18. WORKSHOP STATISTICS: USING TECHNOLOGY TO PROMOTE LEARNING BY SELF-DISCOVERY

Allan J. Rossman
Dickinson College

INTRODUCTION

Technology has been used as an active learning tool in *Workshop Statistics*, a project that involved the development and implementation of curricular materials which guide students to learn fundamental statistical ideas through self-discovery. Using the workshop approach, the lecture-format was completely abandoned. Classes are held in microcomputer-equipped classrooms in which students spend class-time working collaboratively on activities carefully designed to enable them to discover statistical concepts, explore statistical principles, and apply statistical techniques.

The workshop approach uses technology in three ways. First, technology is used to perform the calculations and present the visual displays necessary to analyze real datasets, which are often large and cumbersome. Freeing students from these computational chores also empowers the instructor to focus attention on the understanding of concepts and interpretation of results. Second, technology is used to conduct simulations, which allow students to visualize and explore the long-term behavior of sample statistics under repeated random sampling. Whereas these two uses of technology are fairly standard, the most distinctive use of technology within the workshop approach is to enable students to explore statistical phenomena. Students make predictions about a statistical property and then use the computer to investigate their predictions, revising their predictions and iterating the process as necessary.

Although my direct experience with *Workshop Statistics* concerns introductory statistics courses at the college level, the curricular materials have also been used with secondary students. Moreover, the pedagogical approach and its corresponding uses of technology are certainly appropriate at any educational level. I use the software package Minitab (1995) when I teach the course, but the materials are written in a flexible manner to permit the use of other statistics packages. One can also use a graphing calculator such as the TI-83.

In this paper, I first describe this course and its accompanying textbook/workbook *Workshop Statistics: Discovery with Data* (Rossman, 1996; Rossman & von Oehsen, 1997). I then provide specific examples of the uses of technology that I list above, concentrating on the use of technology for exploring statistical phenomena such as correlation, sampling distributions, and confidence intervals. I conclude by offering some suggestions for research questions to be addressed regarding the role of technology in statistics education.

OVERVIEW OF WORKSHOP STATISTICS

"Shorn of all subtlety and led naked out of the protective fold of educational research literature, there comes a sheepish little fact: lectures don't work nearly as well as many of us would like to think" (Cobb, 1992, p. 9).

"Statistics teaching can be more effective if teachers determine what it is they really want students to know and to do as a result of their course- and then provide activities designed to develop the performance they desire" (Garfield, 1995, p. 32).

Workshop Statistics replaces lectures with collaborative activities that guide students to discover statistical concepts, explore statistical principles, and apply statistical techniques. Students work toward these goals by analyzing real data and by interacting with each other, with their instructor, and with technology. These activities require students to collect data, make predictions, read about studies, analyze data, discuss findings, and write explanations. The instructor's responsibilities include evaluating students' progress, asking and answering questions, leading class discussions, and delivering "mini-lectures" where appropriate. The essential point is that every student is actively engaged with learning the material through reading, thinking, discussing, computing, interpreting, writing, and reflecting. In this manner, students construct their own knowledge of statistical ideas as they work through the activities.

Workshop Statistics focuses on the *big ideas* of statistics, paying less attention to details that often divert students' attention from larger issues. Little emphasis is placed on numerical and symbolic manipulations. Rather, the activities lead students to explore the meaning of concepts such as variability, distribution, outlier, tendency, association, randomness, sampling, sampling distribution, confidence, significance, and experimental design. Students investigate these concepts by experimenting with data, often with the help of technology. Many of the activities challenge students to demonstrate their understanding of statistical issues by asking for explanations and interpretations rather than mere calculations.

In an effort to deepen students' understandings of fundamental ideas, I present these ideas repetitively. For example, students return to techniques of exploratory data analysis when studying properties of randomness and also in conjunction with inference procedures. They also encounter issues of data collection not just when studying randomness but also when investigating statistical inference.

I believe that the workshop approach is ideally suited to the study of statistics--the science of reasoning from data--because it forces students to be actively engaged with real data. Analyzing real data not only exposes students to what the practice of statistics is all about, it also prompts them to consider the wide applicability of statistical methods and often enhances their enjoyment of the material.

Some activities ask students to analyze data that they collect in class about themselves, but most of the activities present students with real data from a variety of sources. Many questions in the text ask students to make predictions about data before conducting their analyses. This practice motivates students to view data not as simply numbers but as numbers with a context, to identify personally with the data, and to take an interest in the results of their analyses.

The datasets do not concentrate in one academic area but come from a variety of fields of application. These fields include law, medicine, economics, psychology, political science, and education. Many examples come not from academic disciplines but from popular culture. Specific examples, therefore, range from such pressing issues as testing the drug AZT and assessing evidence in sexual discrimination cases to less crucial ones of predicting basketball salaries and ranking *Star Trek* episodes.

For the most part, I cover traditional subject matter for a first course in statistics. The first two units concern descriptive and exploratory data analysis; the third introduces randomness and probability; and the final three delve into statistical inference. The six units of course material are divided into smaller topics--at Dickinson I cover one topic per 75-minute class period. I begin each class by posing a series of preliminary questions designed to get students thinking about issues and applications to be studied, and often to collect data on themselves. Then I provide a brief overview of the major ideas for the day and ask students to start working on the activities in the text. The text leaves space in which students record their responses.

Technology plays an integral role in this course. The text assumes that students have access to technology for creating visual displays, performing calculations, and conducting simulations. Roughly half of the activities ask students to use technology. Students typically perform small-scale displays, calculations, and simulations by hand before having the computer or calculator take over those mechanical chores. Activities requiring the use of technology are integrated throughout the text, reinforcing the idea that technology is not to be studied for its own sake but rather is an indispensable tool for analyzing real data and a convenient device for exploring statistical phenomena.

A variety of teaching resources related to *Workshop Statistics* , including information on how the course is implemented and assessed at Dickinson College, is available on the World Wide Web at: http://www.dickinson.edu/~rossman/ws/.

The sections below describe how *Workshop Statistics* uses technology in helping students to discover three important statistical ideas--correlation, sampling distributions, and confidence intervals. Because the activities and questions form the core of the course, I frequently cite examples below and set them off with bullets (•). I want to emphasize that the distinctive features of these questions are that students spend class-time working collaboratively to address these questions, recording their observations and explanations in the text/workbook itself, and that these activities and questions lead students to discover statistical ideas for themselves, as opposed to their being lectured to about them or presented with examples of them.

Example: Correlation

I devote two topics to the fundamental ideas of association and correlation, the first dealing primarily with graphical displays of association and the second with the correlation coefficient as a numerical measure of association. Preliminary questions to get students thinking about both the statistical issues themselves and some of their applications include:

- Do you expect that there is a tendency for heavier cars to get worse fuel efficiency (as measured in miles per gallon) than lighter cars?
- Do you think that if one car is heavier than another that it must always be the case that its gets a worse fuel efficiency?
- Take a guess as to the number of people per television set in the United States in 1990; do the same for China and for Haiti.
- Do you expect that countries with few people per television tend to have longer life expectancies, shorter life expectancies, or do you suspect no relationship between televisions and life expectancy?

Students begin by constructing a scatterplot by hand using a small example of weights and fuel efficiency ratings of cars. They then allow technology to produce the scatterplots, and the students interpret the results from examples dealing with marriage ages, family sizes, and space shuttle O-ring failures. In addition to allowing students to concentrate on interpretation, using technology also frees them to focus on the concept of association itself. At this point, students respond more knowingly to the question about whether there is a tendency for heavier cars to get worse fuel efficiency than lighter cars. They also discover that this tendency does not mean that heavier cars always have a worse fuel efficiency rating than lighter ones, because they can cite pairs of cars for which the tendency does not hold.

The topic on correlation relies very heavily on technology. Students begin by examining scatterplots of hypothetical exam scores for six classes. They classify the association in each scatterplot as positive or negative and as most strong, moderate, or least strong. (They have studied these ideas in the previous topic.) In so doing they fill in Table 1.

Table 1: Example of table students use for correlation lesson

	most strong	moderate	least strong
negative	C	D	F
positive	E	A	B

Students then use technology to calculate the value of the correlation coefficient between exam scores in each of these classes, recording the values in the appropriate cells of the table above (see Table 2).

Table 2: Correlation coefficients for exam score example

	most strong	moderate	least strong
negative	C -.985	D -.720	F -.472
positive	E .989	A .713	B .465

Students next proceed to answer the following questions about properties of correlation by working together in pairs and recording their answers in their work/textbook:

- Based on these results, what do you suspect is the largest value that a correlation coefficient can assume? What do you suspect is the smallest value?
- Under what circumstances do you think the correlation assumes its largest or smallest value; i.e., what would have to be true of the observations in that case?
- How does the value of the correlation relate to the **direction** of the association?
- How does the value of the correlation relate to the **strength** of the association?

In this manner, students discover for themselves (and with assistance from technology) the basic properties of correlation that instructors typically recite for them. Students use technology to produce scatterplots and to calculate correlations in cases that lead them to see that correlation measures only linear association and that correlation is not resistant to outliers.

After having discovered these properties of correlation, students then use data measuring the life expectancy and number of people per television set in a sample of countries. They use technology to examine a scatterplot and to calculate the correlation coefficient, which turns out to be -.804. Students then address the following questions:

- Since the association is so strongly negative, one might conclude that simply sending television sets to the countries with lower life expectancies would cause their inhabitants to live longer. Comment on this argument.
- If two variables have a correlation close to +1 or to -1, indicating a strong linear association between them, does it follow that there must be a cause-and-effect relationship between them?

Students discover for themselves, again working collaboratively and with the help of technology, the fundamental principle that correlation does not imply causation.

The next activity that students encounter in their study of correlation is a guessing game designed to help them judge the value of a correlation coefficient from looking at a scatterplot. I use a Minitab macro to generate data from a bivariate normal distribution where the correlation coefficient is chosen from a uniform distribution on the interval (-1,1). As students execute the macro, they see the scatterplot, make their guess for the correlation coefficient, and only then prompt the macro to reveal the actual value of the correlation. Students repeat this a total of 10 times, recording the results in a table (see Table 3) and proceeding to answer the following questions.

Table 3: Table used by students to enter estimated and actual correlation coefficients

	repetition									
	1	2	3	4	5	6	7	8	9	10
guess										
actual										

- Make a guess as to what the value of the correlation coefficient between **your guesses** for *r* and the **actual values** of *r* would be.
- Enter your guesses for *r* and the actual values of *r* into the computer and have the computer produce a scatterplot of your guesses vs. the actual correlations. Then ask it to compute the correlation between them; record this value below.

Students invariably guess that the correlation between their guesses and the actual values will be much lower than it really turns out to be.

Only at this stage do I show students a formula for calculating the correlation coefficient. I present them an expression involving products of z-scores in the hope that it will make some sense to them. Rather than have them calculate a correlation by hand, I ask them to fill in a few missing steps in such a calculation.

Most of my students seem to catch on very quickly to the basic properties of correlation, and they need no more than a little guidance with using the software. My hope is that they can better understand and apply the idea of correlation, having discovered it to some degree on their own. Technology clearly plays a pivotal role in this process.

Example: Sampling distributions

Technology also plays a key role in helping students learn about sampling distributions in *Workshop Statistics*. Three topics address this issue. The first introduces the idea of sampling in general and random sampling in particular; the second looks at sampling distributions specifically focusing on the notion of confidence; and the third examines sampling distributions by introducing the concept of statistical significance. Preliminary questions to prepare students for these ideas include the following:

- If Ann takes a random sample of 5 Senators and Barb takes a random sample of 25 Senators, who is more likely to come close to the actual percentage breakdown of Democrats/Republicans in her sample?
- Is she (your answer to the previous question) **guaranteed** to come closer to this percentage breakdown?

The principal role of technology in studying sampling distributions is to conduct simulations that reveal the long-term behavior of sample statistics under repeated random sampling. I question students' abilities to understand simulation results and to see in them what instructors expect them to see. To address this, I always have students perform small-scale simulations by hand before proceeding using technology.

Students first experience random sampling by using a table of random numbers to select a random sample of 10 from the population of 100 U.S. Senators. They analyze variables such as gender, party affiliation, and years of service for the senators in their sample; their sample does not mirror the population. Students then use the computer to generate 10 random samples of 10 senators each and analyze the results by responding to the following questions:

- Did you get the same sample proportion of Democrats in each of your ten samples? Did you get the same sample mean years of service in each of your ten samples?
- Create (by hand) a dotplot of your sample proportions of Democrats.
- Use the computer to calculate the mean and standard deviation of your sample proportions of Democrats.

In so doing, students discover the crucial (if obvious) notion of sampling variability and begin to consider the critical issue of sampling distributions. Next, sample size is increased. Students have the computer generate 10 random samples of 40 senators each and then answer the following questions:

- Again use the computer to calculate the mean and standard deviation of your sample proportions of Democrats.
- Comparing the two plots that you have produced, in which case (samples of size 10 or samples of size 40) is the **variability** among sample proportions of Democrats greater?
- In which case (samples of size 10 or samples of size 40) is the result of a **single** sample more likely to be close to matching the truth about the population?

With this activity, students begin to discover the effect that sample size has on a sampling distribution, an idea to which they return often.

To study sampling distributions in the context of the notion of confidence, students take samples of 25 Reese's Pieces candies and note the proportion of orange candies in their sample. Comparing results across the class convinces them of sampling variability; that is, that values of sample statistics vary from sample to sample. The students begin to notice a pattern to that variation, which is that most of the sample proportion values are somewhat clustered together. The following question asks them to begin to think about the notion of confidence:

- Again assuming that each student had access only to her/his sample, would most estimates be reasonably close to the true parameter value? Would some estimates be way off? Explain.

Students then use the computer to simulate 500 samples of 25 candies. This necessitates specifying the value of the population proportion of orange candies (I choose .45), which in turn allows students to determine how many of the simulated sample proportions fall within a certain distance of that specified population value:

- Use the computer to count how many of the 500 sample proportions are within ±.10 of .45 (i.e., between .35 and .55). Then repeat for within ±.20 and for within ±.30.
- Forget for the moment that you have designated that the population proportion of orange candies be .45. Suppose that each of the 500 imaginary students was to estimate the population proportion of orange candies by going a distance of .20 on either side of her/his sample proportion. What percentage of the 500 students would capture the actual population proportion (.45) within this interval?
- Still forgetting that you actually know the population proportion of orange candies to be .45, suppose that you were one of those 500 imaginary students. Would you have any way of knowing definitively whether your sample proportion was within .20 of the population proportion? Would you be reasonably "confident" that your sample proportion was within .20 of the population proportion? Explain.

Students then increase the sample size to 75 and ask the computer for 500 new simulated samples of candies. They compare the results of this simulation to those with the smaller sample size:

- How has the sampling distribution changed from when the sample size was only 25 candies?
- Use the computer to count how many of these 500 sample proportions are within ±.10 of .45. Record this number and the percentage below.
- How do the percentages of sample proportions falling within ±.10 of .45 compare between sample sizes of 25 and 75?
- In general, is a sample proportion more likely to be close to the population proportion with a larger sample size or with a smaller sample size?

In this manner, students again discover the effects of sample size and learn that larger samples generally produce more accurate estimates. The computer simulations also enable students to check the reasonableness of the Central Limit Theorem, because they observe the approximate normality of the simulated sample proportions and verify that the mean and standard deviation of the simulated sample proportions are in fact close to those indicated by the Central Limit Theorem.

In order to study sampling distributions and the concept of statistical significance, students roll dice to simulate defective items coming off an assembly line. The hypothetical question addressed by students is whether finding four defective items in a batch of 15 constitutes strong evidence that the defective rate of the process is less than one-third. Students investigate how often such results would occur by chance alone as they roll dice to represent the manufactured items.

After combining and analyzing the dice results for the class, students use the computer to simulate 1,000 batches of 15 items each with a defective rate of one-third. They discover the idea of significance as they answer these questions:

- How many and what proportion of these 1000 simulated batches contain four or fewer defectives?
- Based on this more extensive simulation, would you say that it is very unlikely for the process to produce a batch with four or fewer defectives when the population proportion of defectives is one-third?
- Suppose again that the engineers do not know whether or not the modifications have improved the production process, so they sample a batch of 15 widgets and find four defectives. Does this finding provide strong evidence that they have improved the process? Explain.
- Now suppose that the engineers find just 2 defective widgets in their sample batch. About how often in the long run would the process produce such an extreme result (2 or fewer defectives) if the modifications did not improve the process (i.e., if the population proportion of defectives were still one-third). Base your answer on the 1000 simulated batches that you generated above.
- Would finding two defectives in a sample batch provide strong evidence that the modification had in fact improved the process by lessening the proportion of defectives produced? Explain.
- Repeat the previous two questions if the engineers find no defectives in the sample batch.

Technology plays an invaluable role in helping students understand the nature of sampling distributions, an idea that they must master in order to comprehend the nature and interpretation of statistical inference.

Example: Confidence intervals

I ask students to study confidence intervals in the context of estimating a population proportion before moving on to population means. One reason for this is that with binary variables the population proportion describes the population completely. Thus, one can avoid complications of dealing with measurement variables by not needing to worry about shapes of distributions or about whether one should look at the mean or median or some other measure of center. Another reason is the simplicity of simulating long-term behavior of sample proportions as opposed to sample means. As I describe above, one can use candies and dice to conduct these simulations; again one need not worry about assumptions about population distributions.

Preliminary questions that students answer as they begin to study confidence intervals include:

- If a new penny is spun on its side (rather than tossed in the air), about how often would you expect it to land "heads" in the long run?
- Spin a new penny on its side five times. Make sure that the penny spins freely and falls

naturally, without hitting anything or falling off the table. How many heads did you get in the five spins?

- Pool the results of the penny spinning experiment for the entire class; record the total number of spins and the total number of heads.
- Take a guess as to what proportion of students at this college wear glasses in class.
- Mark on a number line an interval which you believe with 80% confidence to contain the actual value of this population parameter (proportion of students at this college who wear glasses in class).
- Mark on a number line an interval which you believe with 99% confidence to contain the actual value of this population parameter.
- Which of these two intervals is wider?
- Record the number of students in this class and the number who are wearing glasses.

After presenting students with the familiar expression for constructing a confidence interval for a population proportion, I ask them to produce a confidence interval by hand from the penny spinning data using the data pooled together from the class experiment. Then students use the computer to simulate 200 experiments of 80 penny spins each, assuming the actual proportion of spins that land "heads" is $\theta = .35$, and to construct a 95% confidence interval in each case. They then answer the following as a way of discovering what "confidence" means:

- How many of the 200 confidence intervals actually contain the actual value (.35) of θ? What percentage of the 200 experiments produce an interval which contains .35?
- For each interval that does not contain the actual value (.35) of θ, record its sample proportion $\hat{p}$ of heads obtained.
- If you had conducted a <u>single</u> experiment of 80 penny spins, would you have any <u>definitive</u> way of knowing whether your 95% confidence interval contained the actual value (.35)?
- Explain the sense in which you would be 95% confident that your 95% confidence interval contains the actual value (.35).

Students then investigate various properties of confidence intervals by having a Minitab macro to do the calculations for them. They first examine the effect of the confidence level selected and then the effect of the sample size used. As they use technology to address the following questions, they discover that confidence intervals get wider as one requires more confidence and that they get narrower as one increases the sample size. Students are asked to answer the following questions:

- **Intuitively** (without thinking about the formula), would you expect a 90% confidence interval for the true value of θ to be wider or narrower than a 95% interval? (Think about whether the need for greater confidence would produce a wider or a narrower interval.) Explain your thinking.
- To investigate your conjecture, use the computer to produce 80%, 90%, 95% (which you have already done by hand), and 99% confidence intervals and record the results in this table:

confidence level	confidence interval	half-width	width
80%			
90%			
95%			
99%			

- Do you need to modify your earlier answer in light of these findings?
- How would you **intuitively** (again, without considering the formula) expect the width of the 95% confidence interval to be related to the sample size used? For example, would you expect the interval to be wider or narrower if 500 pennies were spun as opposed to 100 pennies? Explain your thinking.
- To investigate this question, use the computer to produce a 95% confidence interval using each of the sample sizes listed in the table. Fill in the table below:

sample size	sample heads	confidence interval	half-width	width
100	35			
400	140			
800	280			
1600	560			

- Do you need to modify your earlier answer in light of these findings?
- How does the half-width when n = 100 compare to the half-width when n = 400? How does the half-width when n = 400 compare to that when n = 1600?
- What effect does **quadrupling** the sample size have on the half-width of a 95% confidence interval?

In addition to allowing the students to explore properties of confidence intervals, using technology can free students from computational burdens so that they can focus on what inference procedures are all about. For example, I ask students the following to emphasize the ideas that confidence intervals only make sense when one wants to estimate a population parameter from a sample statistic and that using confidence intervals may not produce reasonable estimates when one starts with a biased sample:

- Suppose that an alien lands on Earth, notices that there are two different sexes of the human species, and wants to estimate the proportion of all humans who are female. If this alien were to use the members of the 1994 United States Senate as a sample from the population of human beings, it would have a sample of 7 women and 93 men. Use this sample information to form (either by hand or using the computer) a 95% confidence interval for the actual proportion of all humans who are female.
- Is this confidence interval a reasonable estimate of the actual proportion of all humans who are female?
- Explain why the confidence interval procedure fails to produce an accurate estimate of the population parameter in this situation.
- It clearly does not make sense to use the confidence interval in (a) to estimate the proportion of women in the world, but does the interval make sense for estimating the proportion of women in the U.S. Senate in 1994? Explain your answer.

When students learn about tests of significance, they use technology to discover the duality between tests and confidence intervals. They also explore the distinction between statistical significance and practical significance, learning the importance of using confidence intervals in conjunction with tests of significance (and in *Workshop Statistics*, they use technology to do the computational work). Similarly, students use technology to investigate properties of confidence intervals for population means. At that point, they study the effects of sample variability as well as confidence level and sample size.

Some might argue that instructors could simply tell students about these properties of confidence intervals and provide them with examples, but I hope that by tackling the ideas themselves, students are better able to understand and apply them. Some would also argue that students should investigate the properties by analyzing the formulas and noting, for example, the effect of the square root of the sample size appearing in the denominator. I prefer to try to instill some intuitive sense about statistics and variability in the students rather than have them view the subject as the results of formulas. A third objection might be that students could investigate properties as I have described but without the use of technology. Although this is certainly true, I like to think that technology helps the student to concentrate on the substantive concepts rather than on the numerical manipulations.

RESEARCH DIRECTIONS

The overarching research question that emerges from the *Workshop Statistics* project is the obvious one: Does any of this really work? More specifically, are students' understandings of and abilities to apply fundamental statistical ideas enhanced by using technology to promote learning by self-discovery? I have not designed studies to collect data on this question, but I do have some thoughts on directions that one might pursue. I believe that the three areas that I address above--correlation, sampling distributions, and confidence intervals--provide fruitful ground for pedagogical research.

Correlation should be a prime topic to study for at least two reasons: (1) helping students to understand ideas related to correlation is a goal of most statistics instructors, and (2) most students of *Workshop Statistics* come to discover the basic properties of correlation fairly quickly and easily. Some questions to explore include:

- Do students who study correlation in the manner I describe above really have a firmer grasp of its basic properties than those who study the topic in a more traditional setting?
- Are students who encounter the distinction between correlation and causation by discovering and explaining it for themselves more likely to apply it thoughtfully in new situations than other students?
- Does a computerized guessing game in which students judge the value of a correlation coefficient from a scatterplot help them to become more proficient at doing that?

The study of sampling distributions also lends itself to pedagogical research. Although I firmly believe that computer simulations are very powerful and enlightening tools for understanding the long-term behavior of sample statistics under repeated sampling, I question whether many introductory students really see and understand what instructors want them to when they look at the simulation results. As I describe above, in the hope that they will better understand the computer output, I always ask students to perform real simulations with candies or dice or cards before they use the computer. Even so, my anecdotal impression is that although most students seem to understand sampling distributions when they study them directly, they have much difficulty connecting the ideas to those of confidence intervals and tests of significance when they move to those topics. I would welcome research investigating whether students genuinely understand simulation results and whether they make connections between sampling distributions and confidence intervals or *p*-values.

Confidence intervals provide a third prime topic for research. I question whether many introductory students really understand the long-term repeated sampling interpretation of a confidence interval. More importantly, I would welcome efforts to study whether students who study confidence intervals with the hands-on, self-discovery approach that I describe above develop a better intuition for their properties or are able to apply them more thoughtfully than others. I especially wonder whether such students are better positioned to detect inappropriate applications of the procedure.

At a broader level, I would like to raise questions about the best way to study and evaluate pedagogical and curricular reform. Some of the many questions that arise include:

- How do we operationally define “more understanding” or “better results”?
- Do observational studies and anecdotal evidence have anything to contribute?
- Can we design controlled experiments to address the questions of interest?
- How can the many confounding variables be controlled?
- To what populations can findings be generalized?
- What ethical considerations come into play?
- What work has already been done, and what can reasonably be expected?
- How can teachers of statistics make better decisions about which teaching/learning methods would work best in their classrooms?

Students generally respond very favorably to this workshop approach to learning statistics. They enjoy the productive learning environment created in the classroom, they appreciate the collection and analysis of real data, and they like the collaborative element of their work. Most students perceive the workshop approach as facilitating their learning. With rare exceptions, students have no trouble using Minitab after the first few class periods, with minimal help.

As the instructor, I too enjoy the classroom learning environment, because the students are genuinely engaged with the material at almost all times. They do not have the luxury of copying notes with the intention of figuring things out later; they must immerse themselves in consideration of the issues during class-time. At Dickinson, the class size is limited to 24 students, so I interact with each pair of students several times during every class period, which allows me to identify each student’s strengths and weaknesses. I believe, anecdotally, that they also acquire a deeper understanding of fundamental ideas, but I would welcome research that would shed light on this question.

REFERENCES

Cobb, G. W. (1992). Teaching statistics. In L. Steen (Ed.), *Heeding the call for change: Suggestions for curricular action, MAA Notes*, No. 22, 3-43.

Garfield, J. B. (1995). How students learn statistics. *International Statistical Review*, *63*, 25-34.

Minitab Inc. (1995). *Minitab* [Computer program]. State College, PA: Author.

Rossman, A. J. (1996). *Workshop Statistics: Discovery with data*. New York: Springer-Verlag.

Rossman, A. J., & von Oehsen, J. B. (1997). *Workshop Statistics: Discovery with data and the graphing calculator*. New York: Springer-Verlag.