Available at
http://pvamu.edu/pages/398/asp

Vol. 1, No. 1 (2006) pp. 61 – 81

Applications and Applied Mathematics (AAM):
An International Journal

REVERSE-TRADITIONAL/HANDS-ON: AN ALTERNATIVE METHOD OF TEACHING STATISTICS*

Aliakbar Montazer Haghighi
Department of Mathematics
Prairie View A&M University
P.O. Box 519, Prairie View, Texas 77446, USA
E-mail: amhaghighi@pvamu.edu
Tel: (936) 857-2026, Fax: (936) 857-2019

Roozbeh Vakil
Department of Mathematics
St. Cloud, MN 56301, USA
E-Mail: rvakil@stcloudstate.edu
Tel: (320)308-4964, Fax: (320) 308-4269

Johnson K. Wetiba
Department of Mathematics
P.O. Box 519, Prairie View, Texas 77446, USA
E-mail: jkwetiba@pvamu.edu
Tel: (936) 857-3807, Fax: (936) 857-2019

Received December 15, 2005; accepted March 18, 2006

Abstract

This paper presents a method for teaching statistics called the Reverse-Traditional/Hands-On (RT/HO) method. The method includes student-generated data, use of a statistical computing package such as MINITAB to introduce statistical terms and methods, and a statistical research component. An analytic comparison of students' successes in courses using the Traditional and RT/HO methods is presented.

Key Words: Active Learning Environment; Instruction using Technology

AMS Class: 97D40

1. Introduction

The Traditional method of teaching consists mainly of giving lecture by the instructor and expecting students to be cognitively active but physically inactive, except for note taking. Most students of any age cannot maintain such behavior for a long period of time Cangelosi (2003). One aspect of the Traditional method of teaching is that it has a tendency to view students as passive learners (Steinhorst and Keeler, 1995) because it does not engage them actively.

* This research was partially supported by the National Science Foundation under Grant # ESI-9553502, A. M. Haghighi, PI. The opinions expressed do not necessarily reflect the views of the Foundation.

Extended abstract of this paper is included in the proceedings of the International Conference of Informatization of Society, Astana, Kazakhestan, October 2004.

It is now widely accepted that one way to obtain attention of the new generation of students and make them interested in mathematics and statistics is to get them actively involved through hands-on activities. It is also well accepted that mathematics and statistics should be taught more as an experimental science and less as traditional mathematics. It can even been seen in quite number of research papers that motivation of a paper is mentioned as a real world problem or experience.

Now days, with technology generously available, assignment of laboratory for students is practiced even more than ever. However, the laboratory times is used for practice time rather than teaching. It is perhaps a type of active learning. There is an extensive literature, papers and books, particularly in statistics, during the past decade considering active learning (in its broad sense of definition). See Scheaffer et al. (1996) and references listed there in, for instance. However, it is interesting to note that after long years of talking about infusion of technology in teaching and practically technology being a part of teaching across the curriculum, now we see a flash back. Effectiveness of technology in teaching is now under question. Patrick Allitt (2005), is one of those open critics criticizing American universities how they have wasted billions of dollars in the last decade or two on technology in teaching.

We, in this paper, are not disputing the creative methods some instructors use to teach and actively involve students in teaching and learning processes. We are speaking of the general Traditional method worldwide.

The method described in this paper engages students for the entire instruction session. Because of availability of computers, we emphasize concepts of statistics rather than computations. We are talking about teaching and learning using hands-on experience rather than lab practice. Using the RT/HO method, terms show up and concepts emerge for discussion only when they are needed. Statistical methodology is discussed when terms such as "*p*-value" or "*t*-value" appear on the computer screen. Additionally, required topics of probability are covered only when the concept of inference is discussed. The method puts the emphasis on learning and facilitating rather than teaching. The method allows students to shift from a focus on statistical methods to a focus on statistical thinking as Hahn and Hoerl (1998) state. The RT/HO method avoids mathematical formulas (unless the level of the course calls for them) and thus students can learn statistics without relying on mathematical formulas. We conjecture that this method may be applied to teaching other science and mathematical subjects. This conjecture is under study.

There is ample evidence that the current status of teaching statistics at all levels is not satisfactory. Statistics educators, researchers, and administrators have been calling for and demanding drastic changes in the way statistics is being taught and learned (see Hogg 1991, Watts 1991, Cobb 1993, Garfield 1993, and Moore, Cobb, Garfield, and Meeker 1995, for example). Recent findings in statistics teaching and learning indicate that an active learning environment in the context of discovery methods improves students' understanding significantly (see Garfield 1993, Keeler and Steinhorst 2001, for example). In discovery methods, students are guided to discover concepts on their own through a series of, carefully designed, step-by-step tasks. Despite this compelling evidence regarding students' conceptual understanding, many educators are concerned about the amount of content that can be covered using discovery methods. As a result, some educators move towards leaner contents in statistics courses (Hogg 1992, Moore 1992) and some compromise some computational aspects of statistics (Steinhorst and Keeler 1995). The method described in this paper actively engages students in the process of learning but the formation of concepts is not as time consuming as discovery methods and therefore allows covering the entire standard content in an introductory statistics course.

Although the RT/HO method was developed to teach any mathematics course, it was first used as the means of instruction of statistics in a five-week summer program, called Benedict Pre-college Statistics (BP STAT), in 1996, funded by the National Science Foundation (NSF), with 29 high school students (9^{th} through 11^{th} grade) and was repeated in 1997 with 30 participants, five of them being from the previous year.

The importance of statistics has never been as great as it is today. Statistics has vast application in virtually all areas of science and industry and thus human endeavor. But attention towards statistics goes beyond its benefits, it may very well be used (abused) by those who wish so. The National Council of Teachers of Mathematics (NCTM), in *Principles and Standards for School mathematics* (2000), warns that "Statistics are often misused to sway public opinion on issues or to misrepresent the quality and effectiveness of commercial products" (p. 48). To counter this pitfall, students need to learn to analyze the data properly. To be an informed citizen has become a harder task than ever. To address this concern, educators need to make data analysis available to students as early as possible. The good news is that statistics is not a subfield of mathematics (Moore 1992), it is a methodological discipline and has its own distinctive concepts and reasoning; as such, learning it at the introductory level could be done using limited mathematical skills.

Learning statistics for all has become so important in recent years that all related organizations have acknowledged this fact and each one has tried to do something about it. NSF, for its part, funded a major project for the development of "Data Driven Curriculum Strand for High School Mathematics" and the "Consortium for Policy Research in Education (CPRE)." The American Statistical Association, recipient of this grant, has stated in its report that "statistics provides an interesting practical use of mathematics that students find enjoyable and challenging and which will introduce them to research."

In the BP STAT project, we observed that students demonstrated positive attitudes toward collecting data and were excited to see the results on the computer screen. The RT/HO method was used at Prairie View A&M University (PVAMU) in the fall 2003 and the spring of 2004. We have, consistently, observed positive attitude on the part of students. We also, anecdotally, noticed that students retain the information for a long period of time. These observations concur with the general goals of reform-minded approaches in introductory statistics courses (Keeler and Steihorst, 1995).

The name "Reverse-Traditional/Hands-On" (RT/HO) was chosen because in this method the sequential order of the Traditional method is reversed and the instruction is built upon hands-on activities. In a Traditional format of instruction, an instructor usually starts with an overview of the content to be covered followed by some examples to clarify the mentioned content and finally gives the solution. The RT/HO method starts with raw data (collected by students in most cases and/or using existing real data). After data is entered into a computer, students see answers to several standard parameters on the screen of their computers. At this point the instructor explains the terms on the screen. Gradually, the content to be covered in a given session emerges. This method requires students to work with computers in two distinct ways. First, as an integral part of instruction, students use computers to learn statistical terms and concepts. Second, they use computers to conduct statistical research, which starts from scratch and ends with write-up of a report. Hands-On refers to both of these activities.

2. Description of the Method

The RT/HO method consists of two components: instruction and research. Both components use cooperative learning. To form and monitor the groups, guidelines such as those suggested by Goodsell et al. (1992), Garfield (1993), and Dietz (1993) are used. Thus, generally, students in an elementary statistics class divide into groups of five and maintain membership in their chosen group for the entire course.

2.1 Instruction Component

To prepare the lesson plans, recommendations made by Davidson (1990), for instance, could be used. To begin the instruction, the instructor will ask the students to measure a measurable attribute in groups or in class such as height, age and the room temperature. Then, students are then asked to enter the data into a statistical computer software package such as Minitab. The instructor guides the students through a sequence of steps selected from the software's menus to obtain the basic statistical characterictics such as mean, median, and standard deviation. Then the instructor starts defining and explaining terms that have appeared on the screen and what the numerical values mean. This procedure continues during an instructional session and through the duration of the course. In summary, the idea is to let the computer spit out analyses and then indicate what the output means. The students are guided in the interpretation of the output through questions posed by the instructor.

The computer activities provide the opportunity for students to get a coherent picture of related terms and statistical concepts as opposed to isolated and non-related terms and concepts, which occurs often in Traditional methods of teaching. In other words, the method avoids defining a list of terms in isolation and in advance. Terms are introduced and defined as they are needed. New terms and concepts emerge out of word-problem situations in most cases, as instruction moves ahead. Formulas and their derivations may be mentioned and provided based on the level of the course.

There are two stages for instruction using the RT/HO method: basic, and advanced. It is important to notice that coherence in presenting the materials is central to this method at both stages, i.e., each lesson consists of a collection of related statistical concepts.

At the basic level, for example, students will be exposed to the basic terminologies as a coherent whole such as mean, mode, median, histogram, pie chart, and distribution via a series of computer activities such as navigating through statistical software.

The advanced stage entails two layers; in the first layer students get their first exposure to the statistical analysis using computer statistical software like MINITAB. In the second layer, using ANOVA, students will learn statistical inference. Statistical thinking and conceptual understanding will be facilitated at this stage. Manual computations of formulas will be practiced and comprehended at this level when it is appropriate. What this means, when the instructor is explaining the results shown on the screen, he/she may attempt to point out formulas used and briefly outline the derivation if necessary and appropriate.

The method was tested both at the high school and college levels. Below we provide examples of instructional activities for each stage of the RT/HO method. Examples 2.1.1 and 2.1.2 were used in the BP STAT project and example 2.1.3 was used at Prairie View A&M University (PVAMU) in a statistics class.

Example 2.1.1. A group of five collected data on their heights. The observed heights were 62, 62, 65, 68, and 68 inches. A group member entered the data into the first column of the MINITAB's Worksheet, called C1. (We should repeat that although we have used MINITA, any statistics computer could be used.) Then, from the menu, choosing the sequence of options "Stat," "Basic Statistics," and "Descriptive Statistics," the following output appeared on the MINITAB's screen's Session (above the Worksheet):

```
Variable            N       Mean     Median     TrMean      StDev    SE Mean
C1                  5      65.00      65.00      65.00       3.00       1.34

Variable      Minimum    Maximum         Q1         Q3
C1              62.00      68.00      62.00      68.00
```

At this time, the instructor asked the following questions:

- What can be said about the height of the students in the group overall? (Hinting the average)
- How many were taller and how many were shorter than the average? (Hinting the median)
- What was the height of the tallest person? (Hinting the maximum)
- What was the height of the shortest person? (Hinting the minimum)
- How could you relate these questions to the numbers on your screen? (Hinting terms and their numerical values)

The instructor noticed that the group engaged in a heated discussion to relate their guesses to the numbers on the screen. As the instructor started to define each term, Variable, N, Mean, Median, Tr Mean, StDev, SE Mean, Min, Max, Q1, and Q3, students showed enthusiasm for their correct guesses and pondered about new terms. After this preliminary discussion, the instructor continued with the following question:

- If you were to choose only one of the terms "mean," "median," or "mode" to explain the height of the group, which one would you have chosen?

The answers to such questions gave the instructor a chance to explain terms in more detail and compare terms with similar meaning.

Example 2.1.2. We bought five apples, of the same kind, one each from five different stores. One apple was given to each group of participants. Each group member of a group was asked to take a bite of the apple given to the group and assign a number between 1 and 5 to the sweetness, 1 for not sweet at all, 2 for very little sweet, 3 for sweet, 4 for very sweet, and 5 for excellent. Once the data was ready, students were instructed to key in their collected data into the computers. Following a few menu driven steps, with little or no help from the instructor, the results were shown on the screen. The computer generated a table of terms and a list of related numbers beneath them to view. Since one group had only 5 members, the instructor joined that group to make five groups with six members in each, total of 30 members for this experiment. The idea was to see if the sweetness of apples from different stores was the same. In MINITAB's Worksheet, groups were denoted as C1, C2, C3, C4, and C5. Groups entered their data and exchanged. Thus, each group had the following data:

C1	C2	C3	C4	C5
2	1	5	1	4
3	5	1	2	5
3	3	2	1	5
4	4	1	1	5
4	2	3	2	3
5	1	2	1	4

Participants were then asked to go to MINITAB's menu and do Stat, ANOVA, and One-way (Unstacked). The following appeared on their computer screen "Session":

One-way ANOVA (Unstacked): C1, C2, C3, C4, C5

```
Analysis of Variance
Source     DF        SS        MS        F        P
Factor      4     31.33      7.83     5.62    0.002
Error      25     34.83      1.39
Total      29     66.17
                                   Individual 95% CIs For Mean
                                   Based on Pooled StDev
Level       N      Mean     StDev  --------+---------+---------+--------
C1          6     3.500     1.049                (-----*------)
C2          6     2.667     1.633          (------*-----)
C3          6     2.333     1.506        (------*-----)
C4          6     1.333     0.516  (------*------)
C5          6     4.333     0.816                     (------*------)
                                   --------+---------+---------+--------
Pooled StDev =    1.180                  1.5       3.0       4.5
```

As before, new terms such as ANOVA, One-way, DF, SS, MS, F, and P were defined. By the end of this activity, students were ready to conduct analysis of this data in their research project.

Again at this time, the instructor asked the following questions:

- Since students already know the "average", the instructor asked:
  If you had the average of your four tests given to you as 75 and you had three out of your test papers graded available, could you find the score of the fourth? (Hinting the degrees of freedom)
- What are we trying to do with apples? (Hinting the variability in sweetness)

- What are we doing with data? (Hinting the analysis of data)
- Could we average any group of numbers to find the average for that item? (Hinting the pooled standard deviation)
- How could we tell if there is a significant difference in sweetness of apples from different stores? (Hinting the *p*-value)

The instructor explained that since the *p*-value was very small, 0.002, the inference was that there was a highly significant difference in mean sweetness somewhere among the five groups. This was to show the students how to interpret some numbers and terms shown on the screen.

For further advanced discussion we could have continued the analysis to find out the location of the difference.

Example 2.1.3. This lesson was built upon example 2.1.2. In BP STAT (high school level) program, Example 2.1.2, we did not mention the formulas used for students. However, at PVAMU (college level) since students had a Calculus course, we did explain how the values of SS shown on the screen would have been obtained as follows:

$SS_{among} = \sum_{i=1}^{g} n_i(\bar{x}_{i\cdot} - \bar{x})^2$, $SS_{within} = \sum_{i=1}^{g}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_{j\cdot})^2$, Total $= \sum_{i=1}^{g}\sum_{j=1}^{n_i}(x_{ij} - \bar{x})^2$, g = number of groups, n = total number of members, df = degrees of freedom

$\sum_{j=1}^{5} x_{1j} =$	T_1	T_2	T_3	T_4	T_5	$\sum_{j=1}^{5} T_j =$	T	T^2	$T^2/30$
	21	16	14	8	26		85	7225	240.83
	T_1^2	T_2^2	T_3^2	T_4^2	T_5^2			$\sum_{j=1}^{5} T_j^2$	$\sum_{j=1}^{5} T_j^2/6$
	441	256	196	64	676			1633	272.17

	$C1^2$	$C2^2$	$C3^2$	$C4^2$	$C5^2$	Total	
	4	1	25	1	16	47	
	9	25	1	4	25	64	
	9	9	4	1	25	48	
	16	16	1	1	25	59	
	16	4	9	4	9	42	
	25	1	4	1	16	47	
Total	79	56	44	12	116	307	$= \sum_{j=1}^{5}\sum_{j=1}^{6} x_{ij}^2$

$g = 5$, $n = 30$, $df = 30 - 1 = 29$, $SS_{among} = 272.17 - 240.83 = 31.34$, $SS_{within} = 307 - 272.17 = 34.83$, Total = 66.17 (note a slight difference due to rounding)

$\bar{x}_{1\cdot}$	$\bar{x}_{2\cdot}$	$\bar{x}_{3\cdot}$	$\bar{x}_{4\cdot}$	$\bar{x}_{5\cdot}$
3.500	2.667	2.333	1.333	4.333

The inference discussed in this case was as follows. By looking at the sweetness averages as 3.5, etc., one cannot tell, with high probability, that all apples from different stores are the same. However, since from the *F*-table with 4 and 25 degrees of freedom and 95% confidence interval the value of $F_{0.95}(4, 25)$ is 2.76 and the fact that the p-value is 0.002, we conclude that there is a highly significant difference in mean sweetness somewhere among the five groups.

2.2 Research Component

To complete the course, students are required to conduct a statistical research project. The statistical research project is an integral part of the method. The research, starts from an initial question and ends with written report. The project allows the students to practice using the statistical software that they have learned and to interpret results they obtain using the software.

This component plays an important role in students' participation in the process of learning. Students are seriously involved in practically every step of their research projects, nothing is provided in advance. Students conduct their research in groups. Each team will choose its topic, make questionnaire, revise the questionnaire to make it appropriate for statistical inference, collect data, analyze the data, interpret results, and write a final report. This hands-on activity requires students to develop skills in working independently and as a team member and thus prepare them for real world situations. Each group gains the ability to choose a topic for a statistical research project through an introductory guidance by the instructor in which he/she will provide a general idea about statistical research. The group will then work on finding a research project. The instructor must approve the topic.

In the BP STAT project, most groups had a tendency to select topics that only allowed the use of descriptive statistics. They did not look for causes, estimation of parameters, hypothesis testing or other things that would lead them toward inferential statistics. This may be due to the type of instructions students received within their first week. To direct them toward inferential statistics, the instructor guided them by asking questions based on the purposes of their proposals and made general comments on the subject matters. The instructor's questions and comments usually called for seeking relations between the factors, correlations and truth. Based on these comments, students revised their topics.

We suggest the following time line for implementation of the research part of the method. The first week (or its equivalent number of days) of instruction will be spent on exploring statistics by giving students many examples and computer-assisted exercises. The topic selections are due at the end of the first week. Immediately after approval of the topics, students are asked to design questionnaires for their projects. They go through the same process of approval as for the topics. Here the instructor helps students to build upon previously designed questionnaires, which are basically only survey types, and improve them to include questions leading to concepts such as correlation, causal relationship and hypothesis testing so that they will be appropriate for inferential analysis.

By the end of the fourth week, designs of questionnaires should be completed and groups submit progress reports including the topics, purposes, objectives, and questionnaires. Distribution of questionnaires and collection of data requires some instruction about types of sampling. At this point students are instructed to choose their sampling types based on their purpose and types and geographic location of populations. Instruction now completes its fifth week. The required time for data collection will not be more than two weeks. It is done concurrently with other activities such as literature review.

The collected data will then be keyed into the computer and will be ready for analysis. During the process of analysis, students receive a great deal of assistance while they are using computer software such as Minitab. They are helped on how to do correlation analysis, ANOVA, and regression. All these are done via hands-on computer and using the menu options in the statistical software. This process takes about four weeks. Groups submit a weekly report on the progress of their analysis.

The final step in conducting the research is writing up the final report of the research. The report has the following format: Cover Sheet, Table of Contents, Abstract, Introduction (to include literature review, purpose and objectives), Methodology, Analysis, Conclusion, Discussion and References.

Note that all communications between the instructor and each group are dated and kept in a folder. The end result will be a portfolio consisting of everything from the beginning to the end and all dated. Groups are asked to submit their final projects on diskettes as well as on papers. Groups are required to present results of the researches.

3. Theoretical Perspective

In addition to the research component, the RT/HO method is comprised of three instructional principles: cooperative learning, reversing the Traditional order of instruction, and hands-on activities. We discuss the theoretical aspects of these principles below.

3.1 Cooperative Learning

In a study-team, students interacting with other members of the group will notice the possibility of arriving at a different conclusion. Reynolds et al. (1995) state that “These differences may give rise to conflict among the individuals in the group. Contradictions coming from others at a similar level of conceptual development serve to bring the cognitive differences into ‘sharp focus’, and thus lead to coordination that can resolve the conflict” (p. 20). In our experiences we observed that when students interact with one another, they can focus better and thus the in-class time is used very effectively. We also observed that some students may ask questions that the instructor cannot relate to, but other students see the relevance very well and can provide acceptable answer.

The cooperative learning component of the RT/HO method plays a major role in implementation of it in many ways including: data collection, data analysis, and decision-making. These interactions among group members lead them to heated discussion with one another and thus converting “cognitive differences” to “sharp focus” and then arriving at collective understanding.

3.2 Hands-On Activities

Conceptual understanding is the key to the learning of statistics and the ability to apply it properly in physical and social sciences. According to Hogg (1991), students will learn better if they find and generate their own data, formulate the hypothesis, design experiments and surveys, summarize data, and communicate findings with one another.

What engages students is real world statistical projects/examples that they can relate to. Just by throwing actual data into a text it will not do it. They can best relate to the data if they themselves originate the projects and go through the entire process of collecting data, analyzing data and inferring from the data. Hands-On activity is an important component of the RT/HO method. When students are involved in the process of conducting a project from the beginning they will more likely have scientific curiosity into the results because it is real to them. Snee (1993) puts this fact in sharp focus: “Collection and analysis of data is at the heart of statistical thinking, data collection promotes learning by experience and connects the learning process to reality” (p. 152).

3.3 Reverse-Traditional

Thinking mathematically (statistically) involves looking for connections and making connections builds mathematical (statistical) understanding. Without connections students must learn and remember too many isolated concepts and skills (NCTM Principles and Standards, 2000). Gathering, representing, and analyzing data can develop insight into other statistical and mathematical concepts. The RT/HO method is conducive to seeing the statistical terms in a related context .The “Reverse-Traditional” aspect of the RT/HO method provides a unique opportunity for students to see statistical terms as a coherent whole as opposed to isolated and not related concepts. We observed in example 2.1, since statistical terms and concepts are offered to students in a package, and the answers were on the screen to see, the instructor could guide students to make educated guesses. These mini conjectures which were prompted by the instructor's guidance and availability of answers on their computer screen, help students look at them in a holistic manner. In that example students were encouraged to compare the terms *mean*, *mode*, and *median* to argue which would better represent the height of the class if they were to choose only one. Later on they see how *mean* will give rise to other measures such as *standard deviation*.

4. Assessment

The Traditional evaluation of students’ performance entails in-class-written tests and an occasionally oral presentation. This evaluation does not reflect many aspects of students’ improvement in aspects of performance which is not easy to quantify such as attitude, higher order thinking skills, and problem solving (Franklin 2000). In an effort to address these difficulties and avoid controversial issues concerning grading, we expanded our evaluating apparatus to include: statistical thinking, attitude, cooperation, problem solving using computer statistical software to generate and analyze data, and ingenuity in topic selection appropriate for research as well as the process of developing statistical research from the beginning to the end rather than considering only the resulting report.

Since students work in groups and complete computer based activities, oral communication among them is very natural. This dialogue is a good opportunity for the instructor to walk around and listen to students' conversations. This valuable dialogue, i.e., informal discussion, prepares students for the final presentation of their research paper that is worth 10% of the total grade. The remainder of the grade is distributed as follows: quizzes 10%, two tests and a final exam (counting twice as much as a test) 40%, research paper 30%, and homework assignments 10%. Tests are a combination of traditional questions and questions that have to do with computer software being used.

To assess the BP STAT students' learning we collected a variety of formal and informal information. Since all activities were carried out in the groups, evaluations were conducted in groups as well as individually. Here are some examples of quizzes given to the BP STAT participants.

Example 4.1. The following appeared on the MINITAB's Session (above the Worksheet):

```
Variable         N      Mean    Median  Tr Mean     StDev   SE Mean
C1               5     2.600     3.000    2.600     1.140     0.510
Variable       Min       Max        Q1       Q3
C1           1.000     4.000     1.500    3.500
```

In 5 minutes, your group should explain what percent of data was between 1.5 and 3.5.

Example 4.2. Consider the following highway miles per gallon for 19 selected mini-compact, sub-compact, and compact cars [U.S. Department of Energy and U.S. Environmental Protection Agency, 1991 Gas Mileage Guide and 1991 Fuel Economy Guide Addendum]: 26, 46, 36, 31, 28, 28, 27, 38, 42, 36, 37, 33, 23, 29, 37, 34, 29, 40, 28. Your group should use Minitab to construct a stem and leaf display for the data in 5 minutes.

Example 4.3. What sequential keys would you use to get a box plot? (Individual quiz)

Example 4.4. If correlation coefficient is around 0.3, how do you interpret the relationship between the variables?

Example 4.5. Suppose MINITAB has generated the equation $y = 0.34 - 0.002\,x$. What dose this equation tell? Interpret numbers involved in this equation.

Example 4.6. If in a testing of differences of means from different populations, the *p*-value is 0.001, how would you decide on accepting or rejecting the hypothesis?

5. Results of Implementing the Method

5.1 BP STAT, Testing the Method for Learning Only

In both years of BP STAT, there were more participants from the rural parts of South Carolina than urban. Most of the participants were eleventh graders with more than 95 percent having GPA's of above 3.0 on a scale of 4 points. There was no significant difference between the number of males and females in the program. Participants had at least one year of algebra and more than half two years or more. They were somewhat familiar with Internet, Wordprocessing and computer games at the time they attended the program.

Only RT/HO method was used in the program. In brief, the analysis shows that both females and males learned equally well. The mean posttest grades for the males and female were 60.8% and 62.9%, respectively, with the coefficient of variations of 22% and 18%, respectively. Those from rural area did much better on the posttest than from urban area. The average posttest grade for students from rural schools was 63.2% with a coefficient of variation of 17% while the average posttest grade for students from urban schools was 58.4% with a coefficient of variation of 27%.

The number of mathematics courses taken and the student's academic performance in the program were strongly correlated with the correlation coefficient of 0.834. The average posttest scores of those with one year of high school mathematics, 59.4%, increased to 83.5% for students with 4 years of mathematics.

Judging from the average final test scores of all students for both pretest and posttest and both years 1997 and 1998, the participants scored substantially higher on average on the posttest than on the pretest. In 1997, improvement was from 18.07% on pretest to 64.45% on posttest, while in 1998 it was from 34.20% to 64.44%. To formally test the difference in pre- and post- test scores, a *t-test* on the paired differences, post - pre, by students was performed for each year 1997 and 1998, using MINITAB. The *t*- and *p*- values for 1997 were 20.56 and 0.00, respectively, and for 1998, they were 17.59 and 0.00, respectively. The *p*-value of 0.00 for each year shows a significant statistical evidence that the difference in means is greater than zero. Since a positive difference in this setting means a higher score on the posttest than the pretest, we take this to mean that there is a significant improvement in test scores on the posttest over the pretest (i.e. the students learned a great deal) in each year.

5.2 PVAMU Statistics Classes for Comparison of the RT/HO Method with the Traditional

At PVAMU two sections of Math 3023, Probability and Statistics, were offered during the fall of 2003 and spring of 2004. The same instructor taught both sections and covered the same topics with the same textbook and used the same tests. Average enrollment for each section was 32.

In the Traditional method section, the class involved: definition of terms, examples and discussions. Homework problems were then assigned covering several sections of a chapter. On the other hand, the RT/HO method section was data driven. Students collected data on some attributes such as gender, height, age, systolic blood pressure, their academic classification and so forth. Students then entered data into a Minitab Worksheet for analysis. The terms accompanying the output were then defined.

Students using the RT/HO method answered questions by describing data either by providing summary values or by using histogram, stem-leaf, digidot plots, pie charts, bar charts, Pareto charts, box-and-whisker displays and tallying data for categorical variables. They also answered questions to some problems by simulating data for some common distributions such as Bernoulli, binomial, Poisson, discrete uniform, continuous uniform, normal, gamma, Weibull, Lognormal, and exponential.

Some students using the Traditional method had a hard time understanding the *p*-value approach to hypothesis testing concerning a population parameter, assessing the utility of the model in regression analysis. Students using the RT/HO method continually learned from their peers through group discussions of definitions, concepts and homework assignments.

Students acknowledged learning a great deal from the presentations of research papers which covered topics such as multiple regression, two-way analysis of variance, nonlinear regression (intrinsically linear model), and analysis of contingency tables. Here is the list of projects of the five groups in the fall semester:

Group 1's project was on a two-factor factorial design. Detail of this project is in the Appendix.

Group 2 used data from their experiment. Their problem was to find a model that can be used to predict the percentage of properly sealed/wrapped packages of a product from plate temperature and plate clearance. They also assessed the utility of the model they decided to use by performing residual analysis, and checking the validity of assumptions of the model.

Group 3's objective was to determine if there is a relationship between the distance a golf ball travels and the brand of ball and whether this distance also depended on the type of club used to hit the ball. They used a two-factor factorial design to answer their questions.

Group 4's objective was to determine if drying time can be predicted from the amount of a certain chemical used in a process. Initially, they thought this problem would be solved by applying a simple linear regression model. On plotting the data and analyzing residuals, they discovered that the data followed an exponential distribution, $y = \beta_0 \cdot e^{-\beta_1 \cdot x}$. They transformed the data and regressed the logarithm of the outcome variable (drying time) with the predictor (amount of chemical used). A plot of residuals and ANOVA table showed this model provided a good fit.

Group 5's research project involved testing for independence between two categorical variables: passing car emissions test and type of inspection station.

5.2.1 Analysis of Data

The objective is to determine if there is a relationship between success in the course and method of teaching (Traditional versus RT/HO). Logistic regression is a method used to predict a categorical outcome from a set of independent variables that are not all continuous. Traditionally, discriminant function analysis has been utilized for this purpose, but is not appropriate because the assumption of multivariate normality is not satisfied when some of the predictors are categorical. Discriminant function analysis may also yield predicted probabilities that are less than zero or greater than one. Basic concept of this model can be found in Rosner (1998) or Gujarati (1995), for example.

If p denotes the probability of academic success, then $1 - p$ is the probability of not academic success. Thus, $p/(1 - p)$ is the odds ratio in favor of the academic success. It is known that logit, the (natural) log of the odds ratio, is linear in the variable and in the parameter, i.e., $\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X$, where X is a dichotomous independent variable that takes values 1 and 0, depending uponenrollment in the RT/HO or traditional sectin, respectively. Note that the difference of log-odds for 1 and 0 is β_1. That is, $\ln(odds_1) - \ln(odds_0) = \ln\left(\frac{odds_1}{odds_0}\right) = \beta_1$. See Neter, et al. (1996, p. 577), for example, for more detail. Thus, the interpretation of β_1, the slope, in our case is that it measures the change in the log-odds for a unit change of X. We expected the logit of success for a student to be greater than 1. Of, course, 1 would mean there is no difference in the methods. Thus, the estimated probability of success is $\hat{p} = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}}$.

Our variables (in this case) predictors, are sections of the course. Sections 1 (Traditional) and 2 (RT/HO) are coded 0 and 1, respectively. We have chosen our outcome variable as "academic success" and define it as the final score in the course. Based on our belief that the RT/HO method should increase the number of students passing the course, we consider the academic success as the average course score between 75 and 90, inclusive. Thus, we denote by 0 (not academic success) and 1 (academic success) scores less than and greater than or equal to 75, respectively.

It is worth noting that if we lower the cutoff line for the definition of academic success to a point that almost everyone succeed, score of 60, for instance, or so high that almost no one succeeds, score of 99, for instance, then log odds model will not work because $p/(1 - p)$ will be either infinity or zero and the logit becomes meaningless. In such circumstances, maximum likelihood models are used. However, due to its complexity, we avoid those extreme cases.

Using Minitab (a statistical computer software), applying binary logistic regression (success versus section), the Link Function and Logistic Regression Table are as follows:

<u>Link Function: Logit</u>
Response Information

Variable	Value	Count	
Acad. Success	1	116	(Event)
	0	17	
	Total	133	

<u>Logistic Regression Table</u>

Predictor	Coefficient	SE Coef	Z	P	Odds Ratio	95% CI
Constant	1.2730	0.3024	4.21	0.000		
Section	1.8181	0.6633	2.74	0.006	6.16	1.68 – 22.60

Thus, the estimated $\ln\left(\dfrac{p}{1-p}\right) = 1.2730 + 1.8181\,X$, where $X = 0$ for Traditional method and 1 for RT/HO method. Hence, $\ln(odds_0) = \ln\left(\dfrac{p}{1-p}\right) = 1.2730 + 1.8181(0) = 1.2730$ and $\ln(odds_1) = \ln\left(\dfrac{p}{1-p}\right) = 1.2730 + 1.8181(1) = 3.0911$. Therefore, β_1=1.8181 and the estimated ratio of the odds is $\widehat{OR} = \dfrac{odds_1}{odds_0} = e^{\beta_1} = e^{1.8181} = 6.16$, as is given in the Logistic Regression Table above. This relation expresses the odds in favor of academic success divided by odds in favor of not academic success. That means, students taught by the RT/HO method are six times more likely to succeed in this course than those taught by the Traditional method.

A 95% confidence interval for the true odds ratio from the Logistic Regression Table above is (1.68, 22.60) and since it does not include 1, students taught using the RT/HO method are more likely to succeed, in fact six times more, in this course than those taught by the Traditional method.

It should be noted that analysis showed that class attendance had no significant difference in the teaching method applied.

Remark: If we change the cutoff point to 80, then the Response Information and Logistic Regression Table will be as follows:

Link Function: Logit

Response Information:

Variable	Value	Count	
Acad. Success	1	92	(Event)
	0	41	
	Total	133	

Logistic Regression Table

Predictor	Coefficient	SE Coef.	Z	P	Odds Ratio	95% CI
Constant	-0.5220	0.5868	-0.89	0.374		
Section	0.9014	0.3872	2.33	0.02	2.46	(1.15, 5.26)

The p-value of 0.02 indicates that if the experience is repeated many times under same conditions, only 2 percent of the times we will observe a statistic as large as 2.33 or larger if the slope, β_1, is zero (methods of instruction are equivalent). Since this is a rare event, we conclude that the slope is sufficiently greater than zero. Thus, ‘section’ or method of instruction is a significant predictor of the log-odds of success.

The odds ratio for of 2.46 indicates that students taught by the RT/HO Method are about two and half times more likely to succeed in this course than those taught by the Traditional Method. A 95% confidence interval (1.15, 5.26) does not include 1 confirm the result.

Notice the decrease of the ratio comparing with the cutoff point of 75. That is, even when we raise the academic success bar, we still have a ratio of 2.5 to 1 in favor of success.

6. Discussion/Conclusion

Having observed the students' enthusiasm, curiosity, and astuteness toward working and completing their projects, we came to the conclusion that they have developed a real value for statistical analysis, by observing the emergence of non-trivial information. After original implementation of the method in 1996 and 1997 with high school students, the RT/HO method was used several times in subsequent years at college level introductory statistics course. At both high school and college levels, students demonstrated high level of interest and excitement in conducting research projects. In fall of 2003 the method was used by a faculty member different at Prairie View A&M University, Prairie View, Texas, different from the one in South Carolina in 1996 and 1997. At both the high school and college levels, we observed students' attitude and improvement towards developing an appreciation for statistical analysis. Moreover, we observed that students in RT/HO class demonstrated better understanding of statistical concepts and the write-up and the presentation of their projects were particularly helpful in their understanding of inferential analysis. We conclude that when students develop a project from selection of the topic, participate in the process of data collection, and learn statistical concepts parallel to their needs, they will develop a genuine interest in the subject and are capable of performing meaningful research on their own.

The students' maturity at college level did not make it any easier when it came to topic selection, developing a questionnaire, and distinguishing between descriptive statistics (survey) and inferential analysis. Results of the statistical analysis on the success of students are phenomenon. These observations suggest that statistics education should and can be a part of all students' curriculum at least as early as the beginning of high school.

We would like to expand the scope of the RT/HO method to more advanced courses in statistics as well as some mathematics courses. Attempts are being made toward those directions and we are very hopeful to see the results soon.

References

Allitt, Patrick, *Professor, Stop Your Microchips*, Chronicle of Higher Education, June 24 (2005).

Cangelosi, S. J. *Teaching Mathematics in Secondary and Middle School, Upper Saddle River,* New Jersey: Prentice Hall, (2003).

Cobb, G. *Reconsidering Statistics Education: A National Science Foundation Conference,* Journal of Statistics Education [Online], 1(1) (1993).

(www.amstat.org/publications/jse/v1n1/cobb.html)

Davidson, *N. The Small-Group Discovery Method in Secondary- and College-level Mathematics, Cooperative Learning in Mathematics: A Handbook for Teachers,* Menlo Park, California: Addison Wesley Publishing Company (1990).

Dietz, E. J. "A Cooperative Learning Activity on Methods of Selecting a Sample," *The American Statistician*, 47, 104-108 (1993).

Franklin, J. “Assessing Assessment: Are Alternative Methods Making the Grade?” *Curriculum Update*, Association for Supervision and Curriculum Development, Spring (2002).

Garfield, J. " Teaching Statistics Using Small-Group Cooperative Learning," *Journal of Statistics Education* [Online], 1(1) (1993).

(www.amstat.org/publications/jse/v1n1/garfield.html)

Goodsell, A. Maher, M., and Tinto, V., *Collaborative Learning: A Sourcebook for Higher Education*, University Park, PA: National Center on Postsecondary Teaching, Learning and Assessment (1992).

Gujarati, Damodar N. *Basic Econometrics, third edition*, McGraw-Hill, Inc (1995).

Hahn, G., and Hoerl, R. “*Key Challenges for Statisticians in Business and Industry,” (with discussion*), Technometrics, 40,19 213 (1998).

Hogg, R.V. “*Statistical Education: Improvements Are Badly Needed*,” The American Statistician, 45(4), 342-343 (1991).

Hogg, R. V. “*Towards Lean and Lively Courses in Statistics*,” in *Statistics for the Twenty-First Century*, MAA Notes Number 26, eds. F. Gordon and S. Gordon, Washington D.C.: The Mathematics Association of America, 3-13 (1992).

Keeler, C. M., and Steinhorst, R. K. *“Using Small Groups to Promote Active Learning in the Introductory Statistics Course: A Report from the Field,”* Journal of Statistics Education [Online], 3(2) (1995). (www.amstat.org/publications/jse/v3n2/Keeler.html)

Keeler, C., and Steinhorst, K. “*A New Approach to Learning Probability in the First Statistics Course*,” Journal of Statistics Education [OnLine], 9 (3) (2001).

(www.amstat.org/publication/jse/v9n3/keeler.html)

Moore, S. David. “*Teaching Statistics as a Respectable Subject*,” in Statistics for the Twenty-First Century, MAA Notes Number 26, eds. F. Gorden and S. Gorden, Washington D.C.: The Mathematics Association of America, 14-25 (1992).

Moore, D. S., Cobb, G., Garfield, J., and Meeker, W. Q. “*Statistics Education F in de Siecle*,” The American Statistician, 49(3), 250-260 (1995).

National Council of Teachers of Mathematics, Principles and Standards for School Mathematics, Reston, VA: Author (2000).

Neter, John, Michael H. Kunter, Christopher J. Nachtsheim, and William Wasserman. *Applied Linear Statistical Models*, fourth edition, IRWIN publishing Co. (1996).

Reynolds, B. E., Hagelgans, N. L., Schwingendorf, K. E., Vidakovic, D., Dubinsky, D.,Shahin M., and Wimbish, Jr., G. J., editors. *A practical Guide to Cooperative Learning in*

Collegiate Mathematics, MAA Notes Number 37, Washington D.C.: The American Mathematical Society (1995).

Scheaffer, Gnanadesikan Watkins, & Witmer. *Activity-Based Statistics*, Springer-Velag Publishing Co. (1996).

Snee, R. "What's Missing in Statistical Education?'", *The American Statistician*, 47(2), 149-154 (1993).

Steinhorst, R. K., and Keeler, C. W. "*Developing Materials for Introductory Statistics Courses from Conceptual, Active Learning Viewpoint,*" Journal of Statistics Education [Online], 3(3) (1995). (www.amstat.org/publications/jse/v3n3/Steinhorst.html)

Watts, D. G. "*Why Is Introductory Statistics Difficult to Learn? And What Can We Do to Make It Easier?,*" The American Statistician, 45(4), 290-291 (1991).

Appendix

As described, research is an integral part of the RT/HO method. Here is a detailed example of students' research.

Group 1: Project on a Two-Factor Factorial Design

Members of the group: Eugene Bacon, Ebony Hawkins, Noah Rattler, Chris Berry, Brian Shiloh, and Anthony Johnson

ABSTRACT: A two-factorial design involves analyzing the interaction effect, and the main effects, depending on the outcome of the analysis of the interaction effect.

PROBLEM STATEMENT: The chemical element antimony is sometimes added to the tin-lead solder to replace the more expensive tin and reduce the cost of soldering. An experiment was conducted to determine the effect of antimony on the tin-lead solder joint (Journal of materials Science, May 1986). Tin-lead solder specimens were prepared using one of four possible methods: water-quenched, WQ; oil-quenched, OQ; air-blown, AB; and furnace-cooled, FC and with four possible amounts of antimony: 0%, 3%, 5%, and 10% added to the composition.

DATA: COOLING METHOD. NOTE: Data represent shear strength (MPa)

% ANTIMONY	WQ	OQ	AB	FC
0	17.7, 19.5, 18.3	20.0, 24.3, 21.9	18.3, 19.8, 22.9	19.4, 19.8, 22.9
3	18.6, 19.5, 19.0	20.0, 20.9, 20.4	21.7, 22.9,22.1	19.0, 20.9, 19.9
5	22.3, 19.5, 20.5	20.9, 22.9, 20.6	22.9, 19.7, 21.6	19.6, 16.4, 20.5
10	15.2, 17.1, 16.6	16.4, 19.0, 18.1	15.8, 17.3, 17.1	16.4, 17.6, 17.6

Data source: Tomlinson, W.J., and Cooper, G.A. "Fracture Mechanism of Brass/Sn-Pb-Sb Solder Joints and the Effect of Production Variables on the Joint Strength." Journal of Materials Science, Vol. 21, No. 5, May 1986, p. 1731

Assumptions:

- The distribution of each response factor-level combination, or treatment is normal.
- The treatments have same variance
- All samples are independently and randomly drawn from their respective populations

Definition of Terms:
N- Number of observations
Mean- Average value
Median- Middle ordered observation
Trmean- Mean after discarding the lowest and highest 5% of the values
StDev- Average deviation of the observations from the mean
SE Mean- Standard error or standard deviation of the sample mean
Minimum- Smallest value
Maximum- Largest value
*Q*1- Value that is equal to or larger than 25% of all observations
*Q*3- Value that is equal to or larger than 75% of all observations
DF- Degrees of freedom are the number of parameters allowed to vary after a statistic is computed
SS- Corrected sum of squares, i.e., sum of squared deviations of the values from some mean
MS- Mean square is obtained by dividing the sum of squares by the corresponding degrees of freedom
F-Distribution – Variance ratio distribution
P-Value- Attained significance level which is the probability of observing F-statistic as large as the one calculated or larger given the null hypothesis is true

Minitab Commands:

```
MTB > SET C1 <RETURN>
DATA > 17.6 19.5 ...... 16.4 17.6 17.6<RETURN>
DATA > END <RETURN>
MTB > SET C2 <RETURN>
DATA >12(0) 12(3) 12(5)   12(10) <RETURN>
DATA > END   <RETURN>
MTB > SET C3 <RETURN>
DATA > (20 20 20 25 25 25 30 30 30 35 35 35)4 <RETURN>
MTB > NAME C1 ‘SHEAR’ C2 ‘ANTIMONY’ C3 ‘COOLMETH’ <RETURN>
```

DESCRIPTIVE STATISTICS FOR SHEAR STRENGTH BY ANTIMONY

ANTI MONY	N	MEAN	MEDIAN	TR MEAN	SE MEAN	MIN	MAX	Q1	Q3
0	12	20.175	19.800	20.020	0.567	17.600	24.300	18.575	21.500
3	12	20.408	20.200	20.340	0.386	18.600	22.900	19.125	21.500
5	12	20.617	20.550	20.810	0.517	16.400	22.900	19.625	22.125
10	12	17.017	17.100	17.000	0.296	15.200	19.000	16.400	17.600

DESCRIPTIVE STATISTICS FOR SHEAR STRENGTH BY COOLING METHOD

Cooling method	N	MEAN	MEDIA N	TR MEAN	SE MEAN	MIN	MAX	Q1	Q3
WQ	12	18.640	18.800	18.620	0.542	15.200	22.300	17.225	19.500
OQ	12	20.450	20.500	20.470	0.602	16.400	24.300	19.250	21.650

AB	12	20.175	20.700	20.340	0.734	15.800	22.900	17.550	22.700
FC	12	18.950	19.500	19.010	0.452	16.400	20.900	17.600	20.200

Two-way ANOVA: Shear Strength versus Amount of Antimony, %, and Cooling Method

Source	DF	SS	MS	F	P
Antimony	3	104.19	34.73	20.12	0.000
Cool-method	3	28.63	9.54	5.53	0.004
Interaction	9	25.13	2.79	1.62	0.152
Error	32	55.25	1.73		
Total	47	213.20			

ANALYSIS AND DISCUSSION

TEST OF SIGNIFICANCE FOR INTERACTION

Null hypothesis: There is no interaction between amount of antimony and cooling method Alternative hypothesis: The two variables interact to affect shear strength Significance level = 0.05 Critical region: Reject the null hypothesis if p-value is less than 0.05

Test statistic: P-value associated with calculated F = 0.152

Decision: Fail to reject the null hypothesis and conclude there is no significant interaction Between amount of antimony in the mixture and the cooling method used.

TEST OF MAIN EFFECTS

A: ANTIMONY

Null hypothesis: The four population means are equal Alternative hypothesis: At least two of the means differ Significance level = 0.05

Critical region: Reject the null hypothesis if *p*-value is less than 0.05

Test statistic: The *p*-value associated with calculated *F* is 0.000.

Decision: Reject the null hypothesis at the 0.05 level of significance and conclude that the At least two of the means are different. We will now perform Fisher’s multiple comparison Test to determine the actual means that are different.

FISHER’S LEAST SIGNIFICANCE DIFFERENCE PROCEDURE

The 0%, 3%, and 5% antimony mixtures are significantly higher than the 10% antimony mixture at the 0.05 family error rate (0.01 individual error rate).

B: COOLING METHOD

Null hypothesis: The four population means are equal Alternative hypothesis: At least two of the four means are different

Significance level = 0.05

Critical region: The null hypothesis will be rejected if the calculated value of *F* exceeds the

Maximum value expected when the null hypothesis is true or *p*-value is less than the significance level.

Test statistic: Since the calculated value of *F* (5.53) exceeds the maximum value expected when the null hypothesis is true (4.65), the null hypothesis is rejected at the 0.05 level of significance (*p*-value = 0.004).