Restructuring Introductory Courses at Penn State

William Harkness, Patricia Buchanan, Robert Heckard, and James Rosenberger
Department of Statistics, Penn State University, University Park, PA 16802

Keywords. Large classes; technology usage; collaborative learning; readiness assessment testing; cost effectiveness.

1. Introduction. In the past few years we have been actively involved in redesigning our first level highly enrolled courses in statistics, including a) statistical literacy, b) introductory statistics, c) biostatistics, and d) engineering statistics. In each re-structured course we have transferred responsibility for learning to the students, and provided hands-on, collaborative group work in one or two computer labs each week. We will describe some of the key ingredients of these redesigned courses that we believe have resulted in enhanced student performance, increased student satisfaction and cost savings in our instructional program.

We have had a cadre of instructors at Penn state that have been systematically involved in restructuring introductory courses. These include the following:

Stat 100: Statistical Concepts and Reasoning—introduction to the art and science of decision making in the presence of uncertainty.
Stat 200: Elementary Statistics---basic statistics, probability, binomial and normal distributions, statistical inference, linear regression and correlation.
Stat 250: Introduction to Biostatistics—statistical analysis and interpretation of data in the biological sciences; probability, distributions, statistical inference for one- and two-sample problems.
Stat 401: Experimental Methods in Engineering-- random variables, probability density functions, estimation, statistical tests, t-tests, correlation, simple linear regression, 1-and 2-way analysis of variance, quality control methods.

Enrollments per academic year in these four courses are around 1000 for Stat 100 (basically a course in statistical literacy), 2500 in Stat 200 (mostly from the social and behavioral sciences), 300 in Stat 250, and 475 in Stat 401 (basically, a course in engineering statistics).

In general, our goals have been to

a. provide a learning environment that is student-oriented rather than teacher-oriented.
b. incorporate more active student participation and frequent hands-on experience with data analysis and interpretation of concepts.
c. use computers for testing, web-based courseware, and resource-rich course web sites.
d. provide timely feedback to students in their problem-solving tasks, analysis, and interpretation.
e. increase opportunities for collaboration with other students and with the instructors.
f. incorporate group work in various ways, including low-stakes quizzes and projects.

2. Common Concerns, Tools and Practices. With costs for higher education rising, state support for universities weakening, and greater accountability being demanded by the general public, there is a need to find ways to maintain or enhance the quality of instruction while keeping cost down. Some ways we have found to do this include the following:

a. use technology, especially computers in computer labs, appropriately
b. adopting innovations in instructional strategies (Readiness Assessment Tests and low-stakes on-line quizzing, and collaborative group work, for example)
c. assign greater responsibility for learning to students.
d. take advantage of undergraduate student expertise to assist in instruction.

We now discuss how we have implemented our approach in detail.

Computer Labs. We are fortunate in having computer labs available to provide students with hands-on applications of statistical techniques and for collaborative group work. For our largest enrolled course (Stat 200) there are 16-18 sections of labs, each meeting two times per week, with 82 PC's per lab. For the other (smaller) courses we use labs with 60 PC's . Every student has access to a PC, although they may work in pairs or groups of four.

Technology. Technology-wise, we conduct classes in computer labs, link to the Internet, use the Minitab Statistical Computing Software, and on-line quizzing software called 'TestPilot'.

Readiness Assessment Tests (RATs). Readiness Assessment Tests have three major components, depending on an instructors use of them: (i) an individual component, (ii) a group component, and (iii) an appeal process. They were developed by Larry Michaelson etal [1996] as an instructional and assessment tool. Students are given reading assignments before classes and prior to instruction on the material. The goal of the reading is for students to learn some of the basic concepts of the course on their own. After the reading assignments students come to class and take a RAT, made up of true/false and multiple choice questions. These questions should test knowledge and understanding of general concepts and principles rather than small, detailed facts. The goal of the individual RAT is to ensure accountability and understanding of the reading assignments. RATs are usually given in the Large Group Meetings (LGMs aka Lectures). Students take the individual RAT first and turn it in and then immediately re-take the same test as a group (previously set up) of three to five. The goal of the group RAT is for students to help one another comprehend ideas that they may not have gotten on their own. If the instructor chooses to do so, students are allowed to appeal any incorrect answers based on the quality of the question or a justification for their answer choice. Each student receives an individual and group grade for each RAT. The instructor uses the feedback from the individual and group RAT scores to determine where students still have misconceptions or misunderstandings. The concepts that students did not get on their own can be used to guide and inform instruction. The feedback helps the instructor focus instruction and activities on application of the course content rather than spending time covering concepts students can easily obtain through self-directed reading and learning. Course activities are typically completed in pairs or groups. The RATs and the content covered on them are used as a means to prepare students for the application of the content in problem-based activities.

RATs cover 'natural units', usually one or two chapters in the text. In a given semester four to six RATs are given. RATs provide a powerful motivator for students to read material prior to classes (since it is a major component of their grade) and to keep up with work on a regular basis rather than trying to study at the last minute before an exam.

Projects. Students are given two substantive projects, the first about the fourth week of the semester and the second during the last two weeks of the course. The first is moderately well-structured and the second is fairly unstructured, with just general guidelines. A survey is developed especially for the second project, cooperatively between the students and instructors. Students are given a scenario, like "'A President of a music company has experienced a downturn in his business. He asked the marketing department to collect data on aspects related to music, and the Director of Marketing contacted Penn State to do a survey of college students for them". Naturally the Department of Statistics was contacted to do this. So, with the help of the students a survey is created. As a group project, students are asked to analyze and, interpret the data and then prepare a report for the Director of Marketing at the company. The requirements imposed on the project specify that students are to perform at least six different statistical techniques in their analysis (e.g., compare two means, two proportions, regression, chi-square test, analysis of variance, etc.).

Lab Quizzes. In the last five minutes of the computer labs students take a short lab quiz on-line. There are usually 5 - 8 multiple choice items on the quiz and the students may consult with their lab partner or all of their group members (three or four per group) in answering the questions. The purpose of this, of course, is to encourage 'students teaching students'.

Undergraduate Interns. We have found that using undergraduate students as aides in our large beginning courses has many benefits. We use them

a. in computer labs to 'handle' the instructors' computer with overhead projection, freeing the instructor to interact directly with students as they work on activities.

b. homework graders.

c. as 'assistants' in other ways, including distributing handouts, providing computer help, and, in some cases, to help students with statistical issues.

3. Details. We elaborate on our approaches by describing in greater detail what we do in two of our courses—Stat 100 and Stat 200. The other two large enrolled courses are managed similarly to Stat 200.

A. Stat 100. Stat 100 is intended to help students understand basic principles of statistics to enable them to be better citizens in society, so the course emphasizes statistical literacy, and to sort out the useful and the accurate from the useless and the misleading. The aim of the course is to

- equip students with the statistical tools and concepts that will enable them to make their own interpretation of results emerging from surveys and studies
- enable students to read surveys and scientific studies with a critical eye
- provide students with experience in evaluating surveys and scientific studies so that after they leave the university they can be better informed citizens in our society and make their own judgments on critical issues involving uncertainty.

The course, offered for three credits, is intended for general liberal arts students as an introduction to statistical literacy and fulfills a general education requirement. It is less technically oriented than Stat 200, with the major difference being in the extent of coverage of statistical techniques and the mathematical level of the course. It is not taken by students majoring in Statistics. Each semester two large classes are offered, with enrollments of 240 each, so that in an academic year approximately 1000 students are enrolled in the course.

The course had changed little since its introduction many years ago, except for the replacement of the text initially used by an excellent new one that meets the course objectives content-wise. The course had been conducted in a prototypical manner, with three lectures per week, limited opportunities for class discussion, and a minimum of hands-on collaborative group work. A major drawback of this paradigm is the lack of regular practice in critiquing and evaluating relevant observational and experimental studies as part of the pedagogical process.

Our earlier experience with restructuring the slightly more technical and higher level elementary statistics course--Stat 200--lead us to consider implementing several of the practices (innovations) adopted in it. The instructional goals differ, however, so that blanket adoption of them was not appropriate. For Stat 200, technology played a major role, particularly statistical computing and on-line testing in computer labs. For Stat 100 the development of computational skills is secondary to statistical literacy issues, and hence this became the main focus of our redesign of the course. Specifically, we replaced the traditional class with an enrollment of 240 consisting of three lectures per week with an instructional mode giving students greater responsibility for learning. The new model has

a. a three day (weekly) cycle consisting of an overview lecture in **Class Meeting 1 (CM1)** , laying out the goals for the week and highlighting the main concepts,

b. Readiness Assessment Tests, discussion of assigned homework, and/or small group activities in **Class Meeting 2 (CM2)**; and

c. three breakout classes in **Class Meeting 3 (CM 3)**, formed by splitting the 240 students in the class into three Collaborative Labs (CLs) of size 80, dedicated to group collaborative work (groups of size 5) on activities (short, directed problem-solving efforts), mini-projects (full-period, less directed tasks), and full-size projects (4 per semester) encompassing the four major course modules.

For Class Meeting 3 (CM3), one of the three breakout sessions is led by the course instructor with a TA present, after which the TA becomes the 'facilitator' in the other two CM3's. For the most part, students are working on their own in these CLs with the instructor/TA available mostly to maintain a presence and to answer questions that might arise. In this way, the total class time required of the instructor is kept at 3 hours per week. An undergraduate intern is trained and is available as an assistant in each of the three CLs. Evaluation of the activities and projects in the CLs is student-graded. Assignments done by a group are passed to two other groups for grading. In those cases where there is a divergence of scores assigned by the two groups the TA serves as a 'referee' to determine the final grade. This procedure serves two important purposes: i) it reduces instructor/TA grading time and ii) it functions as a 'reinforcement' to mastering concepts by the students, since they need to understand the topics to be effective graders. For the course, a packet of materials has been created and made available to students. The packet consists of a) Activities, b) Case Studies, c) 'Thought Questions', d) Mini- and Full-Scale Projects, and e) Study Guides.

Since the primary goal of the course is to enhance student literacy of applications of statistics in their lives as informed citizens of our society, we created a pre- and post project assessing the seven critical components of critiquing outcomes of surveys and experimental results. Some of these seven components are the following:

- The source of the research and of the funding
- The researchers who had contact with the participants
- The individuals or objects studied and how they were selected.

  …….

  …….
- The magnitude of any claimed effects or differences (practical vs. statistical significance)

This project is given at the beginning of the semester and again at the end of the semester to determine what gains have been made.

B. Stat 200.

As noted above, our Stat 200 course consists of two types of classes: Computer Labs and Large Group Meetings (LGMs). The two computer labs and one LGM each week are sequenced either as *Lab>LGM>Lab* or as LGM>Lab>Lab, depending on the time the class is offered. We use the recently published text by Utts and Heckard "*Seeing Through Statistics*" as the main resource. It has more emphasis on statistical literacy and interpretation than most existing texts.

With just one LGM each week the amount of time available for formal lecturing is minimal. There are about 6 or 7 lectures in the semester in LGMs. There is some lecturing in labs--up to 10-15 minutes maximum. How do students learn the concepts if we don't 'lecture'? By giving students the responsibility for learning the concepts. Students are given weekly reading and homework assignments and to 'motivate' them to do it in a timely fashion they are given Individual and Group Readiness Assessment Quizzes (RATs), which are described in some detail above.

Students' understanding of the course material is reinforced in the computer labs through individual and group work on activities designed specifically for this purpose. They are given quizzes in about 2/3 of the lab sessions. They are assigned two group projects, the first early in the semester and the second in the last two weeks of the course. Specifically, here is what students are told:

i. There will be Individual and Group Readiness Tests (RATs). These will be given in the Large Group Meetings (LGMs). The Individual RATs will consist of 12-14 multiple choice questions. The Group RAT will be on the same set of questions and be given immediately after the Individual RAT. About half of the items will be on previously discussed topics and the other half on new material not previously covered.

ii. Lectures will be given in the LGMs on days when there are no RATs.

iii. In the computer labs you will be working on 'activities' in pairs or in small groups to apply what was learned in the readings. Lab quizzes consisting of 5-8 questions will also be given in 15-20 of these labs. Items on the Lab Quiz will be based on 1) the activity for the lab and 2) general concepts being illustrated by the activity.

iv. Some classes and labs will be reserved for work involving the integration of course content, such as evaluating scientific articles and completing group projects.

v. Three hours of 'tutorial' sessions are set up weekly, to provide assistance to students having difficulties with the concepts and/or with homework. They may also be used to get answers to questions about the reading assignments prior to the RATs. Tutorials are not designed as 'lectures' or to solve homework problems, but rather to give students feedback on the material.

Our computer labs have PC's, linked to the Internet, which are loaded with Minitab Statistical Computing Software, and on-line quizzing software called 'TestPilot' [see http://www.clearlearning.com]. In the last five minutes of class students take a short lab quiz on-line using TestPilot. Students may consult with their lab partner or all of their group members (three or four per group) in answering the questions. The purpose of this, of course, is to encourage 'students teaching students'. Student responses on the lab quizzes are sent directly to a file, the results are summarized and made available to instructors to assess student understanding of the concepts covered in the lab. At the next class meeting instructors review any concepts that students did not grasp well.

In the first computer lab we take about 4 or 5 minutes to look at what we will be doing during the week. This may be followed by a brief overview of the main concept covered in the lab. We have an excellent course management system which provides students with all of the activities they will be doing during the week, including reading and homework assignments, datasets to be used, lab quizzes scheduled, etc.. Here is the 'This Week' page on the web site for Week 11: March 18-22, 2002:

This Week

Week 11: Monday, March 18, 2002 - Friday, March 22, 2002

Lab 19 March 18, 2002.

1. *What To Read* Sections 13.1-13.2 before today's lab. Read Sections 12.5-12.7 (covered on RAQ 3) before Wednesday's LGM and13.3-13.5 before Friday.
2. *Exercises* Homework assignment 9: Chapter 12--12.17, 12.21, 12.28, 12.30-12.32, 12.35. Chapter 13: 13.1-13.2, 13.5, 13.6, 13.8. Do at least 6 problems from Chapter 12 and 4 from Chapter 13.
3. Today's lab activities will be concerned with inference for one population mean, including paired comparisons: confidence intervals and testing,
4. *Datasets* Use the data from the **Spring Survey**
5. *Study Guides* RAQ 3 is scheduled for Wednesday's LGM. Take a look at the Study Guide for Chapters 9, 10, 12 and 13
6. *Test Pilot* Take the Lab Quiz

LGM Wednesday March 20, 2002

1. RAQ 3 will be given today. Please bring a #2 pencil. Coverage for the RAQ is Sections 9.3-9.7, Chapter 12, and Sections 13.1-1.2

Lab 20, Friday March 22

1. We need to form groups for Project II. Please login on to coursetalk (an asynchronous web-based communication software, like a chat room) and check your membership. If you are not listed, contact Yudan or Eliza (TAs) to correct the situation. If you want to change groups, we can consider that as well. The URL is coursetalk.cac.psu.edu
2. *Activities* Today's activities will be on two proportions and two means. We will look at whether or not the proportion of students who would have sex without being in a committed relationship is the same or not for whose who have driven under the influence and those who have not. In comparing two means we will get confidence intervals and tests.
3. *Datasets* Use the data from the Spring Survey
4. Use the *t* -table* of multipliers
5. One of the questions on the Lab Quiz asks you what type of general 'theme' is most appealing to you. Confer with members of your group to express your preference.
6. *Test Pilot* Take the Lab Quiz

The schedule for Monday's lab is fairly typical. After the brief look at the week's agenda, students begin working. As soon as they enter the lab, they login on their PC, get on the Internet, bring up the week's agenda, and then open Minitab. They return to their 'This Week' page, move their cursor to 'Datasets', and click on 'Spring Survey', then copy and paste it into their Minitab Worksheet.

'Spring Survey' refers to data collected into a data file from the students during the first week of classes. It consists of about 40 items on a variety of variables, such as gender, grade point average, race, eye color, height and weight, ideal height and weight, drinking habits, etc. These data are then used throughout the semester to illustrate statistical concepts learned in class. A student intern hired by the department passes out an 'activity' for the students to work on, while a teaching assistant at the front of the lab does the same thing the students are to do, but on a delayed basis. The instructor walks around the room prepared to help any students having problems or to answer questions.

After the students have worked (individually, in pairs or in groups of four) about twenty-five minutes on their activity, they are asked for their attention again and we discuss the activity. Here is an example of a short activity we use to learn about the sample standard deviation and the empirical rule: (Note: we have already studied confidence intervals):

Activity #013

Age Measurements for the Shroud of Turin

The Shroud of Turin is a linen fabric that, since 1354, has been claimed to be the burial garment of Jesus Christ. In efforts to establish its authenticity, there has been an enormous amount of scientific testing performed on this object. In one study, several small strips were sent to labs in order to perform radiocarbon dating, a process by which the age of items can be estimated (with some degree of uncertainty). Four of the strips were sent to a lab in Arizona in 1988, resulting in the following estimates for their date of origin:

1397, 1298, 1382, 1287 AD, and their dated age in 1988

date of origin	Age in 1988	$(x - \bar{x})$	$(x - \bar{x})^2$
1397	591		
1298	690		
1382	606		
1287	701		
$\Sigma x =$			$\Sigma(x - \bar{x})^2 =$

Use the above table to guide you through the hand calculation of the standard deviation.
Using the empirical rule, does it seem likely that the fabric could be from the time of Jesus Christ's death?
In Minitab, double-check the standard deviation that you calculated. To do this, enter the four data points in a column, then obtain the descriptive statistics (as you've done before).
Now find a 99% confidence interval for the population mean year of origin for the strips sent to the Arizona lab. Go to Stat > Basic Statistics > 1-Sample t. In the "Variables" area, select the column into which you entered the data, then make sure that the confidence interval level is 99. Interpret the range you obtain. What do the results suggest about the authenticity of the shroud (based on the strips sent to the Arizona lab)?

If you think it would probably have been a good idea to send strips to other labs for dating, you are right -- this was actually done. If you are interested in these and other data and information, the Web has many sites dedicated to studying the Shroud of Turin, such as The Shroud of Turin Website, The Shroud of Turin: Genuine artifact or manufactured relic? (by Jack Kilmon), and The Council for Study of the Shroud of Turin. You may find the controversial debates on these sites to be interesting.

About 15 or 20 minutes before the end of class, we get the students attention and discuss the work they have done. This includes going over their solutions and interpretation of the results. It also provides students with valuable feedback on the concepts they have been working on and their understanding of them. In the last 6-8 minutes of the class students take the on-line lab quiz covering the day's concepts and the activity they were working on. Instructors found that by putting questions on the lab quiz about the activity that students were more attentive to their work. As for the individual and group collaborative activities, we were pleasantly surprised at the students' positive reactions to not being lectured to and instead being able to work in groups in the labs to apply what they had learned from the resource

4. Impact on Student Learning

We developed several assessment instruments prior to restructuring the course:

i. A content knowledge test consisting of 18 items was developed prior to the restructuring (in 1998/99). It was administered at the beginning and end of the spring 2000, fall 2000 and spring 2001 semesters. During the spring 2000, two sections were taught in the traditional format (n=340) while one section was taught as a pilot using the revised Stat 200. In the fall and spring semesters of 2000/01 all classes were taught using the new format.
ii. A 20-item test on 'choosing the appropriate statistical technique' from a set of 10 was created in 1996/7 and has been used as part of final exams every semester since then.
iii. Statistics on the number of D's and F's, gpa's ,and dropouts were compiled for the 5-year period 1996/97 through 2001/02
iv. Assessment of student performance in follow-up courses using a subset of the content knowledge test is on-going.
The results:

a. The pilot and redesigned classes outperformed the traditional class on the final test of content mastery by 10% to 13% (60%: traditional class, 66% in the pilot class, 68% in the redesigned classes), The improvement in performance in the redesigned class was greatest on concepts. On technical aspects (working with formulas and reading tables for example) the traditional class performed marginally better.

b. Students in the restructured course were able to identify the correct statistical technique to use about 86.5% of the time, about 11% better than the 78% correct rate for students in the traditional course. This is viewed as a consequence of lab work.

c. The percentage of students receiving a D, F, or dropped the course decreased from a rate of about 12% in the traditional course to about 9.8% in the restructured course, a decrease of about 18%. The average gpa was essentially unchanged: 2.974 in the traditional course and 3.015 in the new course.

5. Other Comments.

i. Assessment of student understanding of concepts using RATs has proven to be very effective in detecting areas in which students are not grasping the concepts, thereby enabling corrective actions to be taken in a timely manner, and in preparing students for higher level activities in the computer labs than previously. As a result, students have been helped in building skills, as the evidence of the pre- and post-test shows. The web page enables more rapid feedback to students, another crucial element in the learning process.

Student perception of the importance of RATs is evident in the results from Innovation and Quality (IQ) survey data where the majority of students (55%) rated the RATs as one of the most important aspects of the class. Seventy-five percent of respondents believed that periodic RATs help them keep up with the readings and that they were vital for their learning and understanding of the content. As voiced in focus groups, students felt that the RATs helped by promoting recognition of holes in their understanding. In addition, students liked the opportunity to work in groups and interact with others in the class. Most students emphatically suggested keeping the Readiness Assessment Testing as part of the course.

ii. As for the individual and group collaborative activities, we were pleasantly surprised at the students' reaction to not being lectured to and instead being able to work in groups in the labs to apply what they had learned from the resource

6. Department Head Perspective

The restructuring of the courses has resulted in savings to the department of over $125,000 and enrollments have increased about 20%, mostly in the basic Stat 200 course. The use of Master's of Applied Statistics and undergraduate interns in the labs provides valuable experience for students in our majors. The instructors responsible for the course changes have indicated that the enhanced student performance may be a consequence of the Hawthorne effect. Whether this is so or not is irrelevant, in that performance has improved and this is a result we were hoping for! The decreased usage of teaching assistants in Stat 200 permits greater flexibility in assignment of duties to them, a substantial departmental benefit. Finally, this also enables international students to develop language skills prior to being assigned teaching responsibilities.

References.

Felder, R. M. and Brent, R.(1996), "Navigating the bumpy road to student-centered instruction," *College Teaching*, **44**(2), pp. 43-47.

Springer, L., Stanne, M.E., and S.S. Donovan (1999), "Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis," *Review of Educational Research*, **69**(1), pp. 21-51.

Michaelson, L.K., Black, R.H., and Fink, L.D (1996)., "What every faculty developer needs to know about learning groups." In L. Richlin (Ed), *To improve the academy: Resources for faculty, instructional and organizational development*, New Forums Press.

Michaelson, L.K., (1999), "Myths And methods in successful Small Group Work," *The National Teaching and Learning Forum*, **8**(6), pp 1-7.