The Wayback Machine - https://web.archive.org/web/20130307064227/http://www.amstat.org/publications/jse…

Teaching Bits: A Resource for Teachers of Statistics

Journal of Statistics Education v.7, n.3 (1999)

Robert C. delMas
General College
University of Minnesota
354 Appleby Hall
Minneapolis, MN 55455
612-625-2076

delma001@maroon.tc.umn.edu

William P. Peterson
Department of Mathematics and Computer Science
Middlebury College
Middlebury, VT 05753-6145
802-443-5417

wpeterson@middlebury.edu

This column features "bits" of information sampled from a variety of sources that may be of interest to teachers of statistics. Bob abstracts information from the literature on teaching and learning statistics, while Bill summarizes articles from the news and other media that may be used with students to provoke discussions or serve as a basis for classroom activities or student projects. Bill's contributions are derived from Chance News (http://www.dartmouth.edu/~chance/chance_news/news.html). Like Chance News, Bill's contributions are freely redistributable under the terms of the GNU General Public License (http://gnu.via.ecp.fr/copyleft/gpl.html), as published by the Free Software Foundation. We realize that due to limitations in the literature we have access to and time to review, we may overlook some potential articles for this column, and therefore encourage you to send us your reviews and suggestions for abstracts.

From the Literature on Teaching and Learning Statistics

"Teaching the Reasoning of Statistical Inference: A 'Top Ten' List"

by Allan J. Rossman and Beth L. Chance (1999). *The College Mathematics Journal*, 30(4), 297-305.

The authors state that instructional principles that have grown from the reform movement in statistics education have been more readily applied to data analysis than to the teaching of statistical inference. To remedy this situation, the authors offer ten recommendations for teaching the reasoning of statistical inference. Each recommendation is followed by examples of activities and exercises that illustrate the suggested approach.

"Using Statistics and Statistical Thinking to Improve Organisational Performance"

by S. B. Dransfield, N. I. Fisher, and N. J. Vogel (1999). *International Statistical Review*, 67(2), 99-122.

This article may be of interest to undergraduate statistics majors, students in graduate statistics programs, and their instructors.

Excerpt from the Summary: This paper describes a strategy that seeks to link measurement to all facets of organisational performance, particularly to desired business outcomes, and also to mesh measurement with process improvement in a natural way. The use of statistics and statistical thinking is then discussed in this context, with particular focus on the opportunity for statisticians to have a key role at the top decision-making level of the organisation. We argue that the role requires skills in both advanced technical statistical modelling and analysis, and in statistical thinking. It also requires a preparedness to form an appreciation of the business and management imperatives faced by the leaders of an enterprise, and a willingness to work from this basis.

The Mathematics Teacher
November 1999, Volume 92, Number 8

This issue of *The Mathematics Teacher* presents a "Focus on Statistics" with many articles that should be of interest to the statistics educator.

"*Titanic:* A Statistical Exploration"

by Sandra L. Takis, 660-664.

Abstract: The tremendously popular movie *Titanic* has rejuvenated interest in the Titanic and its passengers. Students are particularly captivated by the story and by the people involved. Consequently, when I was preparing to explore categorical data and the chi-square distribution with my class, I decided to use the available data about the Titanic's passengers to interest students in these topics. This article describes the activities that I incorporated into my statistics class and gives additional resources for collecting information about the Titanic.

"The Role of Technology in Introductory Statistics Classes"

by George N. Bratton, 666-669.

The author states that incorporating technology (computers or calculators) into introductory statistics courses does three things: it eliminates the need to teach certain topics, provides a better way to teach other topics, and permits the teaching of topics that were not possible in the past. Suggestions and examples are provided for each of the three areas.

"The Double Stuf Dilemma"

by Marie A. Revak and Jihan G. Williams, 674-675.

The authors describe a sweet activity for testing whether Double Stuf Oreo cookies actually contain twice the amount of filling as regular Oreo cookies. In addition to an activity that will appeal to the students' taste buds, the authors describe several methodological issues that can be raised with students that can make this activity a valuable learning experience.

"What Is Normal, Anyway?"

by Maria E. Calzada and Stephen M. Scariano, 682-690.

The authors state the importance of testing whether data can be considered to follow a normal distribution before using the data for inferential statistics. While visual displays such as histograms and normal-quantile plots are useful tools, the authors note that identifying normality based on these plots is not an exact science. The authors suggest the Lilliefors test for normality as a more objective test and illustrate with several examples how the Lilliefors test can be conducted with a TI-83 graphing calculator. The TI-83 program code for the Lilliefors test is provided in an appendix and is available at the NCTM Web site. Modified versions of the code for the TI-85 and TI-86 can be obtained from the authors.

"Discovering an Optimal Property of the Median"

by Neil C. Schwertman, 692-703.

The author uses a practical problem that is faced by many utilities to help students learn about a property of the median: What is the optimal location of a central depot that provides the highest efficiency for distribution of a good or product? The author describes a set of activities that help students discover that the optimal point that minimizes the absolute deviations from the point is the median. Six activities are provided, each one building on knowledge learned in the previous activity.

"Secondary Students' Performance on Data and Chance in the 1996 NAEP"

by J. Michael Shaughnessy and Judith S. Zawojewski, 713-718.

The authors report on findings from the most recent National Assessment of Educational Progress (NAEP). While the NAEP results indicate that teachers are placing more emphasis on data analysis, statistics, and probability in courses, there is still room for improvement given that three-fourths of twelfth-graders have not studied statistics or probability in school. The authors provide two examples from the 1996 NAEP to illustrate difficulties that students have when reasoning about data and chance. These examples can develop students' understanding of chance and data through discussion and debate in the classroom.

"German Tanks: A Problem in Estimation"

by David C. Flaspohler and Ann L. Dinkheller, 724-729.

Excerpt from the Introduction: Estimation is covered extensively in elementary statistics courses. The example discussed in this article describes a real-world situation and a simulation of that problem in which the selection of a suitable estimate is less apparent. This problem can be used at various levels. In a more elementary setting, the problem is useful to describe the concept of estimation. An advanced class can use the problem to discuss unbiasedness, minimum variance, and best estimates. At any level, the problem furnishes an excellent opportunity to make connections to the social studies curriculum and demonstrates an application of statistical techniques.

"Using Simulation on the Internet to Teach Statistics"

by Vee Ming Ng and Khoon Yoong Wong, 729-733.

Introduction: This article shares some reflections about using simulations on the Internet to teach statistics. Because Internet resources keep appearing at a phenomenal rate, giving a comprehensive listing of all such resources within an article is impossible. The authors focus on a few ideas that are helpful for teaching and learning.

"Cooperative Teaching Opportunities for Introductory Statistics Teachers"

by Deborah J. Rumsey, 734-737.

The author describes how recommendations from the statistics education reform movement were used to restructure an introductory statistics teaching program. Emphasis was placed on using a cooperative teaching approach to promote both consistency in the way courses are taught and to support teachers as they incorporate new teaching methods into their courses. Two major resources were developed as the centerpiece of cooperative teaching: weekly teaching meetings and a teaching-resource notebook. The weekly meetings serve as a place to discuss statistical concepts, present teaching and assessment methods, and propose new ideas for discussion and feedback. The author discusses several different formats for the weekly meetings, including collaborative presentations. The teaching-resource notebook serves as a repository of examples, experiments and projects, articles and questions for in-class discussion, and handouts. Each item includes an explanation of the intended use in the classroom and suggestions for assessment. The article ends with additional ways to promote teacher involvement and ownership through collaboration.

"Data Analysis and Baseball"

by Gary Talsma, 738-742.

The author discusses how sabermetrics, a set of statistical approaches used to gain insights into baseball through the analysis of numerical records, can provide a familiar and motivating context for illustrating principles of data analysis. Various sources of baseball data are listed, including periodicals, books, and Web sites.

"Investigating Distributions of Sample Means on the Graphing Calculator"

by Gloria B. Barrett, 744-747.

The author provides examples and methods for using the TI-83 graphing calculator to help students explore the sampling distribution of sample means drawn from normally distributed populations and populations that are not normally distributed.

Stats: The Magazine for Students of Statistics
Fall 1999, Number 26

"The Census and Politics"

by Barbara Bailar, 8-10.

This article presents background information on the controversy surrounding the traditional undercount in the census. The author indicates many of the important uses for census data that make the undercount an issue. Several issues surrounding the undercount are discussed, such as the fact that there are overcounts as well as undercounts, that the undercount does not affect all people equally, that the adjustments made to the undercount are based on conjecture as well as fact, and how recent legislative acts can both facilitate and hinder the processing of census data.

"Ear Growth Revisited"

by Linda J. Young, RaQwin M. Young, Kristin E. Kelly, Suzanne R. Kirby, and Beverly A. Kelly, 11-13.

The article describes the research method and analysis used by two 10-year-old fifth graders to replicate and extend a study reported in the *British Medical Journal* on the growth rate of ear length in males and females aged 30 to 93. The young researchers extended the research design to answer two questions not addressed in the original study: Do children's ears grow at the same rate as adults, and do males' and females' ears grow at the same rate? The research method could make for an interesting class project in an introductory statistics course, and could promote discussion on the advantages and disadvantages of comparable research designs (e.g., a comparison of cross-sectional and longitudinal studies).

"AP STATS"

by Bob Stephenson, 15-18.

The purpose of the "AP STATS" column is to present fundamental ideas that have come from the teaching of Advanced Placement statistics courses. In each column, an exercise or activity that can be used in the classroom or as an assignment is discussed, along with the statistics or mathematics that underlie the topic. This issue's column addresses why the squared deviations from the mean are divided by $(n - 1)$ instead of n when computing the sample variance.

"From Maris to McGwire: A Statistical View of the Homerun Race"

by Chad Cleaver, 19-21.

The author shows that the Poisson distribution adequately models the likelihood that "power hitters" will hit a homerun during a single season. In the conclusion, the author discusses extensions of this study that might provide interesting classroom activities. Examples of extensions are testing whether the Poisson distribution is equally effective in modeling the homerun probabilities of average players and testing whether models based on the data for a season provide accurate predictions for the next season.

"On the Job: A Day in the Life of a Statistician at Southwest Research Institute"

by Janet P. Buckingham, 24-26.

The author describes how her fascination with numbers led her to a B.S. in Computer Science and Mathematics, and then a job as a Research Analyst at Southwest Research Institute in San Antonio, Texas. After 21 years, she is now a Statistical Consultant for the Institute. The position requires her to provide statistical guidance and analysis for any project that originates within the Institute. This provides opportunities to work on a variety of scientific topics and to use an assortment of statistical analysis techniques. In her "typical day," she describes meeting with engineers who are working on automotive emissions research, providing the experimental design and data entry format for a Department of Energy project, meeting with a colleague to discuss problems he's having with a one-factor analysis of variance, and participating in a meeting between Institute personnel and government engineers regarding the results of a reliability study of a new magneto-optic inspection instrument.

"Outlier...s"

by Allan J. Rossman, 27-29.

In this issue's column, Allan looks into the art of procrastination and discovers that you can't get away from statistics, no matter how hard you try. He shows us that statistics are relevant to playing solitaire on the computer, viewing videos, and reading both non-fiction and fiction. As usual, interesting assignments are posed for the reader to tackle while procrastinating.

Teaching Statistics

A regular component of the Teaching Bits Department is a list of articles from *Teaching Statistics*, an international journal based in England. Brief summaries of the articles are included. In addition to these articles, *Teaching Statistics* features several regular departments that may be of interest, including Computing Corner, Curriculum Matters, Data Bank, Historical Perspective, Practical Activities, Problem Page, Project Parade, Research Report, Book Reviews, and News and Notes.

The Circulation Manager of *Teaching Statistics* is Peter Holmes, Peter.Holmes@ntu.ac.uk, RSS Centre for Statistical Education, University of Nottingham, Nottingham NG7 2RD, England. *Teaching Statistics* has a website at http://science.ntu.ac.uk/rsscse/ts/.

Teaching Statistics, Autumn 1999
Volume 21, Number 3

"Squaring the Circle -- Statistically Speaking" by P. Glaister, 66-69.

The author describes two games that are similar in form to Buffon's classic needle problem. One game involves dropping a ring onto a surface marked with parallel lines spaced at equal distances, the distance being larger than the diameter of the ring. The task is to determine the probability that the ring, when dropped, will cross one of the lines. A similar problem is described where a rectangular figure is dropped with the requirement that the distance between the parallel lines is greater than a diagonal of the rectangle. The author offers solutions to these problems and comparisons of Buffon's needle problem with the ring and rectangle problems that can be instructive to students of probability.

"Statistical Nuggets with a Graphics Calculator" by Alan Graham, 70-73.

This is the first of two articles that illustrate the use of the graphics calculator to provide students with insights into major concepts in statistics. One example illustrates how the graphing calculator can help students understand that spurious patterns occur naturally when generating random values. A second example helps students learn how changing the class interval size affects the overall shape of a histogram. Copies of worksheets for conducting the exercises are provided.

"Young Students' Informal Statistical Knowledge" by Ian J. Putt, Graham A. Jones, Carol A. Thornton, Cynthia W. Langrall, Edward S. Mooney, and Bob Perry, 74-78.

The articles summarizes results from a study of students' statistical thinking. In particular, the authors take a broad perspective and look at the ways students organise, describe, represent, and analyse data. Eight students from two elementary schools in the midwestern United States were interviewed about their understanding of bar and tally graphs. Students' responses to the interview probes were classified as either idiosyncratic or higher-level. One implication of the research is that teachers can become more aware of the range of students' statistical thinking by using open-ended questions and probes like the ones illustrated in the article. Their evidence also suggests that students who respond to these probes demonstrate higher levels of analysis and interpretation.

"The Birthday Problem Revisited" by Jonny Griffiths, 78-80.

The author presents a variant of the classic birthday problem that presents a good example of the Poisson approximation to the binomial.

"Having a Ball with Confidence Intervals" by James F. Robison-Cox, 81-83.

The author describes a hands-on activity that can be used as an analogy to the construction and interpretation of confidence intervals. The activity takes 30 to 50 minutes and requires props such as tennis balls, a chalk board, and rulers. A chalkboard with a thin film of chalk covering its surface is required. A vertical line, which

represents the population mean and is labeled "mu," is drawn on the chalkboard. Students take turns tossing a tennis ball at the vertical line. A horizontal line is extended for a specified distance from each mark left by a tennis ball. This can lead to discussions of how many intervals cross the vertical line and how to improve the accuracy and precision of the procedure. Students' suggestions are tried until the accuracy is increased to a 90% level. The author discusses how to draw parallels between this activity and procedures for constructing confidence intervals.

"On Approximate Calculations" by Flavia Jolliffe, 84-87.

The author argues that while calculators and computers allow us to focus more attention on statistical concepts, there are areas in statistics where students benefit from working through small examples or learning to approximate calculations in order to recognize when answers are clearly wrong. Areas where approximate calculations are helpful and suggestions for ways to assess students' understanding and use of approximate calculations are presented.

"An Aid to Intuitively Understanding Bias and Consistency" by Neil H. Spencer, 88-89.

Summary: Students often find the concepts of bias and consistency hard to understand intuitively. This article presents means of demonstrating these concepts using a computer so that the students, once they have seen them in action, can more readily grasp their meanings and implications.

"Directional Data" with an Introduction by the Editor, Gerrald Goodall, 90-92.

The editor presents an article that was used to develop a comprehension question for an A-level mathematics examination in the UK. As stated in the summary, "This article displays some very counter-intuitive problems of 'averages' and shows one area of statistics in which vector methods play a natural role."

Topics for Discussion from Current Newspapers and Journals

"How to Predict Everything"

by Timothy Ferris, *The New Yorker*, 12 July 1999, pp. 35-39.

Princeton physicist J. Richard Gott III has an all-purpose method for estimating how long things will last. In particular, he has estimated that, with 95% confidence, humans are going to be around at least fifty-one hundred years, but less than 7.8 million years. Gott calls his procedure the Copernican method, a reference to Copernicus' observation that there is nothing special about the place of the earth in the universe. Not being special plays a key role in Gott's method. In this interview with *The New Yorker*, Gott explains how he developed the method, and gives some of its many other applications.

In 1969, just after graduating from Harvard, Gott was traveling in Europe. While touring Berlin, he wondered how long the Berlin Wall would remain there. He realized that there was nothing special about his being at the Wall at that time. Thus if the time from the construction of the Wall until its removal were divided into four equal parts, there was a 50% chance that he was in one of the middle two parts. If his visit was at the beginning of this middle 50%, then the Wall would be there three times as long as it had so far; if his visit was at the end of the middle 50%, then the Wall would last 1/3 as long as it had so far. Since the Wall was 8 years old when he visited, Gott estimated that there was a 50% chance that it would last between 2.67 and 24 years. As it turned out, it was 20 more years until the Wall came down in 1989. This success of this prediction spurred Gott to write up his method for publication. (It appeared in the journal *Nature* in 1993.)

The New Yorker has a special interest in knowing how long various theater productions will run. On May 27, 1993, Gott looked up all the shows listed in *The New Yorker*, including Broadway and off-Broadway plays and musicals. He called each theater to find when the shows had opened. He then used his method to compute 95% prediction intervals for the time they would close. Forty-four shows were running at the time, and so far thirty-six have closed, all at times within his confidence intervals. Since the others are all still within the range he predicted, Gott remarks that he is "batting a thousand" so far!

Gott describes a number of other amusing applications of the Copernican method. For example, the reason that you so often find yourself in the longest line at the supermarket is that there is nothing special about you. You are just a random person waiting to check out. When you choose a random person from those waiting to check out you are more likely to get one from a long line. You can find much of this discussion on-line in an article Gott wrote for *The New Scientist* (http://www.newscientist.com/ns/971206/grimreckoning.html).

"Studies of Leisure Time Point Both Up and Down"

by Janny Scott, *The New York Times*, 10 July 1999, B7.

Do Americans have more leisure time than they used to, or less? This article explores why researchers have come up with contradictory answers. One problem is that researchers have different methods for trying to answer this question. They consider trends over different time periods and collect data differently -- some relying on subjects' memories, others on keeping notebooks or diaries. In addition, their definitions of leisure time and work may differ.

The Bureau of Labor Statistics is the largest source of public data on the number of hours per week Americans work. The Bureau regularly surveys 400,000 non-farm businesses for the number of hours worked by production and non-supervisory workers. The Bureau also regularly conducts telephone interviews of 50,000 randomly selected households, asking for the hours worked in the previous week. Interpreting all of these data is another matter, and as Bureau economist Randy Ilg points out, "... one set of statistics will give you one point of view and another set ... will tell you something else.

Social science researchers have generated their own data on the subject. In her 1991 book, *The Overworked Americans*, Harvard economist Juliet Schor reported that time on the job had increased steadily over the past 20 years. But a 1997 book entitled *Time for Life*, by sociologists John Robinson and Geoffrey Godbey, found an opposite trend. The authors reported that men and women were working less than they did in 1965 and had gained an hour a day in free time. Robinson and Godfrey asked their subjects to keep notebooks. Researchers who favor this approach say that relying on memories causes a bias because people tend to exaggerate the amount that they work. Those who rely on the memories of the subjects counter that using notebooks biases the results in the other direction, since people who are busy forget to record some of the work periods.

Some researchers now argue that the average change in the amount of leisure may be the wrong measure to use, because there appear to be two different groups of Americans moving in different directions. One group includes baby boomers, who tend to be workaholics with little leisure time. The other group, which reports increasing leisure time, includes older Americans taking early retirement, and younger couples who start their families later in life and have fewer children. Simply reporting one overall average misses this story.

"The Courts vs. Scientific Certainty"

by William Glaberson, *The New York Times*, 27 June 1999, Sect. 4, p. 5.

The Institute of the National Academy of Sciences reported a study which looked at all the studies that have been carried out to test the effect of silicone breast implants on women's health. They concluded that silicone implants

can cause more local problems than have been previously suggested. However, they found no convincing evidence that the implants cause more serious ailments such as lupus and rheumatoid arthritis, as has been claimed in lawsuits. More than $7 billion in jury awards and settlements has already been paid out in these suits.

This article considers the difference between what scientists require for evidence and what the court does. Scientists tend to require evidence at the 95% level. There are no deadlines for conclusions to be reached, and research continues until convincing evidence is produced. The court, however, does not have the luxury of time; it must act at a specific time on a specific case. In civil cases, juries are told that a "preponderance of evidence" provides enough certainty to favor the plaintiff. The article interprets this as meaning there is at least a 51% chance that the charges are valid. But the standard does not appear to be consistently applied. In the case of silicone implants, the court seems to have believed preliminary evidence backing up the plaintiff's claims, but it now appears this judgment was faulty. On the other hand, until recently it has been difficult for smokers who developed serious illnesses to successfully sue a tobacco company, despite overwhelming long-term evidence that smoking is responsible.

According to this article, some experts feel that we just have to get used to the fact that the legal system is not perfect and cannot act like science. Others say that the tort system is supposed to be based on the assumption that you only pay when you have done harm. In rulings beginning in 1993, the United States Supreme Court has said that judges should act as gatekeepers and be more vigilant in assuring that "scientific evidence is not only relevant but reliable." It remains to be seen what effect these rulings will have.

"To Spot Bias in SAT Questions, Test Maker Tests the Tests"

by Amy Dockser Marcus, *Wall Street Journal*, 4 August 1999, B1.

It has been well-documented that on average black and Hispanic students score below whites on SAT exams, and women on average score below men. Critics of the tests have used these data to charge that the SAT is biased against certain groups. In June, the U.S. Department of Education Office of Civil Rights began circulating draft legal guidelines outlining what it considers bias. One concern is that colleges using the SAT in admissions may face legal action because of differences in the scores of different groups.

The Educational Testing Service (ETS) reports that they have made every effort to avoid bias and make exams fair to all who take them. According to the article,

> The company tests some questions each year on a section of the SAT that doesn't count toward the total score and evaluates how students of various races and sexes fare in answering them. If one group answers a question significantly better than other groups do, the question is banished from future tests.

Here are two examples of questions cited in the article, one from the verbal section and one from the math. These were tried out on students who took the SAT exams in 1998. The questions were found to be biased and will not appear on future SAT exams.

Verbal

DUNE : SAND ::

    a. beach : ocean
    b. drift : snow
    c. wave : tide
    d. rainbow : color
    e. fault : earthquake

Correct Answer: b.

Results: 23 percent more whites than African-Americans and 26 percent more whites than Hispanics answered the question correctly.

Hypothesis: Regional variations. High proportions of African-Americans and Hispanics live in the south and southwest areas of the country, where there is less familiarity with terms associated with extreme winter weather.

Math

> At North Industries, 1,200 employees are in the health plan and half of all company employees are in the savings plan. Of all the company savings plan members, 800 are in the health plan and 250 are not in the health plan. How many employees of North Industries are not in the health plan?

Correct Answer: 900.

Results: 16 percent more whites than Asian-Americans answered this question correctly.

Hypothesis: Asian-Americans tend to do worse on word problems applied to real life situations than whites of the same ability.
(Source: Educational Testing Service)

As might be expected from this description, a number of letters to the editor expressed the opinion that this analysis of bias was itself unfair. In response, Paul A. Ramsey of ETS attempted to clarify the procedure, called Differential Item Functioning (DIF), as follows:

> [The DIF process] produces statistics that identify large differences in group performance on individual test questions; it helps us determine whether these differences relate to what the SAT is designed to measure (verbal and mathematical reasoning) or whether they could be caused by unrelated factors. The DIF procedure first matches test takers from different groups, e.g., males/females, on the basis of their overall test performance, and then flags any test question on which these matched groups perform in substantially different ways.

"Hope for Sale"

by Gina Kolata and Kurt Eichenwald, *The New York Times*, 3 October 1999, Sect. 1, p. 1.

"In Pediatrics, a Lesson in Making Use of Experimental Procedures"

by Gina Kolata and Kurt Eichenwald, *The New York Times*, 3 October 1999, Sect. 1, p. 40.

These two articles describe the difficulties surrounding the evaluation of experimental medical procedures. The first is a long discussion of a controversial breast cancer treatment: high-dose chemotherapy followed by bone marrow transplant or stem cell rescue. Researchers tried for years to conduct a clinical trial to determine if this treatment could save lives of patients with an advanced stage of breast cancer. But women who believed this was their only chance refused to enroll in clinical trials, where there was a chance they would not get the actual treatment. The procedure itself is difficult and carries substantial health risks. It involves harvesting bone marrow or stem cells from the patient prior to chemotherapy. The patient then receives massive doses of toxic drugs targeted at the cancer, but which also destroy bone marrow. After chemotherapy, the stem cells are replaced and

the bone marrow restored. If successful, the drugs will have killed the cancer cells, and the restored bone marrow will grow back before the patient dies of infection.

This experimental procedure was introduced in 1979 by Dr. Gabriel Hortobagyi, and by the late 1980's the results looked promising. Patients with advanced breast cancer who were given transplants had remission rates of 50 to 60% compared with the 10 to 15% remission rates achieved by conventional treatments. This led to pressure from Congress and elsewhere for insurance companies to pay for the expensive treatment, and businesses were established to provide it. But even as the procedure was becoming widely used, some cancer specialists began to have second thoughts. Hortobagyi realized that in the early days of this treatment, patients had been carefully selected: they were younger than 60 and free of conditions such as heart disease or emphysema that would make the procedure even more risky. The success rate, however, had been compared with the entire population of women with advanced breast cancer who were treated with conventional therapies. When Hortobagyi restricted the comparison group to women whose age and overall health matched the transplant group, he found that the survival rates were similar.

Last May, results of five studies -- two large clinical trials in the United States, two large trials in Europe, and a small trial involving 154 women in South Africa -- were reported at the annual meeting of the American Society of Clinical Oncology. In four of the five clinical trials, there was no difference in survival rates for women who were randomly assigned to have transplants and those who were assigned to have conventional therapy. Only in the South African study did those with transplants do better than those with conventional therapy. But a further look at the South African data showed that those who had transplants lived about as long as those in the other studies, while those who had the conventional treatment did far worse, raising some doubt about the validity of this study. The National Breast Cancer Coalition response was that the data speak for themselves: bone marrow transplants have been tested and failed. On the other hand, the American Society of Clinical Oncologists put out a press release saying that these papers presented "mixed early results" and more years of study are required. However, as remarked earlier, researchers are finding it difficult to find subjects for such experiments.

The authors observe that, while Federal rules require that new drugs or devices (such as a heart valve) should be proven safe and effective before being sold to the public, medical procedures (such as bone marrow transplants and surgical procedures) are not regulated by the government. The authors suggest that this reflects the government's reluctance to interfere with doctors' practice of medicine.

The second article explains how pediatric oncologists have solved this problem in their field. These oncologists have organized and have refused to offer experimental treatments outside of clinical trials. They sponsor only credible experiments. With their high quality-control they have been able to persuade insurance companies to pay for the treatments. While only 1% of adult cancer patients enroll in clinical trials, 60% of pediatric cancer patients do. The article documents some success stories that resulted from the pediatric oncologists being able to quickly carry out clinical trials leading to definitive results.

"Mozart Sonata's IQ Impact: Eine Kleine Oversold?"

by Rick Weiss, *Washington Post*, 30 August 1999, A9.

In a 1993 article in the journal *Nature* (Vol. 365, p. 611), psychologist Frances Rauscher and her colleagues announced that listening to a certain Mozart piano sonata produced a temporary improvement in "spatial IQ," as measured by a portion of a standard intelligence test. The specific task involved visualizing the shape that would result if a sheet of paper were cut and folded in various ways. Excitement about the "Mozart effect" spawned a marketing blitz that included a self-help book and CD-ROMs promising to enhance children's performance on standardized tests. Catalogs pitched devices to expectant mothers anxious to broadcast Mozart to the womb. Politicians also got into the act -- the governor of Georgia pledged that the state would buy a classical music recording for every new baby.

Now Christopher Chabris, a cognitive neuroscientist at Harvard, has completed a meta-analysis of all 16 previously published studies on the Mozart effect. He found little or no improvement in standardized test scores. Meanwhile, at Appalachian State College, psychologist Kenneth Steele has performed his own study, and similarly reports no evidence for the effect.

Rauscher and her supporters have criticized Chabris' and Steele's methodologies, and promise that forthcoming studies will support the original claim of an effect. Steele and Rauscher debated their positions on a recent NPR segment. You can listen to it on-line at
http://www.npr.org/news/healthsci/indexarchives/1999/Aug/990825.02.html

Correspondence from *Nature* and a *Psychological Science* paper by Steele and his colleagues are available from Steele's web site http://www.acs.appstate.edu/dept/psych/Faculty/Steele.htm

Even its supporters concede that the original study was over-interpreted. The *Washington Post* article concludes

> If nothing else, it seems, the rise and fall of the Mozart effect may teach the public a lesson about the tentativeness of all scientific discovery. If that happens, then the incomparable composer will have made people wiser after all, if not actually smarter.

"Cloudbusting"

The Economist, 21 August 1999, pp. 69-70.

Seeding clouds to boost rainfall was first attempted in the 1940s. Most people are familiar with the idea, and after all this time it is natural to assume that it has been proven effective. But any objective measure of success requires estimating how much rain would have fallen without seeding. This turns out not to be easy, and the article notes that despite 50 years of anecdotal evidence, we still lack solid proof that seeding actually works.

American researchers at the National Center for Atmospheric Research (NCAR) have spent three years on a rigorous investigation of a new technique called "hygroscopic flare seeding." A hygroscopic compound is one that attracts moisture. Salt is an example, and in clouds over the ocean, rain droplets are known to form around salt particles. It seems reasonable that seeding other clouds with hygroscopic particles would increase rainfall. Evidence for this was found in 1988, when a South African rain researcher named Graeme Mather noticed that emissions from a paper mill seemed to promote raindrop formation in the clouds above it. Chemical analysis showed that the emissions were effectively seeding the clouds with two hygroscopic salts: potassium chloride and sodium chloride. Mather followed up with a series of experiments using flares fired from small planes to seed clouds with hygroscopic compounds. The current NCAR efforts in Mexico now aim at reproducing and testing his results.

As noted earlier, researchers needed an experimental design that allowed comparison between similar seeded and unseeded clouds. The following "double blind" protocol was used. As a storm develops, a plane seeks out clouds for seeding. When one is found, the pilot notifies a controller on the ground, who opens a randomly chosen envelope to reveal an instruction card. The card is equally likely to say "seed" or "don't seed." The controller accordingly tells the pilot whether or not to ignite the hygroscopic flare. But the pilot then chooses a similar envelope of his own, whose instruction card tells him whether or not to obey the controller's instruction. In either case, he continues along the flight path that would be used for seeding, so the controller is unaware of his action. So far, 99 such flights have been completed, 48 of which resulted in seeding. The results support Mather's finding that seeding increases rainfall by 30-40%. Dr. Brant Foote, who oversees the program, points out that at least 150 flights will be needed to obtain convincing statistical evidence. Assuming the effect is real, the researchers will then need to address the practical question of whether the extra rain produces enough economic benefit to justify the cost of seeding.