will students exposed to CSMs perform on a subsequent in-class examination?

<u>Instructional Units</u>

Two units were designed by the researcher which covered important and abstract concepts related to the CLT. One unit was considered a 'traditional' unit while the second unit, which used CSMs using the EXCEL program, was considered the 'experimental' unit. Both traditional and CSMs units began with an overview of the unit followed by a list of student-specific objectives. The objectives were: 1) recognize and understand the properties of the sampling distribution of the sample mean; that is, a) the mean of the sampling distribution for the sample mean is the population mean, and b) the standard deviation of this distribution is the standard error of the mean, 2) recognize and understand Theorem 1: If a random sample of n observations is selected from a population with a normal distribution, the sampling distribution of the sample mean will also be approximately normally distributed, and 3) recognize and understand the CLT: If a random sample of n observations is selected from any population, if n is sufficiently large ($n > 30$), the sampling distribution of the sample mean will also be approximately normally distributed.

The structure, organization, and introduction of concepts and examples in the units were identical for both units. The exception was that students in the CSM group used Excel to perform any experiments for the understanding of the abstract concepts while students in the traditional group read their units and worked with diagrams (similar to a textbook) for their understanding.

Measurement

The research questions were addressed by considering two cognitive measures and one affective measure of interest. Seven open-ended questions were used to measure student understanding of CLT concepts (See Table 1 in the Appendix). These same questions were administered before (Pre) and after (Post) the students were exposed to their units. These questions would address research question one. The affective measure, for research question two, was developed by the researcher to measure student attitudes toward their instructional unit (ATIU). The survey consists of 10 items measured on a 5-point Likert scale where higher scores indicate more positive attitudes toward the instructional unit. Finally, students' understanding of CLT concepts was evaluated later in the semester. Five multiple choice questions were included as a part of a classroom test in order to determine the effect of CSMs over a longer period of time (Follow-up).

Procedure

All participants, who were randomly assigned to groups, signed up as volunteers during the first few weeks of class. Fourteen students (45.2%) were randomly assigned to the computer simulation group (CSM) while 17 (54.8%) belonged to the traditional (T) group. The researcher provided an oral overview of the content of the units before the students began. While the students worked on their units, the researcher assumed the role of a facilitator, answering questions when needed. Other assistants were available to the students for questions regarding Excel as well as to insure that the students were engaged with the material.

The study took place at a large research university. The majority of the participants were female (74.2%) and European American (58%). Over 83% of the participants reported that their grade point average was between 3.6 - 4.0 and 41.9% indicated that they have never taken a statistic course before.

Results

Quantitative Analyses

The percent of students who responded correctly to the pre and post measures are reported in Table 2 in the Appendix. An inspection of the percent responding correctly after exposure to the instructional unit was initially of interest. For questions 1 and 2 on the post measure, it appeared that more students in the T group responded correctly to these questions. For example, 29.4% of the students in the T group responded correctly on the post measure for question 2 compared to 23.1% in the CSM group. For the remaining 5 questions on the post measure, a higher percentage of students in the CSM group answered these items correctly compared to students in the T group.

Although students were randomly assigned to groups and thus, the groups were considered 'equivalent', the small sample sizes as well as the percent correct reported for the pre-measure (i.e., question 6 – CSM percent is 0 and T is 17.7) could indicate possible differences between the two groups on the pre-measure. Therefore, a one-way ANOVA using the pre-measure as the dependent variable was conducted which revealed no statistically significant differences between the two groups ($F(1,29) = .527$, $p = .474$). Based on this result, it was assumed that the two groups were equivalent on their understanding of concepts related to the CLT before exposure to their respective instructional units.

A mixed-model analysis was conducted by considering group as the between-subjects factor and the pre and post measures as the repeated measures or the within-subjects factor (See Table 3). The results revealed evidence that performance on the measures depends on whether the students belong to the CSM or T group ($F (1,28) = 6.6$, $p = .015$). The nature of this interaction was investigated by comparing cell means at each level (simple comparisons) of group and measure. For example, for the within-subjects factor (measure) the simple comparisons between the pre and post measures at each level of Group were of interest. The dependent samples t-test revealed a statistically significant difference between the pre and post measure means (1.29 vs. 3.31) for the CSM group at the .05 level of significance ($t (12) = 4.3$, $p = .001$). This significant effect was associated with a Cohen's d effect size of 1.20, which indicates that the difference between pre and post measure means was estimated to be 1.20 standard deviations (Keppel, 1991, pp. 82-83), a very large effect according to Cohen's guidelines of .2 (small), .5 (medium), and .8 (large). In addition, the 95% confidence interval indicated that on average, students in the CSM group scored between .95 to 2.89 points higher on the post measure after exposure to their instructional unit. There was not a statistically significant difference between the pre and post measure means for the T group at the .05 level of significance ($t (16) = 1.9$, $p = .066$).

For the between-subjects factor (group), the independent samples t-test was used to determine if mean differences existed between the two groups for the pre and post measure. For the pre measure, there were no differences statistically between the CSM and T group means ($t (29) = .726$, $p = .474$), which confirms an earlier analysis of no group differences on the pre-measure means. However, there were statistically significant mean differences (2.12 vs. 3.31) between the two groups on the post measure ($t (28) = 2.35$, $p = .026$), a medium effect of .63. In addition, students in the CSM group score between .156 and 2.22 points higher on average, than students in the T group.

The ATIU survey was designed by the researcher to determine student attitudes about using their instructional unit. Students in the CSM group had more positive attitudes toward their unit than students in the T group. The mean for the CSM group (3.36) exceeded the mean for the T group (2.17), resulting in a statistically significant difference ($F(1, 29) = 21.08$, $p < .001$). The effect size of 2.33 indicated an extremely large effect.

Finally, student understanding of CLT concepts was evaluated later in the semester by including 5 multiple choice questions as part of a classroom test. Although students in the CSM group scored higher than students in the T group on the follow-up test, the results revealed no statistically significant difference in the means ($F(1, 26) = 1.01$, $p = .323$).

Qualitative Analyses

All responses to the pre and post measures were analyzed by condensing and categorizing the responses into distinct categories (Patton, 1990). The categories that emerged were either correct (the correct answer was provided to the question), incorrect/conceptual change (an incorrect answer was given but there was clear evidence of scientific conceptual change), and incorrect/other (an incorrect response). The correct responses were categorized into the categories for scoring purposes for the quantitative analyses. The incorrect responses were also analyzed to further investigate any possible learning effects that were not evident from the right/wrong scoring method. Therefore, the strategy of triangulation using both quantitative and qualitative methods (i.e., mixed-methods design) was also an effort to enhance the validity of the research findings (Payne, 1994).

The *incorrect* responses from the post measure only for the CSM and T group were of interest. For question 1 for the CSM group, 100% of the responses were incorrect. Of the incorrect responses, about 67% revealed some evidence that the students' ideas were beginning a positive conceptual change process. 'Distribution of samples of a population' and 'the distribution that represents the frequencies of means of repeated measures' were example student responses. For the 94.1% incorrect responses in the T group, 37.5% of the responses revealed evidence of conceptual change ('when you take a certain size sample multiple times from the population and average their means').

There were almost 77% and 71% incorrect responses for question 2 for the CSM and T groups, respectively. However, almost 29% of these responses indicated positive evidence of conceptual change for the CSM group ('the mean expected from a random sampling') compared to only 12.5% for students in the T group ('the value obtained if the test is repeated').

For questions 3 and 5 for both groups, there were no responses that were considered 'incorrect/conceptual change'. Responses such as 'normal distribution' and 'sample mean' were other incorrect responses for question 3 while 'skewed' and 'not enough samples to tell' were other incorrect responses for question 5.

There were 57.1% and 82.3% incorrect responses for question 4 for the CSM and T groups, respectively. Sixty percent of the incorrect responses were conceptual change responses for the CSM group ('standard deviation') compared to 44% for the T group ('standard deviation of a population).

Responses for questions 6 and 7 were difficult to categorize as a 'conceptual change' response or not. For example for question 6, 62.5% of the 64.3% incorrect responses for the CSM group reported that the shape should be uniform while 62.5% of the 64.7% incorrect responses for students in the T group reported some other incorrect response ('like an 'M''). For question 7 for the CSM group, 66.7% of the 23.1% incorrect responses reported that the shape would be uniform while other incorrect responses such as 'uniform and normal' and 'symmetrical' were reported for 66.7% of the 52.9% incorrect responses for students in the T group.

In summary, there was no way to evaluate conceptually changing ideas for questions 3 and 5 because there were no responses that fit this category. Similarly for questions 6 and 7, the researchers could not be certain if the simple responses were indicative of conceptual change. However, there was clear evidence that students' conceptual ideas about concepts related to the CLT were beginning a positive change for students in the CSM group regarding questions 1, 2, and 4, more so than for students in the T group. These findings seem to support the quantitative results as well.

Discussion

The theoretical framework that underlies this study advocates that students can learn statistics concepts using computer simulation methods. This study may provide some preliminary evidence that using CSMs may improve student understanding of difficult or abstract concepts related to the Central Limit Theorem. Although these findings may support previous research, it must be noted that the sample size for this study was small and as such, replications and additional research studies are needed for validation.

The evaluation of the qualitative student responses from the post measure appears to corroborate the quantitative findings. The quantitative results alone do indicate that 'something' happened between the pre and post measures but without additional evidence, it may be difficult to reveal exactly what. A further investigation into student individual responses revealed that students in the CSM group did appear to benefit more from the simulation experiences. The students in the CSM group not only scored higher than students in the T group on the majority of the questions, but their incorrect responses also indicated clear evidence of scientific conceptual change, a finding that was not evident for students in the T group.

The findings of this study also suggest that using these methods appear to improve students' immediate understanding of abstract concepts. The results associated with the follow-up test on the effect of CSMs were not as convincing. Although students in the CSM group scored higher on the follow-up test than the students in the T group, there was no evidence to indicate a statistically significant difference. It was assumed that students assigned to both groups would be exposed to similar activities before the follow-up test (i.e., completing homework or computer assignments, attending lecture, studying and preparing for the test, etc.). However, all of these assumptions may not be valid and it is possible that real differences existed between the two groups during the time between the post measure and the follow-up test. Related to this finding then, is assessing student performance over a much longer period of time to determine if true learning has been maintained. According to Weir et al. (1990), it is possible that using CSMs can facilitate deeper processing of abstract concepts, with this change occurring gradually and with the assistance of other learning experiences. Because this research focused primarily on determining the effect of CSMs over a shorter period of time, additional research is required to determine the extent to which this learning method affects performance over much longer periods of time.

Finally, the theoretical framework proposed in this study requires that the learner construct new concepts through numerous learning experiences in order to achieve 'equilibrium' or a changing of one's own thinking, according to Piaget (1970). According to Posner et al. (1982), this conceptual change process is a gradual one, which usually involves many learning activities in which cognitive conflicts could arise, as well as sufficient time for students to fully realize the meaning and implication of the new concepts. Although the students were not exposed to 'numerous' learning stimuli from the pre measure to the post measure, many student responses for the post survey revealed some evidence that concepts were changing. The time between the pre and post surveys was probably not adequate for a complete transformation of ideas; however, there was some indication of the beginning of this process.

Final Thoughts

Technology is a powerful medium that can provide efficient methods for delivering instructional objectives to students. It is gaining acceptance worldwide in academia and empirical research will be important to document the effect of these new learning tools on student achievement. The empirical research on learning in the statistics education literature is relatively scarce. With the emergence of the internet, online and distance education courses, and other related information technologies, additional empirical research studies are needed to evaluate these new approaches to learning statistics. Although empirical research using technology in teaching and research presents some interesting challenges, it affords us with many advantages. It provides valuable feedback for our teaching practices and its impact on student achievement. It will sharpen our technical skills as researchers and teachers. Finally, it appears to motivate our students to learn. With rapid advancements in technology and as today's learning environments continue to embrace the internet, web-based learning, and other related information technology, other important questions about how technology impacts statistics learning and achievement will continue to surface. Hopefully, the preliminary findings of this empirical study will continue to advance the research in statistics education and provide important implications for teaching and learning with technology.

References

Alabama State Department of Education. (1989). *Mathematics: Alabama course of study* (Bulletin 1989, No. 31). Montgomery, AL. (ERIC Document Reproduction Service No. ED 327399)

Becker, B. J. (1996). A look at the literature (and other resources) on teaching statistics. *Journal of Educational and Behavioral Statistics, 21,* 71-90.

Bradley, D. R., Hemstreet, R. L., & Ziegenhagen, S. T. (1992). A simulation laboratory for statistics. *Behavior Research Methods, Instruments, & Computers, 24*, 190-204.

Bruning, R., Schraw, G., & Ronning, R. (1999). *Cognitive psychology and instruction* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Chang, T.C., Lohr, S. L., & McLaren, C.G. (1992). Teaching survey sampling using simulation. *The American Statistician, 31(1)*, 17-23.

Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher, 23*, 13-200.

delMas, R. C., Garfield, J., & Chance, B.L. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education, 7*,(3). Retrieved June 7, 2000, from http://www.amstat.org/publications/jse/secure/v7n3/delmas.cfm

Ericsson, K. A., & Smith, J. (Eds.). (1991). *Toward a general theory of expertise.* Cambridge, NY: Cambridge University Press.

Flusser, P., & Hanna, D. (1991). Computer simulation of the testing of a statistical hypothesis. *Mathematics and Computer Education, 25*,158-164.

Franklin, L.A. (1992). Using simulation to study linear regression. *The College Mathematics Journal, 23*, 290-295.

Giesbrecht, N. (1996). *Strategies for developing and delivering effective introductory-level statistics and methodology courses*. Alberta, Canada. (ERIC Document Reproduction Service No. ED 393668)

Greeno, J., Collins, A., & Resnick, L. (1996). Cognition and learning. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational Psychology* (pp. 15-46). New York: Macmillan.

Halley, F. S. (1991). Teaching social statistics with simulated data. *Teaching Sociology, 19*, 518-525.

Hesterberg, T. C. (1998). Simulation and bootstrapping for teaching statistics. *Proceedings of the American Statistical Association Section on Statistical Education*, USA, 44-52.

Kalsbeek, W. D. (1996). The computer program called sample: A teaching tool to demonstrate some basic concepts of sampling (version 1.01). *Proceedings of the American Statistical Association Section on Statistical Education, USA,* 103-108.

Karley, L. M. (1990). Using computer graphics in statistics. *Mathematics and Computer Education*, *24*, 232-239.

Keppel, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Upper Saddle River: Prentice Hall.

Marasinghe, M. G., Meeker, W. Q., Cook, D., & Shin, T. (1996). Using graphics and simulation to teach statistical concepts. *The American Statistician, 50,* 342-351.

Mills, J. D. (2003). A theoretical framework for teaching statistics. *Teaching Statistics, 25(2)*, 56-58.

Mittag, K. C. (1992, April). *Using computers to teach the concepts of the central limit theorem.* Paper presented at the annual American Educational Research Association, San Francisco, CA.

NCTM (2000). *Principles and standards for school mathematics.* Reston, VA: National Council of Teachers of Mathematics.

Ng. V. M., & Wong, K. Y. (1999). Using simulation on the internet to teach statistics. *The Mathematics Teacher, 92*, 729-733.

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park: Sage.

Payne, D. A. (1994). *Designing educational project and program evaluations: A practical overview based on research and experience.* Norwell, MA: Kluwer.

Piaget, J. (1970). Piaget's theory. In P.H. Mussen (Ed.), *Carmichael's manual of child psychology.* New York: Wiley.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66*, 211-227.

Prybutok, V. R., Bajgier, S. M., & Atkinson, M. (1991). Statistics instruction: The next generation. *Journal of Computers in Mathematics and Science Teaching, 10*, 15-27.

Richardson, T. H. (1991). Reproducible bad data for instruction in statistical methods. *Journal of Chemical Education, 68*, 310-311.

Scheaffer, R. (1988). Statistics in the schools: The past, present and future of the quantitative literacy project. *Proceedings of the American Statistical Association Section on Statistical Education*, USA, 71-78.

Schwarz, C. J. (1997). StatVillage: An on-line, www-accessible, hypothetical city based on real data for use in an introductory class in survey sampling. *Journal of Statistics Education, 5*(2). Retrieved May 1, 2002, from http://www.amstat.org/publications/jse/v5n2/schwarz.html

Schwarz, C. J., & Sutherland, J. (1997). An on-line workshop using a simple capture-recapture experiment to illustrate the concepts of a sampling distribution. *Journal of Statistics Education, 5*(1). Retrieved May 1, 2002, from http://www.amstat.org/publications/jse/v5n1/schwarz.html

Shibli, M. A. (1990). A two-stage exercise on the binomial distribution using Minitab. *Collegiate Microcomputer, 8*(1), 55-60.

Shulte, A. P. (1979). A case for statistics. *Arithmetic Teacher, 26*, 24.

Weir, C. G., McManus, I. C., & Kiely, B. (1990). Evaluation of the teaching of statistical concepts by interactive experience with Monte Carlo simulations. *British Journal of Educational Psychology, 61*, 240-247.

West, R. W., & Ogden, R. T. (1998). Interactive demonstrations for statistics education on the world wide web. *Journal of Statistics Education, 6*(3). Retrieved November 7, 2000, from http://www.amstat.org/publications/jse/v6n3/west.html

Appendix

Table 1

Open-Ended Measure (Pre and Post)

1. What is a sampling distribution?
2. What is meant by expected value?
3. What is the expected value of the sampling distribution of the sample mean?
4. What is meant by the standard error of the mean?
5. If we take a random sample of 10 observations from a normal population, what do we expect the shape of the sampling distribution of the sample mean to be?
6. If we take a random sample of 10 observations from a uniform population, what do we expect the shape of the sampling distribution of the sample mean to be?
7. If we take a random sample of a very large number of observations from a uniform population, what do we expect the shape of the sampling distribution of the sample mean to be?

ATIU

1-strongly disagree 2-disagree 3-don't know 4-agree 5-strongly agree

1. This unit helped me to understand the material better.
2. This unit made learning the material more interesting.
3. Learning using this unit made the material more difficult to understand.
4. I liked this unit and hope more concepts in statistics can be taught this way.
5. I believe a unit like this illustrated in a textbook will be a waste of time.
6. I believe I can understand other concepts in statistics using this type of unit.
7. A unit like this is not the best way to introduce new concepts.
8. I believe units illustrated like this in a textbook will improve my grades.
9. This unit will not improve my understanding of concepts in statistics.
10. This unit should not be used to help students learn statistics.

5-Item Multiple Choice – Follow-Up

1. The CLT is important in statistics because
    a. for a large n, the population is normally distributed
    b. for any population, the sampling distribution of the sample mean is approximately normally distributed, regardless of sample size
    c. for a large n, the sampling distribution of the sample mean is normally distributed, regardless of population
    d. for any size sample, the sampling distribution of the sample mean is normally distributed

2. Which of the following statements about the sampling distribution of the sample mean is **<u>incorrect</u>**?
    a. the sampling distribution is approximately normal whenever the sample size is sufficiently large ($n > 30$)
    b. the sampling distribution is generated by repeatedly taking samples of size n and computing the sample means
    c. the mean of the sampling distribution is μ
    d. the standard deviation of the sampling distribution is σ

3. Which of the following statements concerning the CLT is correct?
   a. The CLT states that the sample mean is always equal to the population mean
   b. The CLT states that for large samples, the standard error of the mean is equal to the population standard deviation
   c. The CLT states that for large samples, the sampling distribution of the population mean is approximately normal
   d. The CLT states that for large samples, the sampling distribution of the sample means is approximately normal

4. A random sample of 9 scores is selected from a very large population of self-esteem scores that is normally distributed with mean 10.0 and standard deviation .01. If x denotes the sample mean self-esteem score for the 9 scores in the sample, which of the following statements is true about the sampling distribution for x?
   a. since the sample is not large, we do not know the sampling distribution of x
   b. the sampling distribution x is normally distributed with mean 10.0 and standard deviation .01
   c. the sampling distribution x has a uniform distribution with mean 10.0 and standard deviation .003
   d. the sampling distribution x has a normal distribution with mean 10.0 and standard deviation of .003
   e. both a and c are true

5. The CLT applies to
   a. normal distributions only
   b. uniform distributions only
   c. skewed distributions only
   d. all of the above
   e. a and b only

Table 2
Descriptive Statistics for Open-Ended Measure

Question	Pre Percent Correct	Post Percent Correct	
1	0*	0*	
	0	5.9	
2	0*	23.1*	
	5.9	29.4	
3	21.4*	57.1*	
	0	23.5	
4	0*	42.9*	
	5.9	17.7	
5	64.3*	78.6*	
	82.4	52.9	6
	0*	35.7*	
	17.7	35.3	
7	42.9*	76.9*	
	47.1	47.1	

Note. The * denotes the percent correct for the CSM group. No asterisk denotes the percent correct for the T group.

Table 3
Mixed-Model Analysis of Variance Results

	CSM Group		
	n	M	SD
Pre	14	1.29*	.914
Post	13	**3.31*	1.437
ATIU	14	3.36*	.589
Follow	11	3.36	1.433
	T Group		
	n	M	SD
Pre	17	1.53	.943
Post	17	**2.12	1.317
ATIU	17	2.17*	.804
Follow	17	2.71	1.829

Note. Means comparisons are 1) Pre and Post for CSM (*), 2) Post for CSM and T (**), and 3) ATIU for CSM and T (*).
*p < .01. **p < .05.