ENHANCING LEARNING IN STATISTICS CLASSES THROUGH THE USE OF CONCRETE HISTORICAL EXAMPLES: THE SPACE SHUTTLE CHALLENGER, PEARL HARBOR, AND THE RMS TITANIC*

WALTER R. SCHUMM
Kansas State University

FARRELL J. WEBB
Kansas State University

CARLOS S. CASTELO
Kansas State University

CYNTHIA G. AKAGI
University of Kansas

ERICK J. JENSEN
Kansas State University

ROSE M. DITTO
Kansas State University

ELAINE SPENCER-CARVER
Kansas State University

BEVERLYN F. BROWN
Kansas State University

A HOST OF SCHOLARS have acknowledged the difficulties experienced by students learning research methods and statistics, especially those in the social sciences who may not be as mathematically inclined as students in other fields of study (Blalock 1987; Bridges, Pershing, Gillmore, and Bates 1998; Potter 1995; Rushing and Winfield 1999; Salkind 2000; Watts 1991; Wybraniec and Wilmoth 1999). Students' anxieties may cause them to delay taking methods courses, adversely affecting their ability to understand material in other classes that incorporate these tools (Bridges et al. 1998:15). Many suggestions for easing student anxiety have been offered, including the use of humor (Schacht and Stewart 1990), small group work (Borreson 1990; Longmore, Dunn, and Jarboe 1996), M&M's candies (Auster 2000), use of local county data (Renzulli 2000), a hands-on approach (Fischer 1996; Longmore et al. 1996), and active learning or learning by doing (Potter 1995; Rushing and Winfield 1999; Takata and Leiting 1987; Wright 2000). Generally, active or collaborative learning has been suggested as a key method for improving learning in the area of research methodology.

The senior author has been teaching research methods classes for over twenty years, with generally favorable student ratings. He incorporated the use of computers into classroom assignments and classroom work, as recommended by other scholars (Fischer 1996; Kain 1987; Karp 1995; Wilmoth and Wybraniec 1998). However, despite favorable student-teacher ratings and student comments of various kinds, as well as successful collaborations with students on classroom projects leading to joint publications, he experienced a growing concern

*We thank Helen Moore and several anonymous reviewers for their helpful comments on previous drafts. The senior author's father, Rear Admiral Brooke Schumm, served for four years in the Pacific during World War II as gunnery officer and executive officer of the U.S.S. San Diego, an anti-aircraft cruiser that participated in at least 15 sea battles (Schumm 2001; Whitmore 2000). A revised version of this paper was presented at the National Council on Family Relations annual conference, Rochester, NY, November 9, 2001 as a teaching round table. The paper is dedicated to author Carlos S. Castelo, who passed away on October 21, 2001 and was awarded a posthumous Ph.D. degree by Kansas State University on December 7, 2001. Dr. Castelo is survived by his wife and eleven children.

Please address all correspondence to Walter Schumm, School of Family Studies and Human Services, Justin Hall, Kansas State University, 1700 Anderson Avenue, Manhattan, KS 66506-1403; e-mail: Schumm@Humec.ksu.edu or WRSchumm@aol.com

Editor's note: The reviewers were, in alphabetical order, Fred Halley, Alisa Potter, and Judith Stull.

with apparent failures of student "learning." First, some students never seemed to get over their anxiety about statistics; secondly, students occasionally returned a year or two later to ask questions that indicated that they had not retained, at least with any confidence, the material they had learned in the classroom. Blalock (1987:166) noted the problem of students believing they understood a statistical issue only to find they cannot explain the issue later, for example on an exam or during a lab exercise. A further concern was that for students who were not social science majors or had no intention of ever using statistics to solve "sociological" or "social science" problems, the course was merely another required hurdle to clear on the way to graduation. As a result, the use of sociological examples in class did not retain their attention or maintain their interest. In addition, some students had a difficult time grasping abstract concepts (e.g., does quality of premarital counseling as measured by a certain scale predict subsequent marital satisfaction as measured by a different scale?). The specific classes involved in this discussion were beginning-level graduate courses in research methodology with enrollments of between six and 12 students; however, these issues would be relevant for advanced undergraduate courses as well. The students in these classes ranged in age from their early twenties to their mid-fifties, were predominately females, and represented a moderate degree of ethnic diversity.

Two approaches were developed to deal with these concerns of learning and retention. First, instead of reliance upon "lecture notes" per se, the plan of instruction instead incorporated "data analysis learning objectives." Handouts were developed that specified learning objectives for each class day and the kinds of problems the student should be able to solve if those objectives had been met. Each class ended with a quiz that could not be completed successfully unless the student had mastered the day's material at the recall, comprehension, analysis, and evaluation and synthesis levels (Steen, Bader, and Kubrin 1999:171-72) in accordance with Bloom's taxonomy (Bloom et al. 1956). The second instructional approach included finding and applying examples that would meet several criteria for sustaining student interest and enhancing their learning and confidence with statistics. These examples were chosen to fit a complex set of criteria as outlined in Figure 1. The examples must involve concrete issues such as historical events rather than what students might consider more esoteric issues; the examples should be something most students would recognize from current or past events. Another criterion was that important issues be involved, possibly life and death issues. Also, each example needed to appeal to students who might not be planning to use statistics in their careers. Examples had to accommodate a variety of simple-to-complex statistical tests so that they could be used throughout the course and not just at the end of the term. The number of cases involved in the examples had to be small enough to permit students to enter the data themselves in class, on personal or laptop computers. Finally, the example must lend itself to role-playing how a professional might deal with the particular issue (Figure 1).

One example matching the criteria was developed during a class in the spring of 2000, while the others were developed shortly thereafter and used in a class in the fall of 2001. Three examples used in the classes follow; each example meets the above criteria.

THE CHALLENGER EXAMPLE

On January 28th, 1986, the space shuttle Challenger blew up shortly after launch, killing all the astronauts aboard, including one of the first black and one of the first female astronauts in the U.S. space program (Vaughn 1996). Martz and Zimmer (1992:42) recalled that the Challenger disaster "interrupted for a time one of the most successful engineering and scientific achievements in human history" and "also

Figure 1. Suggestions for Selection of Historical Examples

Criteria for Selecting Historical Examples

1. Use intense examples that possibly involve life and death issues.
2. Use examples that appeal to a wide audience, including those who may not make a career of research methods or statistics.
3. Use examples that allow for a variety of statistical tests, from simple to complex, so that the example can be used throughout the course.
4. Use examples that are familiar to students, through personal experience or well-known history.
5. Use examples that involve a small enough number of cases so students can input data by themselves if necessary or desired for training.
6. Use examples that lend themselves to role-playing the use of statistics in professional situations.

evoked a national resolve to more clearly and accurately determine the risks of space flight." What is often forgotten is that before the flight, shuttle engineers worried about launching the Challenger at the predicted launch temperature of 31 degrees Fahrenheit. Three O-rings (of a total of six on the shuttle) had failed on the only flight to have been launched at the lowest temperature to date (53 degrees), and there had been indications that O-rings failed to expand to seal joints properly at a "cold" temperature of 50 degrees (Dalal, Fowlkes, and Hoadley 1989:946). Failure of O-rings to seal critical launch vehicle joints had the potential to increase the probability of catastrophic failure of a shuttle flight, as O-rings kept hot gases from spreading to adjacent parts of the fuel system where fuel had not yet ignited. The night before the attempted launch, flight engineers had considered the use of statistics to resolve the question of whether temperature was related to O-ring failure (Dalal et al. 1989:945); however, they did not include the data for flights that had no problems (Lavine 1991:919), thus limiting the sample size to the seven flights that had experienced O-ring failure. Using that small sample (a key teaching point about statistical power), it was nearly impossible to determine the risks of launching the shuttle at low temperatures. That determination was important for Challenger's expected launch temperature, since the rocket engines had not been qualified for launch below 40 degrees (Dalal et al. 1989:947).

The Challenger example meets the teaching example criteria outlined in Figure 1. The example is very concrete: when parts fail, machines can fail. Student connection is possible as a few students may remember the terrible and graphic television coverage of the shuttle explosion, with shattered shuttle pieces trailing streamers of smoke and fire as they fell into the ocean. The sample size is manageable, involving only 23 cases if each flight is used as a case or only 138 cases if each O-ring on each flight is used as a case. The proper use of statistics is crucial to the example as it might have delayed the launch to a safer, warmer temperature and saved the crew. The Presidential Commission (1986:148) concluded: "A careful analysis of the flight history of O-ring performance would have revealed the correlation of O-ring damage and low temperature. Neither NASA or Thiokol carried out such an analysis; consequently, they were unprepared to properly evaluate the risks of launching the 51-L mission in conditions more extreme than they had encountered before." Although the data are essentially bivariate, the statistics used can range from an examination of means and standard deviations to the use of ordinary least squares models. More advanced students could look into variable transformations, curve fitting models, and the use of logistic regression, one of the more difficult-to-grasp statistical methods in sociology (Lottes, Adler, and DeMaris 1996; Walsh 1987). The following role-play introduction

[1]The actual launch temperature was 36 degrees, although temperatures at various parts of the space vehicle ranged between 28 and 50 degrees (Presidential Commission 1986:70-81).

Table 1. Temperature, O-Ring Failure, and Number of O-Rings Failed for Each of 23 Space Shuttle Flights Prior to Challenger

Flight Number	Temperature at Launch	O-Ring Failure	Number of O-Rings that Failed
1	66	NO	None
2	70	YES	One
3	69	NO	None
4	68	NO	None
5	67	NO	None
6	72	NO	None
7	73	NO	None
8	70	NO	None
9	57	YES	One
10	63	YES	One
11	70	YES	One
12	78	NO	None
13	67	NO	None
14	53	YES	Three
15	67	NO	None
16	75	NO	None
17	70	NO	None
18	81	NO	None
19	76	NO	None
20	79	NO	None
21	75	YES	Two
22	76	NO	None
23	58	YES	One

Source: Christensen (1997:55; Dalal et al. 1989:946, 949).

is designed to make the students collaborators in the process rather than mere spectators and serves as anticipatory socialization. The role-play begins as follows:

> I am Dr. Smith, as you all know. The Challenger is set to launch tomorrow at an expected temperature at launch site of 31 degrees. Some of your colleagues have expressed reservations about launching at a temperature that is below that for which the shuttle's engines were engineered. I respect their concerns for the safety of the crew, but we have important political and customer pressures to proceed with the launch. Because of recent delays of various flights for various reasons, we can't afford yet another delay—we don't want the program to look like it has insurmountable technical difficulties. At the same time, I am willing to press the issue anyway, but I need some hard statistical data to make my case. Without that data, I cannot hope to overcome the political pressures. We don't have much time. I need your analyses back within a few hours. Give me whatever you can. Thanks.

At this point, I ask the students to use their current knowledge to obtain a solution, which may range from showing how 31 degrees compares to the mean and standard deviations of previous flights to using logistic regression to indicate more than a ninety percent chance of O-ring failure. Chi-square, t-test, correlational, ANOVA, and linear regression statistics may be used at various times during the class. Even though some statistics are more appropriate than others for the interval (temperature, number of O-ring failures) and nominal (O-ring failure) data, all may lend clues to assessing the risks of this particular launch temperature. It is rare that all the critical assump-

[2]Of course, the name "Dr. Smith" and the entire conversation is hypothetical and is not intended to resemble any actual conversations, but rather to involve the students in the role-play.

tions of a statistical test can be met perfectly, which should lend humility to all research(ers) (Blalock 1987:167). Parametric or nonparametric statistics may be used. For example, the Pearson zero-order correlation between temperature and number of O-ring failures (not including the Challenger flight data, of course) is -0.56 ($p \leq .01$). Students may need to collapse the number of O-ring failures into two or three categories rather than four categories for some analyses. The outliers that occurred at 75 degrees (two flights at O-ring failures) may be discussed as an issue (Lorenz 1987), since those outliers helped convince the scientists at the time that there was probably no relationship between temperature and O-ring failure. The assignment may be completed in class or as homework.

Ultimately, students may be referred to Christensen (1997:56), who argued that logistic regression would predict a probability of failure of about 94 percent, or to Lavine (1991), who mentions the problems of making predictions at any temperature so much lower than all previous flights. However, both Dalal et al. (1989) and Lavine (1991) estimated that four to six of the six shuttle O-rings would have failed at 31 degrees. Dalal et al. (1989) estimated that the probability of total shuttle catastrophe was at least 13 percent at 31 degrees compared to about two percent at 60 degrees. Nevertheless, no matter what procedures are used, almost all students will suggest that a launch at 31 degrees involved a higher, probably unacceptable, risk of O-ring failure.

THE PEARL HARBOR EXAMPLE

After the successful use of the Challenger example in class, the senior author found another example that appeared to meet many of the same criteria. While the losses experienced at Pearl Harbor in December 1941 (nearly 20 ships sunk, over 2,300 American lives lost, 68 Japanese lives lost) are well known, there has been a heated debate about whether the losses—and the subsequent war—were avoidable. A revisionist hypothesis has been developed by some scholars (Beach 1995; Stinnett 2000; Toland 1982) that the White House had advance warning of the "surprise" attack and did nothing to stop it as a politically expedient means to get the United States into World War II as a British ally. While other scholars have argued strongly against that hypothesis (Ambrose 1995; Millis 1947; Spector 1985), Stinnett (2000) presents the most current and detailed line of reasoning to support this "revisionist" hypothesis. However, at least one path allows students to compare the two theories statistically. Stinnett (2000:152), Beach (1995:47), and Kimmel (1955:62-63) report that on November 27, 1941 the Navy Department in Washington, DC ordered the most modern ships to leave Pearl Harbor on November 28 for maneuvers south of Wake Island and on December 5 for maneuvers south of Midway Island, in stark contrast to Admiral Kimmel's plan to send ships and planes north of Hawaii to ward off any Japanese threats from that direction. Unlike some aspects of the arguments for the revisionist hypothesis, the relative age of the ships at different locations can be evaluated statistically. The answer to the relative ages of the ships remaining at Pearl Harbor versus the ages of those ordered out of Pearl Harbor before December 7 is not intuitively obvious because some of the most modern, most recently-commissioned vessels were at Pearl Harbor while a few older ships were outside of Pearl Harbor.

The data (name and type of ship, location, year of commissioning) on the Pacific fleet's status are available in at least two sources: the Hearings before the Joint Committee on the Investigation of the Pearl Harbor Attack (U.S. Congress 1946) and Silverstone (1965, 1984). Naval task force locations as of December 7, 1941 are available in Wallin (1968).

In addition to the ships in Pearl Harbor, there were four major groupings of U.S. Navy ships elsewhere on the morning of December 7, 1941. One heavy cruiser, the

Minneapolis, and four World War I vintage destroyers were operating on a special mission some miles south of Oahu (labeled Task Force Zero in Table 2). Task Force 3, under Vice Admiral Wilson Brown, with the heavy cruiser Indianapolis and five World War I-vintage destroyer-minesweepers, was testing new landing craft in the vicinity of Johnston Island far south of Hawaii. Approximately two dozen other ships, mostly support vessels, were operating on their own, scattered for the most part to the south and west of Hawaii. Lastly, there were eighteen ships (a mix of destroyers, submarines, and support vessels) at Mare Island Navy Yard near San Francisco, California.

Of greater interest, Task Force 8, led by Vice Admiral "Bull" Halsey, consisting of the aircraft carrier Enterprise and three heavy cruisers and nine destroyers, was operating southeast of the small Hawaiian island of Kaula (southwest of Kaui), returning to Pearl Harbor after delivering Marine Corps fighter planes to Wake Island a few days earlier. Meanwhile, Task Force 12, led by Rear Admiral J. H. Newton, including the carrier Lexington and three heavy cruisers and five destroyers, was heading toward Midway Island (about four hundred miles away), to deliver military aircraft (Wallin 1968:54).

With a little work, students can assemble a database consisting of each ship, its date of completion or commission, the damage sustained on December 7 (if desired), and its location on December 7 (see Table 2 for a summary of this information). Students will find that the ships listed as being at Pearl Harbor vary from one source to another because some sources list smaller unnamed vessels (yardcraft, garbage scows, etc.) or named vessels (such as the tugboat Sotoyomo) (Slackman 1990) and some do not. In our analyses, those smaller vessels are not included. Students could be drawn into a real-life role-play by the following instructions:

> I am an aide to your state's senior senator, who has been reading Stinnett's (2000) book about Pearl Harbor. He has asked me to determine if there is any hard evidence to support Stinnett's claims. He wants statistical evidence, not mere subjective opinion. If he can obtain such evidence, he thinks he can persuade the Department of Defense to release still-classified messages that might confirm Stinnett's arguments. He needs the results back within three days. Can you do it?

Some of the results that might be obtained follow. The most obvious approach would be to compare the completion dates of vessels at Pearl Harbor with those of the two, three, or four primary task forces (0, 3, 8, and 12) at sea. The most relevant comparison would be with Task Forces 8 and 12 because it appears they were those most specifically directed to leave Pearl Harbor prior to the attack. However, it is not unreasonable to include the Task Force Zero since it was so close to Oahu. Because Task Force 3 was so far away, it might be reasonable to exclude it.

By using SPSS (Norusis 1993) to compare all the ships outside of Pearl Harbor, including those at Mare Island, with those at Pearl Harbor, one obtains a mean of 12.3 years (SD = 8.23) for those ships outside and an independent samples $t(118.08) = 2.05$ ($p \leq .05$), using unequal variances because the Levene test for equality of variances was significant ($p \leq .003$). Comparing Pearl Harbor with Task Forces 0, 8, and 12, we obtain a mean of 10.1 years (SD = 6.46) for ships outside the harbor, with an unequal variances $t(63.41) = 3.33$ ($p \leq .002$), with a significant Levene test ($p \leq .001$). Comparing Pearl Harbor with only Task Forces 8 and 12, we obtain a mean of 7.73 years (SD = 3.38) for ships outside of the harbor, with an unequal variances $t(98.04) = 6.23$ ($p \leq .001$), with a significant Levene test ($p \leq .001$). If we include Task Force 3, but not the Mare Island ships, we obtain a mean of 12.0 years (SD = 7.57), with an unequal variance t(71.55)

[3]More detailed data than presented in Table 2 are available from the senior author upon request.

Table 2. Years of Commissioning of U.S. Naval Vessels in the Pacific, December 7, 1941

	Number of Vessels with Each Location/Task Force					
Year of Commissioning	TF Zero	TF 3	TF 8	TF 12	Mare Island	Pearl Harbor
1907						1
1909						2
1912						1
1914					1	2
1915						1
1916					1	4
1917					1	1
1918	2	2			2	17
1919	2	2				2
1920		1				8
1921						6
1922					2	4
1923						2
1927				1		1
1929			3			
1930				1	1	
1932						
1933	1			1	1	1
1934						4
1935		1		3	3	8
1936			5	2	2	3
1937			2			10
1938			3			6
1939					2	3
1940					2	4
1941						2

Note: No ships were commissioned for those years that are missing in the table.
Sources: Silverstone (1965, 1984) and U.S. Congress (1946).

= 2.07 ($p \leq .05$), with a significant Levene test ($p \leq .001$). Using nonparametric tests yields similar or even more significant results. The effect sizes involved are substantial; for example, the effect size comparing the Enterprise Task Force to Pearl Harbor's ships is 0.86 (15.4 - 7.0)/9.73 (difference in means divided by the larger of the group's standard deviations). The overall analysis of variance is presented in Table 3.

Therefore, the scientific answer suggests that the ships sent out from Pearl Harbor were significantly more modern (in years) than those left behind, supporting Stinnett's (2000) contention and indirectly lending support to his overall hypothesis. Students should be reminded that data are subject to more than one interpretation. However, some facts are pertinent. First, the U.S. Navy already had a war plan, WPPac-46, that envisioned a carrier group off Wake Island counterattacking a Japanese force (Beach 1995:105, 164). Secondly, both Task Forces 8 and 12 "were in a status of wartime alert, fully armed and ready for any emergency" (Wallin 1968:54). Thirdly, the Chief of Naval Operations later testified that he ordered Kimmel to send planes and ships to Wake and Midway in order to catch

Table 3. Means and Standard Deviation for Age of Ships (Years) for Task Forces and for Ships at Pearl Harbor

Tsk Force/Site (Lead Ship)	Number of Ships	Mean	S.D.	Oldest Ship	Newest Ship
Zero (Minneapolis)	5	20.6	6.50	1918	1933
Three (Indianapolis)	6	20.5	6.66	1918	1935
Eight (Enterprise)	13	7.0	3.51	1929	1938
Twelve (Lexington)	9	8.8	3.07	1927	1936
Pearl Harbor	94	15.4	9.73	1907	1941
Mare Island Navy Yard	18	12.8	9.53	1916	1940
All Ships	145	14.3	9.32	1907	1941

An ANOVA for the above group yields $F(5,139) = 3.94$, $p \leq .003$, although the Levene test for homogeneity of variance is significant, $F(5,139) = 17.91$, $p \leq .001$. If the analysis is performed only for surface combat ships (excluding minelayers, seaplane tenders, supply ships, etc.), then $F(5, 91) = 4.53$, $p \leq .002$. with the mean scores for Task Forces Zero, Three, Eight, and Twelve remaining the same, while the mean for Pearl Harbor becomes 13.4 years and that for Mare Island being 9.7 years. It is interesting that both the most modern (8 ships built between 1939 and 1941) and the oldest (6 ships built between 1907 and 1915) vessels at Pearl Harbor were not surface combat ships, as were 11 of the 17 ships at Pearl Harbor that had been commissioned in 1918.

any enemy raids on those islands (Kimmel 1955:74), although Admiral Kimmel was never made aware of that intent nor of any specific risks of attack, other than sabotage, on Pearl Harbor. Lastly, while an attack on Pearl Harbor had long been considered a possibility, most Naval officers felt it was extremely unlikely and that an attack would probably occur elsewhere first (Holloway 1981:106; Layton 1981:278-79; Wellborn 1981:95-96), with Pearl Harbor as a possible follow-on objective (as happened in the Battle of Midway in 1942).

One interpretation of these facts offered to the students is that the Navy Department anticipated a smaller-scale Japanese attack on Wake or Midway to the west of Pearl Harbor and tried to set an ambush for such an attack, only to find itself embarrassingly ambushed by a much larger than anticipated Japanese force at its major port to the east, Pearl Harbor. Regardless of how others may interpret these data, Stinnett's comment that the ships in Pearl Harbor were older remained merely an unsubstantiated opinion until it was tested scientifically, using statistics. One of the purposes of statistics is to help professionals in a variety of careers to make decisions on something other than mere subjective opinion or conjecture.

THE TITANIC EXAMPLE

One reviewer of this article suggested that the Challenger and Pearl Harbor examples might appeal less to sociology students than to engineering or ROTC students, and cited a colleague who used the Titanic story to illustrate the impact of social class. The Titanic is indeed a particularly useful example for illustrating the complex interplay between theory and statistics. Many students will have seen at least one of the movies based on the Titanic disaster of April 15, 1912. The Titanic carried approximately 2,200 passengers but only enough lifeboats to accommodate 1,178, providing room to save all the women and children and 39 percent of the men on board

(Wormser 1994:27). However, far fewer passengers survived because the lifeboats were not filled and most did not return to the scene to rescue survivors from the icy waters. Data are readily available (see Table 4) on survival rates on the basis of social class (first-, second-, and third-class passengers and crew) and gender or family member status (male adult, female adult, child). When asked, most students will recall something about a higher percentage of women and upper-class passengers surviving. However, a variety of contemporary sociological theories might suggest that social class and gender should predict survival rates, with males and upper-class passengers using their social status and resources to gain priority for entry into lifeboats. For example, the wealthy, male owner of the ship, Mr. Bruce Ismay, demanded and obtained a seat in one of the life boats (Wormser 1994:32), even as social norms at the time dictated that males give up their potential seats on lifeboats to women and children. He was later scorned for his actions. How might theory about social norms interact with theories about social class and gender? As an example of what perhaps could be called "ancient" theory about social norms, students can be referred to Proverbs 30:9, which suggests that it is better to be of the middle class, because those in the lower class will violate social norms in order to survive while those in the upper class will violate social norms simply because they have the resources to get away with it. What evidence can be found in the Titanic data regarding each of these theories?

Gender theory is supported because some men accepted places in lifeboats, taking seats that "should" have gone, by contemporary social norms, to women and children who were still left on board. This idea can be tested using one-sample tests (actual survival rates of women versus a hypothetical survival rate of 100%). Social class theory receives support because women and children from the upper classes had higher survival rates than their third-class ticket counterparts. Ancient theory receives support because a higher proportion of second-class ticket holders' (middle-class) children survived than in either of the other two social classes, while a lower percentage of middle-class male adults survived than did adult males from either the upper or lower class (as shown by significant quadratic relationships). This suggests that the social norm of "women and children first in the lifeboats" received greater support among middle-class males than among the other two classes of men. At the same time, gender theory may be challenged to some degree

Table 4. Survivors and Casualties as a Function of Social Class and Status from the Titanic

Status	Class (Ticket)	Lost	Saved	Total Aboard
Men	First Class	118	57	175
	Second Class	154	14	168
	Third Class	381	75	456
	Crew Members	674	189	863
Women	First Class	4	139	143
	Second Class	15	79	94
	Third Class	89	76	165
	Crew Members	3	18	21
Children	First Class	1	5	6
	Second Class	0	23	23
	Third Class	53	26	79

Source: Butler (1998:239).

by the finding that a much higher percentage (86%) of female crew members survived than male crew members (22%), and an even greater ratio (10.5; 84%/8%) of female to male survivors will be found among the middle-class passengers. Because of the significant interaction effect between social class and gender/child status, it is necessary to analyze the effects of social class within each gender/child status in order to get a more complete understanding of how social class and gender relate to survival rates from the Titanic. Students can determine the relative amounts of variance explained by each theory, comparing main effects, interaction effects, and linear versus quadratic effects. Comparisons can be made between results obtained from analysis of variance approaches and statistical approaches designed for dependent variables with only two levels of outcome.

DISCUSSION

From the students' perspective, the graduate classes in which the historical examples were used (Spring 2000 and Fall 2001) were a success, with the classes being rated among the top five percent for overall course excellence and among the top ten percent for progress on relevant objectives. Notably, they were rated among the top one percent for effective analysis and critical evaluation of ideas and for improving student attitudes about the subject matter, the latter perhaps an indication of anxiety reduction. The ratings were not achieved on the basis of the instructor's personality, as the instructor was only rated among the top twelve percent of instructors (a somewhat lower rating than for the other course characteristics). Nevertheless, the ratings are probably above average for most methodology courses.[4] Student feedback was positive. Comments included:

> Excellent course. Really appreciated hands-on learning with SPSS and understandable teaching style professor uses in explaining complex concepts.
>
> Thank you for the effort that you put into class preparation.
>
> The Challenger example was excellent.
>
> A great class for learning and critical thinking.

We have not yet measured changes in long-term retention of statistical knowledge as a result of these techniques. However, pre-test and post-test results were available on several five-point items concerning self-reported competence for 10 students who participated in the Fall 2001 graduate class, as presented in Table 5. Reports for each item ranged from "not at all competent" (1) or "slightly competent" (2) to "somewhat" (3), "very" (4), or "extremely" competent (5). Results from Table 5 suggest that students generally believed they progressed from a sense of being "slightly competent" to "somewhat" or "very" competent. Most improvements were significant ($p \leq .05$, one-tailed).

We believe that these examples can be adapted to a wide variety of courses, both in and outside of sociology departments. Both examples show how even fairly basic statistics can be used to explore genuine historical situations involving life and death decisions. Of course, there are many other examples that are of interest, even if they do not involve actual or potential fatalities. In both cases, the data could be found without too much effort. The datasets involved are small enough that students can gain training in data entry if that is a part of the class; if

[4]Data from an IDEA report for the class, compiled by the IDEA Center, Kansas State University, www.idea.ksuledu, based on comparisons with 35,000 classes rated during the 1993-1994 and 1994-1995 academic years. However, the data for analysis and critical evaluation of ideas was a new item, normed against only 3,668 classes.

[5]One reviewer suggested simple analyses of the success of pumpectomies and age and of Caplovitz's (1963) thesis that the poor pay more, which has been debated in a variety of papers since (Andaleeb 1994; Sturdivant 1973; Williams 1973).

Table 5. Changes in Student Attitudes from Beginning to End of Semester Concerning Confidence in Dealing with Various Statistical Issues

Variable	Pre-test Scores		Post-test Scores		t	df	p<
	Mean	S.D.	Mean	S.D.			
Entering data	2.44	1.13	3.90	0.57	3.49	11.52	.003
Checking data entered	2.22	1.30	3.30	0.82	2.13	13.27	.027
Correcting data entry errors	2.44	1.24	3.50	0.97	2.05	15.20	.030
Using SPSS to compute basic descriptive statistics	2.56	1.74	4.10	0.88	2.40	11.53	.018
Interpreting SPSS Frequencies program output	2.44	1.74	3.70	0.82	1.98	11.15	.038
Using SPSS to compare mean scores	2.44	1.51	4.10	0.74	2.99	11.35	.007
Interpreting mean score comparisons	2.22	1.30	3.60	0.52	2.97	10.24	.008
Using SPSS to compute basic correlations	2.44	1.51	3.90	0.88	2.54	12.56	.013
Interpreting basic correlations	2.33	1.50	3.30	0.67	1.78	10.86	.052
Using SPSS to perform cross tabulations	2.33	1.50	3.10	0.99	1.30	17.00	.102
Interpreting chi-square results and cross-tabulations	2.22	1.48	3.00	0.82	1.40	17.00	.070
Using SPSS to solve real world problems	1.78	1.09	3.40	0.84	3.64	17.00	.002

Note: When degrees of freedom feature decimal points, unequal variances t-tests were performed because of a significant lack of homogeneity of variance. Otherwise, equal variances t-tests were used. One-tailed tests of significance reported.

not, the datasets are small enough to be given to students on computer diskettes. Both examples can be used in a "bare-bones" fashion using only the historical information provided in this report or they can be expanded with more detailed historical introductions to the climate of the times and the specific circumstances surrounding each example. It is possible that use of such storylines may help to arouse or maintain student involvement in the class material. Woodberry and Aldrich's (2000) discussion of how to conduct classroom-based exercises lends some insight to how we will improve the use of these examples in future classes. In our class, we did not try to time the use of the exercise to any portion of the class period, but Woodberry and Aldrich (2000:242) are probably correct in making the case that the exercise will stimulate the class more effectively if used at the beginning of the class period. We agree with their assessment that exercises that get students emotionally involved will generate excitement and keep students on task, which was part of our experience with the Challenger example. We also think that at least some examples should be introduced at the beginning of the course and used throughout the term, with more difficult statistical methods applied to solve the problems as the students gain skill. As suggested by Woodberry and Aldrich (2000:243), specific instructions were given for the exercise and specific results were expected: a briefing on the findings to the simulated administrator. While we didn't call it that at the time, a debriefing did occur, as recommended by Woodberry and Aldrich (2000:245) in which the issue of statistics overcoming the subjectivity issue was clarified. In both the Challenger and the Pearl Harbor examples, one might easily develop a subjective hypothesis about the situation, but it would be difficult to assess the certainty of one's assessment. The use of statistics helps clarify whether the outcome is merely an unconfirmed perception or perhaps a valid, meaningful observation (as opposed to an artifact of random error). As for logistics and evaluation, the activities may be completed in class or as homework and either graded or not graded, depending on the situation. We think that the initial exercises should not be graded and used more to stimulate interest, while later in the course grading might be a valid way to assess student learning in terms of comprehension and analysis.

In future use of these examples, we will elicit more detailed feedback on how to improve the material and on additional suggestions for similar, concrete historical examples that students might come across in their studies, in accordance with Woodberry and Aldrich's (2000:245-46) goal of using regular student feedback as a part of practicing continuous improvement of classroom exercises. In addition to their use in methodology courses, these types of examples could be also used in content courses that include topics such as inequality, bureaucracy, social organization, sociology of work and occupations, social control, and the sociology of history. Such use might minimize the problem Weiss (1987:191) identified as the artificial separation of method and content. If class sizes are larger or not enough computers are available, students could work the assignment in small groups as recommended by Longmore et al. (1996) without losing too much of the learning experience. Another alternative would be to use statistical packages within a UNIX environment, which has fewer limitations in terms of class size.

The use of these particular examples overcomes some of the problems that Halley (1991) tried to solve with simulated data: that large data sets often had substantial missing data problems, while small data sets had insufficient statistical power to allow for many significant findings. These examples involve small data sets but have no missing data and do yield significant results. These kinds of real-life examples may also help bridge a link to short-term (out-of-classroom) experiential learning, as advocated recently by Wright (2000).

Some instructors may note that our examples will challenge conventional assumptions about the efficiency and trustworthiness of some government officials; it is possible that students may resist analyses that challenge government competence or credibility (Silver and Perez 1998). Finally, we believe that use of concrete examples, followed up by more abstract examples or more detailed out-of-classroom experiences, will help students *overlearn*, rather than merely overmemorize, statistics (Fischer 1996; Blalock 1987).

REFERENCES

Ambrose, Stephen E. 1995. "'Just Dumb Luck': American Entry into World War II." Pp. 93-102, *Pearl Harbor Revisited*, edited by Robert W. Love, Jr. New York: St. Martin's Press.

Andaleeb, Syed S. 1994. "Do the Poor Pay More? A Developing Country Perspective." *Journal of International Consumer Marketing* 7:59-72.

Auster, Carol J. 2000. "Probability Sampling and Inferential Statistics: An Interactive Exercise Using M&M's." *Teaching Sociology* 28:379-85.

Beach, Edward L. 1995. *Scapegoats: A Defense* *of Kimmel and Short at Pearl Harbor*. Annapolis, MD: Naval Institute Press.

Blalock, Hubert M., Jr. 1987. "Some General Goals in Teaching Statistics." *Teaching Sociology* 15:164-72.

Bloom, Benjamin S., Max D. Engelhart, Edward J. Furst, Walker H. Hill, and David R. Krathwohl. 1956. *Taxonomy of Educational Objectives, Handbook I: Cognition Domain*. New York: Longman.

Borreson, Robert C. 1990. "Success in Introductory Statistics with Small Groups." *College Teaching* 38:26-28.

Bridges, George S., Jana L. Pershing, Gerald M. Gillmore, and Kristin A. Bates. 1998. "Teaching Quantitative Research Methods: a Quasi-Experimental Approach." *Teaching Sociology* 26:14-28.

Butler, Daniel A. 1998. *Unsinkable: Story of RMS Titanic*. Mechanicsburg, PA: Stackpole Books.

Caplovitz, David. 1963. *The Poor Pay More: Consumer Practices of Low-Income Families*. New York: The Free Press.

Christensen, Ronald. 1997. *Log-Linear Models and Logistic Regression*. 2d ed. New York: Springer-Verlag.

Dalal, Siddhartha R., Edward B. Fowlkes, and Bruce Hoadley. 1989. "Risk Analysis of the Space Shuttle: Pre-Challenger Prediction of Failure." *Journal of the American Statistical Association* 84:945-57.

Fischer, Henry W., III. 1996. "Teaching Statistics from the User's Perspective." *Teaching Sociology* 24:225-30.

Halley, Fred S. 1991. "Teaching Social Statistics with Simulated Data." *Teaching Sociology* 19:518-25.

Holloway, James L. 1981. "Cutting the Red Tape." Pp. 106-108 in *Air Raid: Pearl Harbor! Recollections of a Day of Infamy*, edited by Paul Stillwell. Annapolis, MD: Naval Institute Press.

Kain, Edward L. 1987. "Using a Research Team and Microcomputers to Teach Methodological Principles." *Teaching Sociology* 15:136-43.

Karp, David R. 1995. "Using SPSS for Windows to Enhance, Not Overwhelm, Course Content." *Teaching Sociology* 23:234-40.

Kimmel, Husband E. 1955. *Admiral Kimmel's Story*. Chicago, IL: Henry Regnery Company.

Lavine, Michael. 1991. "Problems in Extrapolation Illustrated with Space Shuttle O-Ring Data." *Journal of the American Statistical Association* 86:919-21.

Layton, Edwin T. 1981. "Admiral Kimmel Deserved a Better Fate." Pp. 272-85 in *Air Raid: Pearl Harbor! Recollections of a Day of Infamy*, edited by Paul Stillwell. Annapolis, MD: Naval Institute Press.

Longmore, Monica A., Dana Dunn, and Glen R. Jarboe. 1996. "Learning by Doing: Group Projects in Research Methods Classes." *Teaching Sociology* 24:84-91.

Lottes, Ilsa L., Marina A. Adler, and Alfred DeMaris. 1996. "Using and Interpreting Logistic Regression: A Guide for Teachers and Students." *Teaching Sociology* 24:284-98.

Lorenz, Frederick O. 1987. "Teaching about Influence in Simple Regression." *Teaching Sociology* 15:173-77.

Martz, H.F. and W.J. Zimmer. 1992. "The Risk of Catastrophic Failure of the Solid Rocket Boosters on the Space Shuttle." *The American Statistician* 46:42-47.

Millis, Walter. 1947. *This is Pearl: The United States and Japan - 1941*. New York: Morrow.

Norusis, Marija J. 1993. *SPSS for Windows: Professional Statistics*. Chicago, IL: SPSS, Inc.

Potter, Alisa M. 1995. "Statistics for Sociologists: Teaching Techniques That Work." *Teaching Sociology* 23: 259-63.

Presidential Commission. 1986. *Report of the Presidential Commission on the Space Shuttle Challenger Accident*. Retrieved May 8, 2000 (http://history.nasa.gov/rogersrep/51lcover.htm).

Renzulli, Linda A. 2000. "Connecting the Classroom to County Characteristics." *Teaching Sociology* 28:249-60.

Rushing, Beth and Idee Winfield. 1999. "Learning about Sampling and Measurement by Doing Content Analysis of Personal Advertisements." *Teaching Sociology* 27:159-66.

Salkind, Neil J. 2000. *Statistics for People Who (Think They) Hate Statistics*. Thousand Oaks, CA: Sage.

Schacht, Steven and Brad J. Stewart. 1990. "What's Funny about Statistics: A Technique for Reducing Student Anxiety." *Teaching Sociology* 18:52-56.

Schumm, Walter R. 2001. "The Admiral's Son with an Independent Spirit." *Marriage and Family Review* 31(3/4):155-79.

Silver, Ira and Gina Perez. 1998. "Teaching Social Theory Through Students' Participant-Observation." *Teaching Sociology* 26:347-53.

Silverstone, Paul H. 1965. *U.S. Warships of World War II*. Garden City, NY: Doubleday.

_____. 1984. *Directory of the World's Capital Ships*. New York: Hippocrene Books.

Slackman, Michael. 1990. *Target: Pearl Harbor*.

Spector, Ronald H. 1985. *Eagle Against the Sun: The American War with Japan*. New York: The Free Press.

Steen, Sara, Chris Bader, and Charis Kubrin. 1999. "Rethinking the Graduate Seminar." *Teaching Sociology* 27:167-73.

Stinnett, Robert B. 2000. *Day of Deceit: The Truth about FDR and Pearl Harbor.* New York: The Free Press.

Sturdivant, Frederick D. 1973. "Rationality and Racism in the Ghetto Marketplace." *Social Science Quarterly*, 54:380-83.

Takata, Susan R. and Wanda Leiting. 1987. "Learning by Doing: The Teaching of Sociological Research Methods." *Teaching Sociology* 15:144-50.

Toland, John. 1982. *Infamy: Pearl Harbor and its Aftermath*. New York: Doubleday.

United States Congress. 1946. *Hearings Before the Joint Committee on the Investigation of the Pearl Harbor Attack.* (Part 12, Exhibits 1 through 6, pp. 345-349). Washington, DC: U.S. Government Printing Office.

Vaughn, Diane. 1996. *The Challenger Launch Decision.* Chicago, IL: University of Chicago Press.

Wallin, Homer N. 1968. *Pearl Harbor: Why, How, Fleet Salvage, and Final Appraisal.* Washington, DC: U.S. Government Printing Office.

Walsh, Anthony. 1987. "Teaching Understanding and Interpretation of Logit Regression." *Teaching Sociology* 15:178-83.

Watts, Donald G. 1991. "Why is Introductory Statistics Difficult to Learn? And What Can We Do to Make It Easier?" *The American Statistician* 45:290-91.

Weiss, Gregory L. 1987. "Using a Local Research Center to Enhance the Teaching of Research Methodology." *Teaching Sociology* 5:191-94.

Wellborn, Charles, Jr. 1981. "The Fog of War." Pp. 95-97 in *Air Raid: Pearl Harbor! Recollections of a Day of Infamy*, edited by Paul Stillwell. Annapolis, MD: Naval Institute Press.

Whitmore, Fred. 2000. "The Unbeatable Ship That Nobody Ever Heard of." *U.S. Navy Cruiser Sailors Association Newsletter* 9(3):1-5.

Williams, Walter E. 1973. "Why the Poor Pay More: An Alternative Explanation." *Social Science Quarterly* 54:375-79.

Wilmoth, Janet and John Wybraniec. 1998. "Profits and Pitfalls: Thoughts on Using A Laptop Computer and Presentation Software to Teach Introductory Social Statistics." *Teaching Sociology* 26:166-78.

Woodberry, Robert D. and Howard E. Aldrich. 2000. "Planning and Running Effective Classroom-Based Exercises." *Teaching Sociology* 28:241-48.

Wormser, Richard. 1994. *The Titanic*. New York: The Trumpet Club.

Wright, Mary C. 2000. "Getting More Out of Less: The Benefits of Short-Term Experiential Learning in Undergraduate Sociology Courses." *Teaching Sociology* 28:116-26.

Wybraniec, John and Janet Wilmoth. 1999. "Teaching Students Inferential Statistics: A 'Tail' of Three Distributions." *Teaching Sociology* 27:74-80.

Walter Schumm, Ph.D. (Purdue University), is professor of family studies at Kansas State University and director of the Ohio Desert Storm Research Project. He is a co-editor of *The Sourcebook of Family Theories and Methods*. His current research interests include military families, the role of religion in family life, evaluation of premarital education, and research methodology in the study of the family. His articles have appeared in *Journal of Marriage and the Family, Social Forces, Journal of Family Issues, Journal of Political and Military Sociology, Marriage and Family Review, Contemporary Family Therapy, The Family Journal, The Hispanic Journal of Behavioral Sciences, The American Journal of Family Therapy, The Journal of Comparative Family Studies, Family Relations, Review of Religious Research, International Journal of Sociology of the Family, Families in Society, Military Medicine, Military Review, and Military Chaplains' Review*, among others. He has lectured in family theory, advanced family life education, survey research methods, and advanced family data analysis. His invited autobiography was recently published in *Marriage and Family Review* (Schumm 2001).

Farrell Webb, Ph.D. (University of Minnesota), is associate professor of family studies at Kansas State University and lectures in social theory, diversity, family policy, human sexuality, and research methodology. His research interests include racial/ethnic attitude development, adolescents with HIV/AIDS, and human sexuality. His articles have appeared in *Journal of Comparative Family Studies, Psychological Reports, and Health Values*, among others. E-mail: fwebb@ksu.edu

Carlos Castelo, Ph.D. (Kansas State University), passed away on October 21, 2001. His research and teaching interests included remarried families and advanced research methodology.

Cynthia Akagi, Ph.D. (Kansas State University), lectures at the School of Education, University of Kansas in sexuality, health, and nutrition. Her research interests include evaluation of health education and adolescent sexuality issues. She is the author of two

interests include evaluation of health education and adolescent sexuality issues. She is the author of two sexuality education books, *Dear Michael* and *Dear Larissa*, published by Gylanta Publishing Company, Littleton, CO.

Erick Jensen, Ph.D. (Kansas State University), lectures in life and social sciences at the public magnet school, Sumner Academy of Arts and Science, Kansas City, KS.

Rose Ditto, Ph.D. (Kansas State University), is in private practice in Wamego, Kansas and lectures regionally and nationally on issues related to child development and parent education.

Elaine Spencer-Carver, M.S.W. (University of Kansas), is a doctoral student in family life education at Kansas State University and lectures on issues related to family violence and family relationships.

Beverlyn F. Brown, M.S.W. (University of Kansas), is a doctoral student in family life education at Kansas State University.