Statistical Literacy Curriculum Design

Milo Schield
Augsburg College
Minneapolis MN, USA

Abstract

College-level students pursuing majors that don't require a quantitative course still need a statistical literacy course that helps them develop the skills to evaluate arguments that use statistics as evidence. Such a course should entail utility in everyday use such that statistical literacy results in a lasting appreciation of the value of statistics as needed in everyday life, civic life, and professional life as a data consumer. A course designed to promote statistical literasy should help students understand and analyze various influences on the size and direction of a statistical association and should include key topics in conditional probability, confounding, and the vulnerability of statistical significance to confounding. This paper describes some new ways of presenting these ideas that are based on the results of field trials conducted in connection with the W. M. Keck Statistical Literacy project at Augsburg College. After studying statistical literacy, 43 percent of Augsburg students strongly agreed that the course helped them develop critical thinking skills and 18 percent strongly agreed that successful completion of the course should become requirement for graduation.

Focus and Philosophy of a Statistical Literacy Curriculum

In an oral presentation at IASE in Korea, Moore (2001) distinguished statistical literacy ("What every educated person should know about statistical thinking") from statistical competence ("roughly the content of a first course for those who must deal with data in their work … or what we hope a statistics student will retain five years later"). In his power point slides, Moore indicated that statistical literacy involves two clusters of "big ideas": 1) "The omnipresence of variation, conclusions are uncertain, avoid inference from short-run irregularity, [and] avoid inference from coincidence." 2) "Beware the lurking variable, association is not causation, where did the data come from? [and] observation versus experiment."

Using Moore's definitions, statistical literacy is for data consumers while statistical competence is for data producers.

Changes in Data in the 21st Century

A sound curriculum should be based on the subject matter of the discipline. In statistics, the subject is data—not just data as numbers studied by mathematics but data in context. Several pervasive and important changes are occurring in the forms and kinds of data available to the general public. Data from large-scale studies is increasingly available. As sample sizes get larger, smaller differences become statistically significant. For example, a particular result from the National Longitudinal Survey of Youth (NLSY), which involved 12,000 subjects, found a 0.4 point difference in IQ test scores between men and women to be statistically significant at a 5 percent level (Bureau of Labor Statistics, 1979; www.bls.gov/nls/). A statistically literate person is equipped to recognize the affect of sample size on statistical significance. Thus in large data sets a statistically-literate person should anticipate that statistical significance is almost meaningless since relations between any two variables are expected to be statistically significant. They can anticipate that understanding the influence of context and confounding is much more relevant than understanding random variation and statistical significance. Such knowledge and skills are becoming increasingly important as population data (i.e., studies based on large samples) become more common influences on decision-making for both governments and for businesses.

Additional changes in the available data that impact the statistical literacy curriculum relate to the fact that observational studies are becoming more common in research. For example, in medical journals, articles involving observational studies (37.2 percent) were 50 percent more prevalent than those involving randomized trials (24.7 percent). Among related news stories, those involving observational studies (58 percent) were 10 times as common as those involving randomized trials (6.2 percent).[1] Doctors and the public are increasingly exposed to conclusions and recommendations arising from non-random, unblinded studies – both of which are conditions that influence the appropriateness of generalizing the results to other persons and settings.

Statistically literate people understand and can react appropriately to the increasing number of so-called experiments that have properties of observational studies. One example of such research comes from education, where Chance (2002) noted that teachers and students can be randomly assigned to classes or teaching methods but not effectively blinded as to which teaching methods are being used. Students should be taught to distinguish statistical experiments from observational studies. A controlled medical study involving random sampling is not necessarily a randomized experiment. Consumers of information arising from non-blinded experiments need to be aware that this limitation affects the conclusions that can be drawn from the study and that this limitation is not nullified by the presence of other elements of true experimentation (e.g., randomization).

The limitations of non-randomized studies are profound, yet observational studies are playing an increasingly important role in political and public policy debates and decisions. For example, John Lott's (1998) text *More Guns; Less Crime* used observational longitudinal crime data to advocate states passing concealed-carry laws. Title 1 funding for education is now tied to results from an observational study called the National Assessment of Educational Progress (NAEP).

Another major limitation of observational research is its greater tendency, over that of experimental research, to suffer the affects of a phenomenon called Simpson's Paradox (a reversal of an association after accounting for the influence of a lurking variable or confounder). Terwilliger and Schield (2004) found 52 reversals (10 percent) of statistically-significant differences in state NAEP scores for one confounder. Simpson's paradox reversals stem from non-randomization, which is inherent in all observational studies. The average data consumer needs the types of educational experiences that, at the very least, encourage them to question and intelligently consider the role that non-randomization and non-blindedness may have played in producing the reported research findings.

Motivation of Students to Study Statistics

A sound curriculum should take into account student motivation toward the subject. During the first week, 47 Augsburg business students taking statistics were asked "Would you take the following business courses even if they were not required?" Figure 1 shows the percentage by course who answered 'Absolutely' (rather than Almost Certainly, Likely, Unlikely or Absolutely Not). For statistics, 24 percent said 'Absolutely Not' while 52 percent said they were 'Unlikely' to do so. One explanation for this difference is that prospective judgments are lower than retrospective. Most students are seniors so their judgments on other courses are retrospective while those on statistics are prospective.

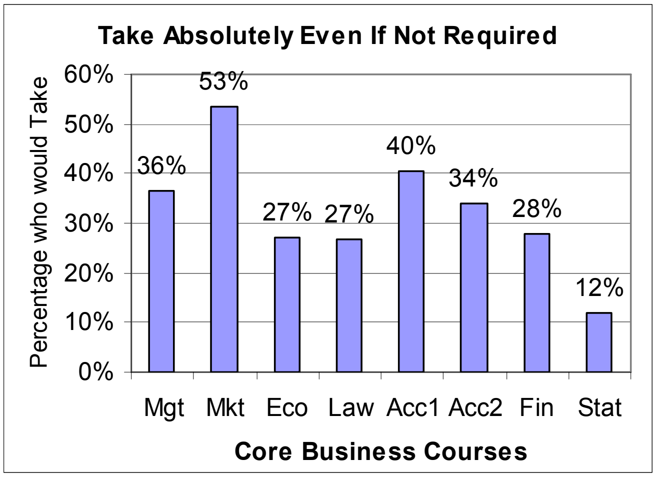

Figure 1: Willingness to Take the Course

These students were also asked about their major and to their attitude toward mathematics on a five-point scale: Like Very Much (4), Like Somewhat (3), Neutral (2), Dislike Somewhat (1), and Strongly Dislike (0). Binary groups were formed from all three variables: willingness to take statistics as an elective, major, and attitude toward math. Like Very Much (4) and Like Somewhat (3) were combined into Like; the rest were combined into Dislike.

Table 1
Percentage of Business Majors Who Would Absolutely or Almost Certainly take Statistics as an Elective – Classified by Major and by Attitude toward Math

Business Students	*Attitude Toward Math*		
MAJOR	ALL	Like	Dislike
ALL	22%	29%	16%
Accounting/Finance/Economics/MIS	28%	38%	21%
Management/Marketing	7%	12%	0%

As shown in Table 1, those who would absolutely or likely take statistics if it were an elective are 22 percent of these business students, 16 percent of those who dislike mathematics, and 0 percent among the less-quantitative majors (marketing and management) who dislike mathematics: a combination that describes many liberal arts majors.

This lack of perceived value for statistics has been documented by Schau (2003) who conducted the Survey of Attitudes toward Statistics (SATS). After completing a traditional course, student's attitudes were more polarized (increased standard deviations) than before and there was a statistically significant 8 percent decrease in the value they perceived in the subject.[2] Robert Hayden (private communication) noted, "Students will not use what they learn in a statistics course (of any kind) unless they believe they learned something usable. So we have to both provide something useful AND sell them on it." This lack of value for statistics was evidenced when 190 students at Pomona College ranked Critical Thinking first in value among 10 core competencies but ranked Data Analysis last (Taylor, 1999). Macnaughton (2004) has argued that the primary goal of an introductory statistics course should be to give students "a lasting appreciation for the value of statistics." By that criterion, the traditional course may not be achieving the Macnaughton goal. One explanation for this may be a lack of focus on observational studies and causation. About 70 percent of those studying statistics are in majors that focus primarily on using observational data.[3]

Role of Statistics in Everyday, Business and Civic Life

A sound statistical literacy curriculum should be based on the role of statistics in a student's daily, civic or professional life. Taking a sample publication related to each arena, Table 2 illustrates differences in the degree and types of exposure to statistical information that the regular reader would experience. Specifically, the table shows the prevalence of published articles using specific statistical terms. Articles involving confidence or significance are as high as 18 percent of all articles in *Nature* but no more than 0.1 percent of all articles in *The Economist*.[4]

Table 2
Prevalence of Articles Using the Term per 100,000 Articles by Source

STATISTICAL TERM	*Nature Magazine*	*USA Today*	*The Economist*	*Yahoo Search*
Mean/Median/Mode	36,456	7,996	890	316
"more than"	33,943	16,945	33,992	368
Rate	31,763	7,647	22,870	1,837
Sample	29,263	778	1,598	807
Risk	20,962	4,003	16,474	837
Percentage	14,968	3,234	3,745	260
Random	12,592	662	1,243	458
Probability	7,685	109	481	134
Chance	6,306	7,550	13,096	586
Standard deviation	4,058	3	0	19
Random Sample	686	10	15	11
"likely than"	488	113	391	14
Percentage Points	64	717	1,849	13
CONFIDENCE				
Standard Error	2,164	0	0	7
Confidence Interval	1,817	0	0	9
Confidence Level	336	26	0	6
Level of Confidence	62	9	9	4
Margin of Error	32	359	44	4
SIGNIFICANCE				
Statistically significant	6,512	14	47	12
Statistical significance	3,474	1	3	6
P-value	2,982	0	12	7
Level of Significance	449	0	0	2

Table 3 summarizes the percentage of young workers (18-25) in business who use a particular statistics topic (see Holmes, 1981). While 54 percent read and interpret tables of data and 37 percent make decisions using data, only 6 percent use a statistical test of significance. These data involve those working in business, but if statistical topics are over-emphasized for those in business, it seems that such topics would definitely be over-emphasized for average citizens.

Table 3
Frequency of Use

60%	draw up tables of data
54%	read and interpret tables of data
53%	assess the accuracy of someone else's data
53%	write reports based on data for others
52%	decide what data to collect
51%	calculate the mean
40%	detect and estimate trends
38%	simplify tabulated data
37%	allow for variability in data
37%	make decisions using data
35%	make projections
27%	draw bar charts and time series graphs
20%	use words such as likely and uncertain
19%	calculate variance or standard deviation
19%	use logarithm or other specialist scales
19%	draw trend lines; read/interpret histograms
17%	calculate median and quartiles
17%	assign probabilities to events
15%	allow for non-response to questionnaires
14%	select the questions on questionnaires
14%	read and interpret scatter diagrams
13%	use statistical tests to compare sets of data
13%	use probability as a measure of uncertainty
12%	read and interpret results of simulations
9%	calculate correlation coefficients
8%	calculate moving averages
6%	use a statistical test of significance
4%	use the normal distribution
2%	calculate index numbers

Stroup and Jordan (1982, 1983) surveyed 151 business statistics teachers and 1,495 business managers regarding the content of their statistics courses. Figure 2 illustrates the percentage of these statistics teachers who teach a given topic versus the percentage of those teaching a topic who teach it extensively. While both data sets are over 20 years old, we are unaware of anything more recent and based on our experience in business, these data appear to remain fairly accurate.

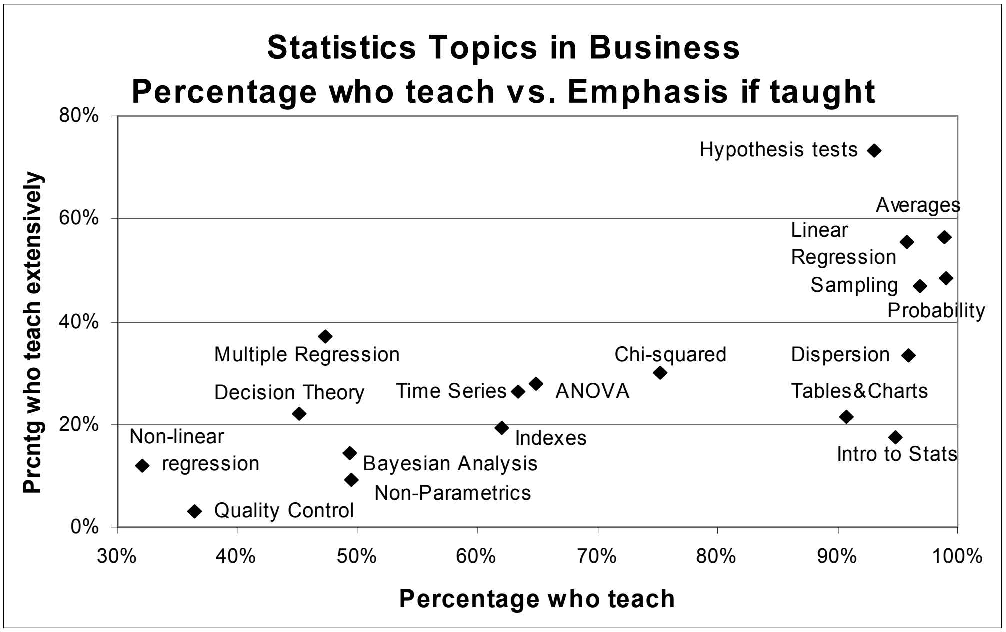

Figure 2. Topics Taught in Business Statistics

Figure 3 illustrates topics in introductory business statistics courses classified by how extensively they are taught and their usage in business; the horizontal axis is the percentage of business managers who, according to Stroup & Jordan (1982, 1983), use this topic. Note the negative association between what is taught and what is used.

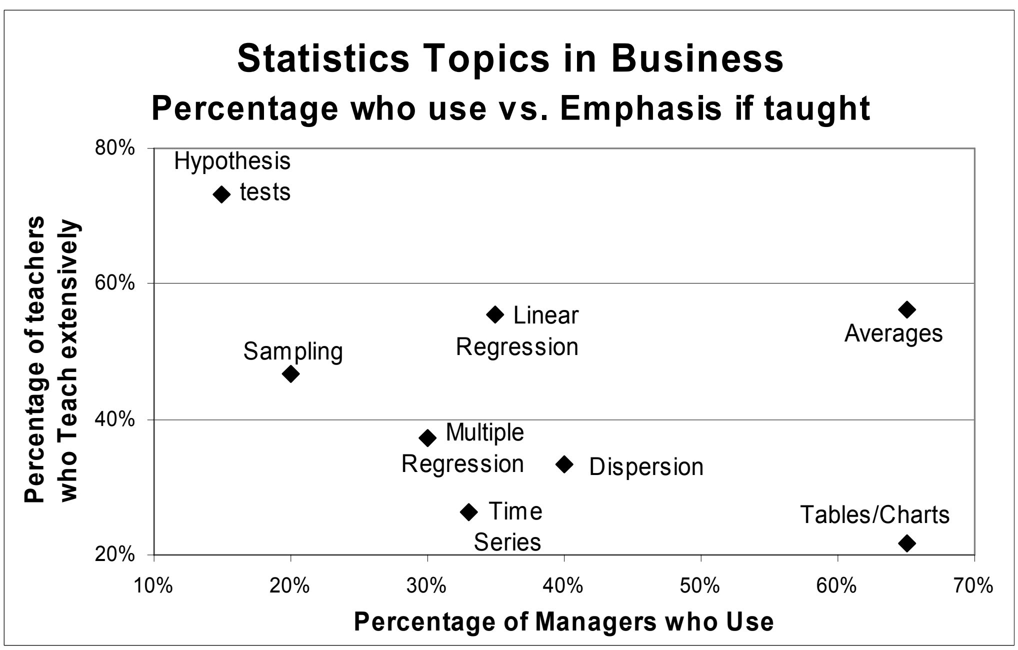

Figure 3. Statistics Topics Used in Business

The mismatch illustrated above, compounded by the lack of emphasis on college statistics on the Graduate Management Admission Test (GMAT)[6], may contribute to business majors' lower level of motivation to learn statistics. If usage in business were the standard for determining the content of instruction, then hypothesis tests are over-taught while multiple regression, time series, dispersion and tables/charts are under-taught. If we look at the use of statistics as evidence in public policy or historical analysis it appears that multivariate analysis is more central than statistical significance. For example, in "*The Bell Curve*" Herrnstein and Murray (1994) used multivariate analysis to argue that low IQ had a stronger association with social problems than did low socioeconomic status, and in "*More Guns; Less Crime*", John Lott (1998) used multivariate analysis to argue that passing concealed-carry handgun laws was associated with reduced crime. In addition, in "*The Great Breakthrough and Its Causes*", Julian Simon (2000) used historical data to argue that the increase in the size of the free, skilled population is the fundamental cause of the dramatic improvements in human welfare.

Key Topics in Statistical Literacy

A sound curriculum should reflect the key topics, tools and principles in that discipline. Examining what the leaders in the field actually do can identify these. This approach may work well in deciding what to teach data producers (statistics majors or minors for whom statistics is what statisticians do), but this may not work at all in deciding what to teach data consumers in a statistics service course. Since the subject of statistics is data, the key statistical ideas can be organized in terms of data.

One way of organizing statistical topics is by the method of acquiring the data – the kind of study involved. In small-sized, well-designed experiments (treatments with random assignment), chance (random error) dominates. In poorly-designed studies bias (systematic error) can dominate regardless of the size of the study. In populations or large-scale, well-designed observational studies, confounding (the influence of a lurking variable) dominates. As large, well-designed observational studies become increasingly common, the study of confounding becomes increasingly important. Holmes (2003, p. 5) said:

> When you have huge data sets, which are essentially populations, it isn't the sampling variability that's important. It is, the actual figures themselves and what are the connections between them. That is an important part of what I would now put in statistical literacy – which I wouldn't have put in 20 or 30 years ago, because there wasn't so much of this sort of stuff around.

A second way of organizing statistical topics is by the kind of data required, whether univariate, bivariate or multivariate. Chance and inference apply to all three kinds of data as does modeling so univariate or bivariate thinking is sufficient. But confounding requires multivariate data and multivariate thinking. Note that most observational data are multivariate. The results may be presented one predictor at a time (pair-wise using bivariate techniques), but the underlying data are multivariate. Multivariate data are seldom studied in the introductory course but are generally the focus of a second course on regression. While an estimated 60 percent of college students take an introductory statistics course, very few take a second course that introduces multivariate thinking (Schield, 1999a). As a result, the understanding of those taking just the introductory course is biased towards inference and away from confounding as compared with the understanding of those taking both courses.[7] The emphasis on statistical inference rather than on multivariate thinking in the introductory course may reflect the difficulty students have with conditional reasoning ("How likely is this outcome *if* due to chance?" versus "How likely is this outcome *to be* due to chance?"- P(Data|Chance) versus P(Chance|Data)). It may also reflect the historic difficulty of teaching multivariate thinking without first teaching an entire course on multivariate regression.

The controversy over hypothesis tests is longstanding (Morrison and Henckle, 1970; Harlow et al, 1997). A reduced emphasis on hypothesis testing has been recommended by a business statistics association called Making Statistics More Effective in Schools and Business (MSMESB) and by the American Psychological Association (APA)—with the MSMESB also advocating an increased emphasis on confidence intervals (see www.MSMESB.org)[8] and the APA advocating increased emphasis on effect sizes (see www.apa.org/science/tfsi.html for the 1996 report and www.apa.org/monitor/may99/task.html for the 1999 report). The use of multivariate summary data is more common in the social sciences and business.

Podehl (2003) advocates using official statistics to assist students in dealing with current social issues. Bregar (2003) shows how Economic Statistics can be taught using official statistics.

Statistical Thinking and Statistical Reasoning

A sound curriculum should reflect the kinds of reasoning used in the discipline. As a branch of mathematics, statistics uses deductive logic and conditional reasoning in reasoning from a population to random samples. There are disagreements on the importance of conditional reasoning—debates about which are documented on the MSMESB website (www.MSMESB.org). With regard to conditional probability, the MSMESB argued for less emphasis. Rossman and Short (1995) raised concerns while Dawes (2001) advocated more emphasis on informal uses of conditional probability. While Berry (1997) argued persuasively that students should be exposed to Bayesian thinking[9] in the first course, Moore (1996) argued conversely noting that students lacked the required skill in conditional reasoning..

As an art, statistics must deal with data-related questions internal to the data: what outliers to ignore, what transformations to make, what level of non-response to ignore, which model is best, what interactions to model, and how sensitive the model is to a change in the data. In addition to these, statistical literacy must also deal with data-related questions external to the data. How do the following relate to the argument at hand: the choice of the population and the outcome of interest; the choice, definition and connotation of each variable; and the vulnerability of both the value and the statistical significance of an observed association to an unknown confounder?

In reasoning 'beyond' the data, statistical literacy should be concerned with informal, inductive, pre-inferential, reasoning as well as with formal, deductive, inferential reasoning. Finzer and Erickson (2005, p. 195). encourage a stronger focus on pre-inferential reasoning, claiming that "many students have trouble with inferential statistics later because they do not have enough experience with making these more obvious quantitative arguments and displays" Pfannkuch and Horring (2005) also call for more attention to "informal inferences" to help school teachers teach statistics.

Curriculum Determinants

A sound curriculum should reflect the time available during the course in relation to the time required to present the underlying concepts necessary to properly understand a topic of interest. To include more advanced topics within a given time limit, educators must find ways to do so without giving a detailed presentation of all the supporting topics: teaching about the binomial distribution without a detailed study of combinations and permutations, teaching conditional probability without teaching algebra, teaching logistic regression without a full exposition of maximum likelihood, teaching multivariate thinking without teaching multiple regression, and teaching statistical inference without deriving the sampling distribution from the binomial. The time required to cover the underlying concepts for each topic may determine which topics can be effectively presented within a single course.

Thinking and Experience of Leading Statistical Educators

A sound curriculum should relate to the current thinking of the leaders in that field. Statistical educators agree on the importance of statistical literacy, as evidenced by the selection of this topic as a theme for the 2001 IASE Satellite Conference in Seoul and the Sixth International Conference on Teaching Statistics (ICOTS-6) in Cape Town in 2002 (See www.swin.edu.au/maths/iase/statlit.html and http://icots6.haifa.ac.il/icots6.html). There are many ideas on the definition and nature of statistical literacy—with some views focusing directly on statistical or quantitative literacy (Best, 2001; Holmes, 2003; Macnaughton, 2004; Moore, 1998a, 2001; Moreno, 1997, 2002; Steen, 2001, Scheaffer, 2001; Stroup et al., 2004).[10] But at present there is no general agreement on the relation between statistical literacy, reasoning and thinking[11] much less between statistical literacy[12], numeracy (Crockhoft, 1982) and quantitative literacy (Steen, 1989, 1990, 1997, 2001). Whatever literacy means, it seems that it should provide utility or perceived value in everyday life. Thus, statistical literacy should give students a lasting appreciation of the value of statistics in their everyday lives as decision makers and citizens. Of course, not all students may see the full utility of a course by the end of the course. But the fewer the students seeing utility by that time, the weaker the evidence for arguing that the course teaches statistical literacy.

Conclusion

Many of the key topics needed for statistical literacy are currently taught in the first two college statistics courses on inference and regression. Most students do not take the second course that contains important ideas involving confounding that are essential to evaluating the strength of evidence provided by statistics obtained in observational studies. This is the dilemma of statistical education today – the choice between statistical inference and multivariate thinking in a single course – since statistical educators are not willing to argue for two required courses (even though the MAA supports a two-semester statistics-modeling course for business majors) and do not seem ready to make serious cuts in the presentation of statistical inference.

Background and Development Process

Background of the W. M. Keck Statistical Literacy Curriculum

The statistical literacy course at Augsburg College originated in the Department of Business Administration in 1993. A second course in General Studies was initiated in 1997. In 2001, the W. M. Keck Foundation granted Augsburg College funds to develop tools and materials for teaching statistical literacy. In the W. M. Keck Statistical Literacy curriculum, statistical literacy is defined as "critical thinking about statistics used as evidence in arguments" (Schield 1999a, 2000a, 2002). Augsburg's program has strong similarities with those described by Moreno (1997, 2002 and 2004), Steen (2001), Scheaffer (2001), Moore (2001), Best (2001), Stroup et al. (2004), Hayden (2004) and Klass (2004).

In reviewing the W. M. Keck Statistical Literacy curriculum, Holmes (2003, p. 7) noted,

> The Augsburg course [in Statistical Literacy] is different. It has a different emphasis from many other courses to establish statistical literacy. It comes from a different background, but it has a lot of overlaps. And in many ways it reflects better the amount of the data that comes as part of every day life, certainly from large observational studies." He goes on to say what the Augsburg course "puts together is unique. That's not to say that the individual things are necessarily unique. But the package as a whole comes off as a very different package. It draws on ideas from areas which have not been in the traditional mainstream of statisticians. But they are there and they are statistical and we should be drawing on them." And, the Augsburg "approach to statistical literacy goes beyond numeracy by focusing on reading and communicating those topics studied in numeracy." The emphasis of this course "is much more in line with the sort of statistical literacy needed by most people in everyday life to read the news, by those who are in business commerce or management and by policy makers.

W. M. Keck Statistical Literacy Curriculum: Statistical Perspective

From a statistical perspective, the goal of Statistical Literacy is "to teach students how to read and interpret data." Dr. John Cerrito, Chair of Business Administration at Augsburg College, has argued consistently and persuasively since 1994 for this goal. To understand the influence of randomness, a sound course in statistical literacy should instruct students about the big ideas of chance, statistical inference (confidence intervals and statistical significance) and confounding (multivariate reasoning). While teaching statistical inference is critical for small-sized surveys and experiments, statisticians cannot afford to withhold the value of multivariate thinking from data consumers who must deal primarily with large scale observational studies where confounding typically has a greater effect on the statistical association than does sampling variability. The real question is how to do all of this within a single semester course where the students have limited mathematical skills and little motivation to study the subject. The subsequent sections present new tools used in the W. M. Keck Statistical Literacy curriculum to present the ideas of statistical significance and the influence of confounding within a single-semester course. To teach less is almost educational negligence given the changes in data and the uses of statistics in commerce and in social policy.

Curriculum Description

A curriculum is determined by the goals, the tools, and the time available, as well as by the motivation, background and interests of the students. The W. M. Keck Foundation has developed or tested new tools involving graphs for use in teaching statistical literacy (See Figures 4, 6 and 8-12). With these simple tools, one can focus clearly on the big ideas at hand in ways that students can readily understand. A common feature is the ability to measure the influence of a confounder so that students can work numeric problems on the influence of confounding.

Teaching Conditional Probability using Ordinary Language

Conditional probability is taught using natural language to describe ratios (percentages and rates). Table 4 illustrates a two-way half table of percentages. [10 percent of male smokers (whole) are runners (part).] These rate-style half tables convey more information in a smaller space.

Table 4
Two-way Half Table of Percentages

PERCENTAGE WHO ARE RUNNERS			
	Non-smoker	Smoker	Total
Female	50%	20%	40%
Male	25%	**10%**	20%
Total	37%	15%	30%

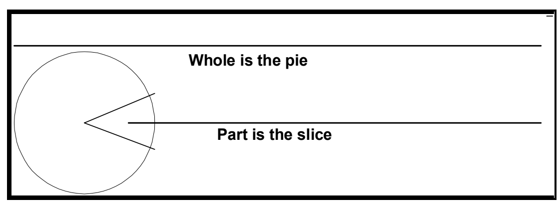

Figure 4. Part-whole Pie

Figure 4 is a device for analyzing the underlying data into the components of a part-whole ratio. Schield (2000b) indicates how to describe ratios in percent grammar and percentage grammar. Schield (2001) presents techniques needed to read tables of rates and percentages. Students find these activities surprisingly challenging and are self-motivated to master them. Since all the margin values in Table 4 are averages, both indexes are wholes and the part is in the title: runners[13]. In percent grammar, "10 percent of male smokers are runners." In percentage grammar, "The percentage of male smokers who are runners is 10 percent" or "Among male smokers, the percentage of runners is 10 percent." Studying this part-whole ratio grammar helps students realize that $P(A|B) \neq P(B|A)$. Utts (2003) identified this "confusion of the inverse" as one of seven key statistical ideas commonly misunderstood by citizens. See also Dawes (2001) and Gigerenzer (2002).

Measuring Associations using Ordinary Language Arithmetic Comparisons

Arithmetic comparisons are a powerful form of association (See Schield, 1999b, 2000b, 2001). Comparing ratios takes more factors into account. The following diagrams show two kinds of ratio comparisons. Figure 5 displays a common-part comparison of ratios: "*Girls grades 9-12 in Wyoming are 10 times as likely to chew tobacco as those in New York.*" Alternatively, Figure 6 displays a distinct-part comparison of ratios: "*Widows are more likely among suicides than are widowers*." This diagram shows that '*suicide*' is a whole – not a part.

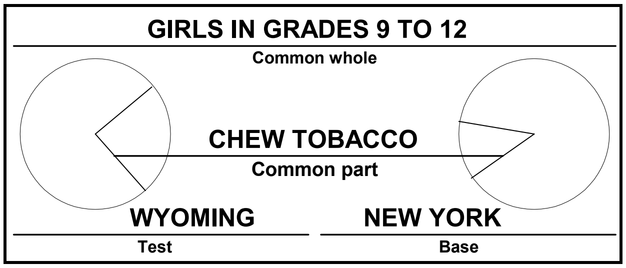

Figure 5. Common-part Comparison

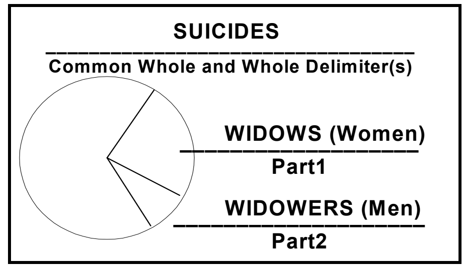

Figure 6. Distinct-parts Comparison

Diagramming Confounding

Statisticians typically focus on what is "in" the data and on the proper design of studies to ensure that plausible confounders are included. But data analysts must think "outside" the data. What unmeasured variables could influence the observed association? A commonly used diagram, Figure 7, can illustrate the associations between three variables.

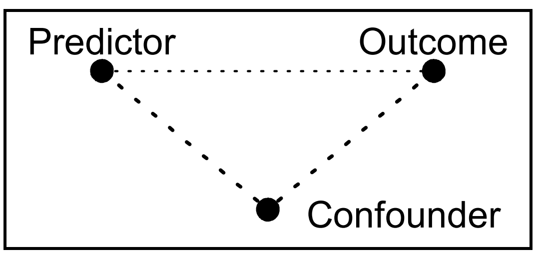

Figure 7. Three-Factor Diagram

Students generally see the importance of the relationship between the confounder and the outcome of interest but take longer to see the importance of the relation between the confounder and the predictor. They see alternate causes for an outcome but cannot see alternate explanations for an association.

Measuring the Strength of a Confounder

Schield (1999c) introduced the outcome comparison diagram in Figure 8 to help students compare the strengths of associations involving outcome, predictor and confounder that are binary. To reverse an observed association, the link between confounder and outcome must be stronger (greater difference or ratio) than that between predictor and outcome. Since the relative prevalence of death is greater for patient condition than for hospital in Figure 8, taking into account patient condition could reverse the hospital-death association.

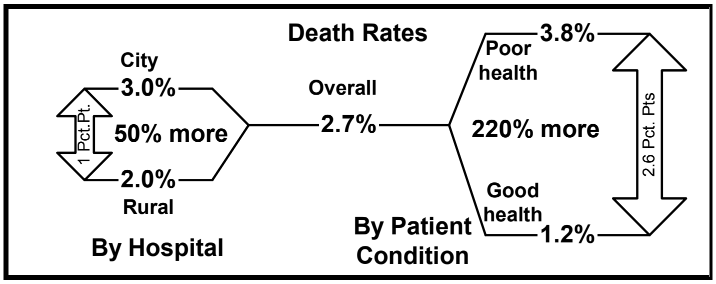

Figure 8. Outcome-Difference Diagram

Standardizing for the Influence of a Confounder

Figure 9 and Figure 10 illustrate a new graphical technique to standardize associations for the influence of a binary confounder (Schield, 2004; Wainer, 2004). This outcome-mixture graph helps students see confounder influence as a change in mix in a weighted average. The binary confounder is the horizontal axis; the outcome of interest is the vertical axis.

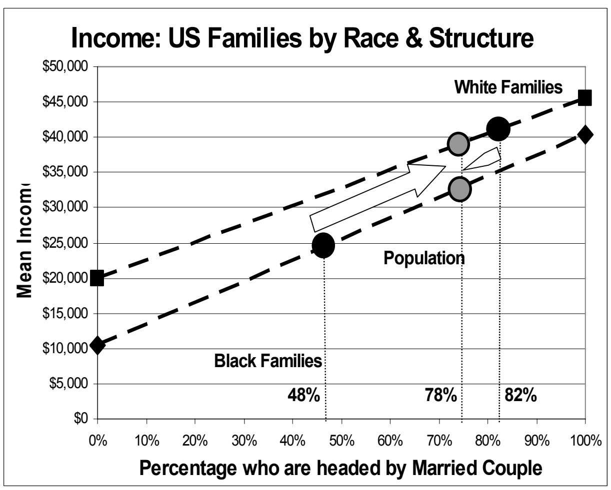

Figure 9. Family Income by Race

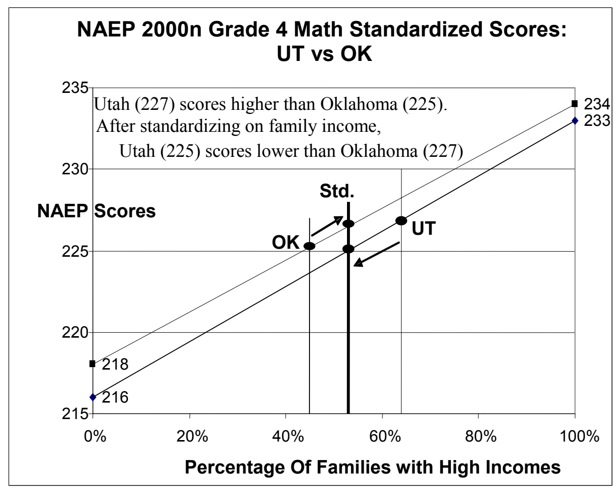

Figure 10. NAEP Scores by State

The binary predictor groups are the two weighted-average lines. The vertical lines are the confounder prevalence for the two groups. The weighted average for each predictor group depends on the confounder prevalence for that group, the vertical lines on the left and right. The vertical line in the middle is the confounder prevalence for both groups taken as together. The standardized values are obtained by using this common confounder prevalence, the vertical line in the middle. By focusing on binary predictors and confounders so that an interactive multivariate model is fully-saturated, there is no need to focus on the adequacy of the model or on model diagnostics, although one must still focus on assuming independence between the outcome and the confounder prevalence.

In Figure 10, family income is 64 percent ($16K) greater among whites ($41K) than among blacks ($25K). After standardizing on their common family structure (78 percent of all these families are headed by a married couple), family income is 18 percent ($6K) greater among whites (39K) than among blacks ($33K). Thus, 62 percent ($10K) of the original black-white family-income gap ($16K) is accounted for or explained by family structure. This graph links quantitative literacy (Steen, 2001) and statistical literacy (Moore, 2001).

Confounder-based Estimates

At this point students become aware that a confounder can either increase or decrease an observed association – and in some cases can actually reverse it (Simpson's Paradox). Using a non-interactive model, some results can be inferred. See page 5 of Schield and Burnham (2003) or Unit D5 (p. 209) in Abramson (1994) on the Exclusion Test and Direction Rule. When students are given numeric comparisons between two or three factors (Figure 11 or Figure 12), they are expected to infer the relation between the *whole effect* (57 percent higher or 2 percentage points higher) and the *direct effect* (the partial slope), the effect that would be obtained after taking into account the influence of a confounder in a non-interactive model.

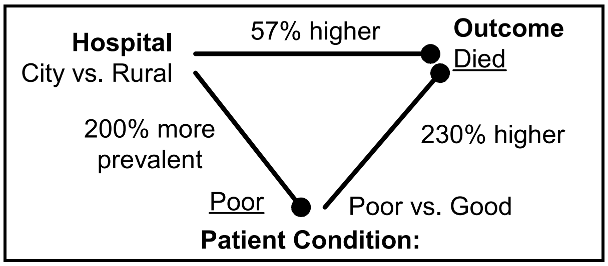

Figure 11. Percentage Difference Triangle

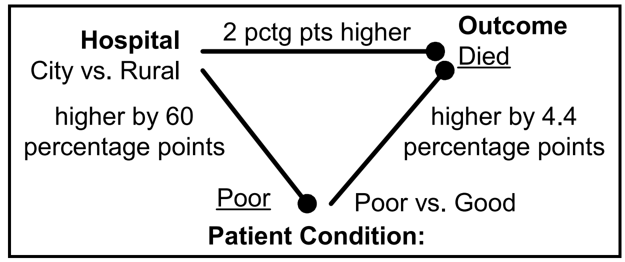

Figure 12. Simple Difference Triangle

Will the direct effect be greater, smaller or the reverse of the whole effect? Given the data in Figure 11 students should conclude that a Simpson's Paradox is possible (not impossible). Given the data in Figure 12 they should conclude that a Simpson's Paradox reversal must occur after taking patient condition into account. Being able to reach these conclusions without actually doing the multivariate regression is a useful skill in reading and interpreting data. See Schield and Burnham (2002 and 2003) for a complete discussion of these data and an indication of how "direct" and "whole" effects relate to the standardization graphs shown previously. For a discussion of confounder resistance and confounder intervals, see Schield and Burnham (2004).

Statistical Significance using Confidence Intervals

Even though statistical significance is not common in the everyday press, it is an important statistical concept. But as typically taught, it takes many class hours to cover all the underlying machinery. Giere (1996) utilized a short cut approach using just binary data (difference in proportions). This short cut bypasses the problem of distinguishing the standard deviations of the population, the sample and the sampling distribution. By using just the most conservative confidence intervals, the standard error depends only on the confidence level (Z) and the sample size (n). And by using the lack of overlap for confidence intervals as a sufficient condition for statistical significance, the teaching of statistical significance is shortened considerably.

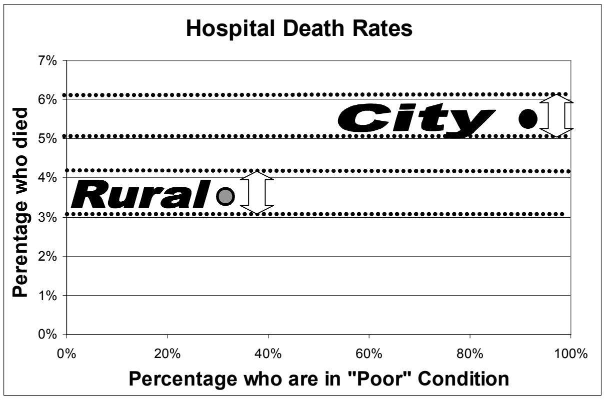

Figure 13. Statistically Significant

Given a random sample of patients from two hospitals, students calculate the 95 percent confidence intervals associated with the associated death rates as shown in Figure 13. If these intervals do not overlap, the observed difference is statistically significant at the 5 percent level. Those who see statistical significance as over-emphasized might disavow using a more valuable concept (confidence intervals) to teach a less valuable one (statistical significance). But this approach provides a quick and easy way to show the influence of confounding on statistical significance.

Vulnerability of Statistical Significance to Confounding

Figure 14 illustrates the crown jewel of a complete course in statistical literacy: showing the vulnerability of statistical significance to confounding in observational studies. (Schield, 2004) Confidence intervals are generated around the standardized value. Since these intervals overlap, this standardized difference in death rates may not be statistically significant.

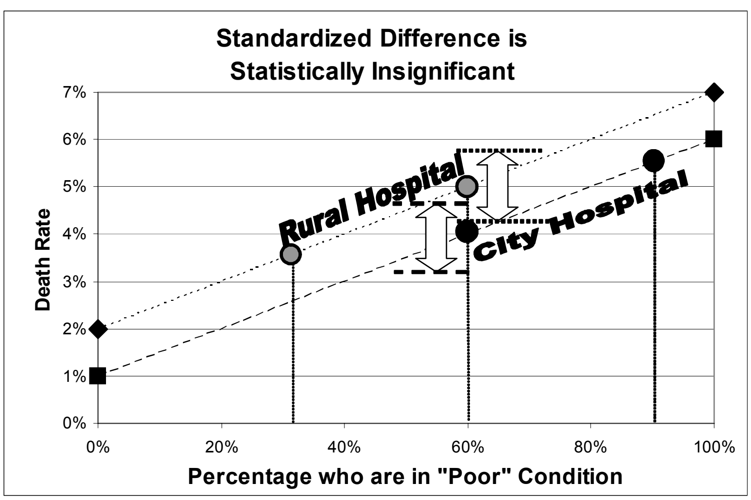

Figure 14. Statistically-Insignificant

We are unaware of any other introductory course that teaches this topic. If 517,000 U.S. college graduates study statistics each year then all too many students may be leaving with the mistaken impression that statistical significance is absolute regardless of the kind of study. Statistical educators at the IASE 2004 Roundtable support remedying this situation. After reading and discussing this paper, they were asked if students should be shown "*that statistical significance can be influenced by a confounder in all introductory statistics courses?*" In an anonymous survey, eight circled 'strongly agree', seven circled ‘generally agree’, and one circled ‘indifferent/undecided.’ The real issue is the cost. When these same people were asked, “*Should introductory statistics teach students more about confounding even if that means less time for statistical significance?*”, seven circled ‘generally agree,’ four circled ‘indifferent/undecided’ and five circled ‘disagree.’

Pilot and Implementation Results

A goal of this project is to develop teaching materials that are useful to students and usable by other teachers. Due to the continual design and infusion of new tools and topics, the teaching materials are still under development. It takes time to integrate these new techniques into a coherent curriculum and much trial and error is required. Student surveys indicate the progress. Sixty-six students taking the general studies course in statistical literacy at Augsburg College were surveyed after completing this course. The data in Table 5 indicates the percentage of students who agree strongly or agree at least moderately with the following statements.

Table 5
Percentage who agree with these statements

Agree Strongly	*Agree at least Moderately*	*Statement*
43%	81%	Developed critical thinking skills
33%	77%	Practical/relevant to major or work
33%	75%	Practical/relevant to personal/civic life
18%	57%	Should be required of all students

Note that ‘agree at least moderately’ includes ‘strongly agree.’ There is no comparable data for those taking other courses, so any conclusion is disputable. Still, these percentages are encouraging – especially since many of these students started the course saying they would not take statistical literacy unless it was required. Since respondent credibility is always an issue, the last item is perhaps the most telling. It may be easy to agree with nebulous outcomes (e.g., developed critical thinking skills). It seems harder to agree with an outcome involving very definite costs.

Outstanding Issues

In “More Damned Lies and Statistics,” Best (2004) titled his last chapter, “Toward Statistical Literacy?” He questioned whether statistical educators will take responsibility for this area. In talking about Quantitative Literacy, Scheaffer (2001, p 4) argued that statistics could and should:

> Some of the QL leaders tend to see statistics as only the small part of QL that deals with chance. I see statistics as much broader than that; in fact, I see it as encompassing much of the QL litany of topics that deal with data and its practical use in everyday situations. Thus, there is an important role for statisticians to play in this expanding interest in QL…. Statistics is the branch of the mathematical sciences that deals with numbers in context – data – and makes systematic study of how to reason under uncertainty. Statistics must be a key part of QL! … Over the years, statistics has lost out on many initiatives that should have been its province because of lack of interest or lack of foresight. Isn't it time we wholeheartedly embrace one that we can see coming?

Conclusion

Statistical educators should develop a college-level statistical literacy course for students in majors that do not require mathematics or statistics. Statistical literacy should focus on everyday arguments that use statistics as evidence. It should highlight the difference between causation and association, and between experiments and observational studies. It should uphold the power of randomized experiments and identify some specific pitfalls of observational studies in using statistical associations as evidence of causal connections. By adding a graph-and-language based course focusing on context alongside the math-based course focusing on statistical inference, statistical educators can serve a greater variety of students and in the process help them think more critically about the role of statistics as evidence in the arguments they will encounter in daily life.

Acknowledgements

To the W. M. Keck Foundation for their grant to Augsburg College "to develop statistical literacy as an interdisciplinary curriculum in the liberal arts." This grant has enabled development of many of the tools and techniques presented herein to teach statistical literacy. To my colleagues in the humanities and professions at Augsburg who encouraged development of this course for their students. To my colleague and co-author, Thomas V.V. Burnham, external reviewers, Dr. Joel Best and Peter Holmes, principal Augsburg internal reviewers, Dr. William Jasperson and Dr. Julie Naylor, project research assistant, Lena Zakharova, adjunct reviewers, Dr. John Stein, Dr. Michael O'Neal, Boyd Koehler and Marc Isaacson, and external consultants, Donald Macnaughton and Dr. Larry Copes, for their patience in reviewing the class and the materials – and for their candid evaluations, their pointed criticism, and their innovative suggestions for improvements. To the many hundreds of Augsburg students who took statistical literacy, studied these new techniques, and gave feedback on what worked and what did not. To my colleagues in statistical education, critical thinking and philosophy who guided my development. To my close colleagues in business administration at Augsburg – Dr. John Cerrito (Department Chair of Business Administration), Professor Peggy Cerrito (Director of Academic Enrichment) and Dr. Magda Paleczny-Zapp (International Business) – for their ongoing encouragement of this entrepreneurial project. And to the IASE Curriculum Design roundtable chair, Gail Burrill, and the IASE reviewers who noted problems, provided direction and gave encouragement.

Sources and References

Abramson, J. H. (1994). *Making Sense of Data.* Second Edition. NY: Oxford University Press.

Berry, D. A. (1997). Teaching elementary Bayesian statistics with real applications in science (with discussion). *The American Statistician, 51*: 241-274.

Best, J. (2001). *Damned Lies and Statistics: Untangling Numbers from the Media, Politicians, and Activists*. Berkeley and Los Angeles, CA: University of California Press.

Best, J. (2004). *More Damned Lies and Statistics: How Numbers Confuse Public Issues.* Berkeley and Los Angeles, CA: University of California Press.

Bregar, L. (2003). Teaching economic statistics in the internet era. International Association for Statistical Education (IASE) Satellite Conference on Statistics Education and the Internet in Berlin, Germany. Available online at: www.stat.auckland.ac.nz/~iase/ publications/6/Bregar.pdf

Burnham, T. (2003). Purpose and definition of statistical literacy. Spring 2003 Keck Statistical Literacy Project Review. Available online at: www.StatLit.org/pdf/2003BurnhamStatLit.pdf.

Chance, B. (2002). What can we learn? Classroom-based research in statistics. *2002 ASA Proceedings: Papers presented at the American Statistical Association Joints Statistical Meetings.* Alexandria, VA: American Statistical Association, Section on Statistical Education.

Cockroft, W. H. (Editor) (1982). *Mathematics Counts: Report of the Commission of Inquiry into the Teaching of Mathematics in Schools.* London: Her Majesty's Stationary Office.

Cryer, J. D. and Miller, R. (1994). *Statistics for Business: Data Analysis and Modeling*. Belmont, CA: Duxbury Press.

Dawes, R. (2001). *Everyday Irrationality: How Pseudo-Scientists, Lunatics, and the Rest of Us Systematically Fail to Think Rationally.* Boulder, CO, and Oxford: Westview Press.

del Mas, R. (2002a). Statistical literacy, reasoning and learning: A commentary. *Journal of Statistical Education*, *10*(2). Available online at: www.amstat.org/publications/jse/v10n3/delmas_discussion.html

del Mas, R. (2002b). Statistical Literacy, Reasoning and Learning: A Commentary. *Journal of Statistical Education*, *10*(3).

Finzer, W. & Erickson, T. (2005). Curriculum innovations using census microdata: a meetings of statistics, mathematics and social science. In G. Burrill and M. Camden (Eds.), *Curriculum Development in Statistics Education: International Association for Statistics Education 2004 Roundtable.* Voorberg, the Netherlands: International Statistics Institute.

Gal, I. (Editor) (2000). *Adult Numeracy Development: Theory, Research, Practice.* Cresskill, NJ Hampton Press.

Gal, I. (2003). Teaching for statistical literacy and services of statistics agencies. *The American Statistician*, *57*(2): 80-84.

Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistical Education*, 10(2).

Garfield, J., Hogg, B., Schau, C. & Whittinghill, D. (2002). First course in statistics: The status of educational reform efforts. *Journal of Statistical Education*, *10*(2).

Giere, R. (1996). *Understanding Scientific Reasoning.* Fourth Edition. Fort Worth, TX: Holt, Rinehart & Winston.

Gigerenzer, G. (2002). *Reckoning with risk: Learning to live with uncertainty.* London: Penguin Books.

Gnanadesikan, M., Scheaffer, R. L., Landwehr, J.M., Watkins, A.E., Barbella, P., Kepner, J., Newman, C.M., Obremski, T.E. & Swift, J. (1987). The *Quantitative Literacy Series*. Palo Alto, CA: Dale Seymour.

Harlow, L., Julaik, S. & Steiger, J. (Eds). (1997). *What If There Were No Significance Tests?* Mahwah, NJ: Erlbaum.

Hayden, R. (2004). Planning a statistical literacy course. *2004 ASA Proceedings: Papers presented at the American Statistical Association Joints Statistical Meetings.* Alexandria, VA: American Statistical Association, Section on Statistical Education.

Herrnstein, R. J. and Murray, C. (1994). *The Bell Curve: Intelligence and Class Structure in American Life.* New York: The Free Press.

Holmes, P. (1981). *Statistical needs of young non-specialists.* Nottingham, UK: Royal Statistical Society Centre for Statistical Education. Available online at: www.rsscse.org.uk/resources/publications.asp.

Holmes, P. (2003). Statistical literacy, numeracy and the future. Talk given at Augsburg College. Available online at: www.augsburg.edu/statlit/pdf/2003HolmesAugsburg.pdf

Li, J. (2005). Statistics education for junior high schools in China. In G. Burrill and M. Camden (Eds.), *Curriculum Development in Statistics Education: International Association for Statistics Education 2004 Roundtable.* Voorberg, the Netherlands: International Statistics Institute.

Klass, G. (2004). Statistics for political science majors. *2004 ASA Proceedings: Papers presented at the American Statistical Association Joints Statistical Meetings.* Alexandria, VA: American Statistical Association, Section on Statistical Education.

Lott, J.R. Jr. (1998). *More Guns; Less Crime: Understanding Crime and Gun Control Laws (Studies in Law and Economics).* University of Chicago Press.

Macnaughton, D. (2004). The introductory statistics course: The entity-property approach. Available online at: www.MatStat.com/teach.

Minkel, J., Choi, C. & Musser G. (November, 2002). Data Points Fit to Print. *Scientific American, News Scan Briefs,* p. 34.

Moore, D. (1996). Reply to Donald A. Berry on teaching elementary Bayesian statistics. *The American Statistician 51*: 275.

Moore, D. (2001). Statistical literacy and statistical competence in the new century. (Power Point Slides.) International Association for Statistical Education (IASE) Satellite Conference on Statistical Literacy in Seoul, Korea. Available online at: www.stat.auckland.ac.nz/~iase/publications/7/Moore.pdf

Moreno, J. (1997). Statistically literacy: Statistics long after school. *Proceedings of ICOTS-5.* Voorburg, the Netherlands: International Statistical Institute. Available online at: www.stat.auckland.ac.nz/~iase/publications/ 2/Topic4d.pdf

Moreno, J. (2002). Toward a statistically literate citizenry: What statistics everyone should know. *Proceedings of ICOTS-6*. Voorburg, the Netherlands: International Statistical Institute. Available online at: www.stat.auckland.ac.nz/~iase/publications/1/1b6_more.pdf

Moreno, J. (2004). Concepts and methods in one intro stats course? *2004 ASA Proceedings: Papers presented at the American Statistical Association Joints Statistical Meetings.* Alexandria, VA: American Statistical Association, Section on Statistical Education, pp. 2758-2760.

Morrison, D.E. & Henckle, R.E. (1970). *The Significance Test Controversy.* Chicago: Aldine.

Podehl, M. (2003). Statistics in the classroom: Learning to understand societal issues. IASE Satellite Conference on Statistics Education and the Internet in Berlin, Germany.

Pfannkuch, M. and Horring, J. (2005). Developing statistical thinking in a secondary girls school: A collaborative curriculum development. In G. Burrill and M. Camden (Eds.), *Curriculum Development in Statistics Education: International Association for Statistics Education 2004 Roundtable.* Voorberg, the Netherlands: International Statistics Institute.

Rumsey, D.J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistical Education*, *10*(3).

Rossman, A.J. & Short, T.H. (1995) Conditional probability and education reform: Are they compatible? *Journal of Statistical Education*, *3*(2). Available online at: http://www.amstat.org/publications/jse/v3n2/rossman.html

Scheaffer, R. L. (2001). Quantitative literacy and statistics. *Amstat News,* 293(Nov.): 3-4.

Schau, C. (1999). Survey of Attitudes Toward Statistics (SATS). Available online at: www.unm.edu/~cschau/satshomepage.htm.

Schau, C. (2003). Students attitudes: The "other" Important outcome in statistics education. *2003 ASA Proceedings: Papers presented at the American Statistical Association Joints Statistical Meetings.* Alexandria, VA: American Statistical Association, Section on Statistical Education, pp. 3673 - 3681.

Schield, M. (1999a). Statistical literacy: Thinking critically about statistics. *Of Significance, 1.* Available online at: http://web.augsburg.edu/~schield/MiloPapers/984StatisticalLiteracy6.pdf..

Schield, M. (1999b). Common errors in forming arithmetic comparisons. *Of Significance, 1*. Available online at: http://web.augsburg.edu/~schield/MiloPapers/984OfSigCompare3.pdf.

Schield, M. (1999c). Simpson's Paradox and the Cornfield Conditions. *1999 ASA Proceedings: Papers presented at the American Statistical Association Joints Statistical Meetings.* Alexandria, VA: American Statistical Association, Section on Statistical Education. Available online at: http://web.augsburg.edu/~schield/ MiloPapers/99ASA.pdf

Schield, M. (2000a). Statistical literacy and mathematical reasoning. Presented at the Ninth International Conference on Mathematics Education (ICME-9), Tokyo, Japan. Available online at: http://web.augsburg.edu/~schield/MiloPapers/2000ICME9.pdf

Schield, M. (2000b). Statistical Literacy: Difficulties in Describing and Comparing Rates and Percentages. *2000 ASA Proceedings: Papers presented at the American Statistical Association Joints Statistical Meetings.* Alexandria, VA: American Statistical Association, Section on Statistical Education. Available online at: http://web.augsburg.edu/~schield/MiloPapers/2000ASA.pdf

Schield, M. (2001). Statistical Literacy: Reading Tables of Rates and Percentages. *2001 ASA Proceedings: Papers presented at the American Statistical Association Joints Statistical Meetings.* Alexandria, VA: American Statistical Association, Section on Statistical Education. Available online at: http://web.augsburg.edu/~schield/MiloPapers/2001ASA2.pdf.

Schield, M. (2002). Three kinds of statistical literacy: What should we teach? *Proceedings of ICOTS-6.* Voorburg, the Netherlands: International Statistical Institute. Available online at: http://web.augsburg.edu/~schield/MiloPapers/ 2002ICOTS6web.pdf

Schield, M. (2004). Three graphs to promote statistical literacy. ICME-10 at Copenhagen, Denmark. Available online at: www.augsburg.edu/ppages/~schield/MiloPapers/2004SchieldICME.pdf

Schield, M. & Burnham, T. (2002). Algebraic relationships between relative risk, phi and measures of necessity and sufficiency in a 2×2 table. *2002 ASA Proceedings: Papers presented at the American Statistical Association Joints Statistical Meetings.* Alexandria, VA: American Statistical Association, Section on Statistical Education. Available online at: http://web.augsburg.edu/~schield/MiloPapers/2002JSM.pdf.

Schield, M. and Burnham, T. (2003). Confounder-induced spuriosity and reversal: Algebraic conditions using a non-interactive model for binary data. *2003 ASA Proceedings: Papers presented at the American Statistical Association Joints Statistical Meetings.* Alexandria, VA: American Statistical Association, Section on Statistical Education. Available online at: http://web.augsburg.edu/~schield/MiloPapers/2003JSM.pdf.

Schield, M. and Burnham, T. (2004). Confounder resistance and confounder intervals for a binary confounder. *2004 ASA Proceedings: Papers presented at the American Statistical Association Joints Statistical Meetings.* Alexandria, VA: American Statistical Association, Section on Statistical Education, pp. 2781-2788. Available online at: http://web.augsburg.edu/~schield/MiloPapers/2004SchieldBurnhamASA.pdf

Simon, J. (2000). *The Great Breakthrough and Its Cause.* Ann Arbor, MI: University of Michigan Press.

Sowey, E. (2003). The getting of wisdom: Educating statisticians to enhance their clients' numeracy. *The American Statistician*, *57*(2): 74-79.

Steen, L.A. (Editor) (1989). *Everybody Counts: A Report to the Nation on the Future of Mathematics Education.* Washington, DC: The National Academies Press. Available online at: www.nap.edu/books/0309039770/html/.

Steen, L.A. (Editor) (1990). *On the Shoulder's of Giants: New Approaches to Numeracy.* Washington, DC: The National Academies Press. Available online at: http://books.nap.edu/books/0309042348/html/index.html

Steen, L.A. (editor) (1997). *Why Numbers Count: Quantitative Literacy for Tomorrow's America.* New York: College Entrance Examination Board.

Steen, L.A. (Editor) (2001). *Mathematics and Democracy: The Case for Quantitative Literacy.* Princeton, NJ: National Council on Education and the Disciplines.

Stroup, D. and Jordan, E.W. (1982). The organizational context of statistics. Graduate School of Business, University of Texas at Austin, *Working Paper Series*, *81/82*: 3-7.

Stroup, D. and Jordan, E.W. (1983). Statistics: Monster in the university. *1983 A ASA Proceedings: Papers presented at the American Statistical Association Joints Statistical Meetings.* Alexandria, VA: American Statistical Association, Section on Statistical Education, pp. 135-138.

Stroup, D., Goodman, R., Cordell, R., & Scheaffer, R. (2004). Teaching statistical principles using epidemiology: Measuring the health of populations. *American Statistician, 58*(1): 77-84.

Taylor, J. (1999). Evaluation Report of the Senior Student Survey on PAC [Perception, Analysis, and Communication] Courses. A presentation to the Pomona College Curriculum Committee. Available online at: http://72.14.203.104/u/pomonaedu?q=cache:1T3zxtgs6OcJ:pom-oit-nt6.pomona.edu/gened/PACEvaluation1999.htm+%22critical+thinking%22+1999&hl=en&ie=UTF-8.

Terwilliger, J. & Schield, M. (2004). *Frequency of Simpson's Paradox in NAEP Data.* Presented the Annual Meeting of the American Educational Research Association, San Diego, CA. Available online at: www.augsburg.edu/ppages/~schield/MiloPapers/2004AERA.pdf

Utts, J. (2003). What Educated Citizens Should Know About Statistics and Probability. *The American Statistician*, 57(2), 74-79.

van Groenestijn, M. (2002). *A Gateway to Numeracy: A Study of Numeracy in Adult Basic Education.* Utrecht, the Netherlands: CD β Press.

Wainer, H. (2004). Three Paradoxes. Draft submitted to the *American Statistician*.

Endnotes

[1] In a *Scientific American News Scan Brief,* Minkel, J., Choi, C. & Musser G. (2002) summarized an article from the British Medical Journal, July 13, 2002. "Researchers from the University of Bristol and the University of Bern looked at 1,193 medical journal articles and determined which ones were accompanied by press releases and subsequently picked up by two newspapers. Notably, the papers were not inclined to describe results from randomized trials, which generate the strongest kind of scientific evidence." The journals were the *British Medical Journal* and *Lancet*; the newspapers were the *Times* and the *Sun*.

[2] Attitudes involve affect, cognitive competence, value and difficulty. The survey involved 287 undergraduates in 11 sections of the introductory statistics course offered by a Math-Stats Department. Increase in standard deviations: Difficulty (46 percent), Affect (28 percent), Value (25 percent) and Cognitive (16 percent). Of the changes in the mean scores, only the 8 percent decrease in Value was statistically significant. Alternate explanations for this decrease include a fatigue effect (post scores are generally lower than pre) and a delayed recognition effect (it takes time to appreciate the value).

[3] This percentage is estimated by relating the kind of data to the major. Randomized experiments are the hallmark of modern psychology and the health sciences; observational studies are the hallmark of business and the social sciences. Assume that all students receiving Bachelor's degrees in these four disciplines study statistics. In 1995 that involved 234,000 in business, 125,000 in the social sciences, 84,000 in the health sciences, and 74,000 in Psychology per Table 287, 2001 US Statistical Abstract. If all students studying statistics are in these disciplines, then of the US students studying statistics (517,000), 70 percent are in majors that focus primarily on observational data (359,000

[4] Search 12/14/2003 within each source. Rates were calculated by dividing hits by the total number of articles in that source. The total number of articles was obtained by searching on 'the': a word common to all articles. Nature total of 100,000 articles may be a ceiling which means the prevalences shown are overstated. The 100,000 was used since the prevalence of 'rate' and 'risk' in Nature using that total are similar to those in The Economist.

[5] Teachers were asked how much they taught the topic: extensive, moderate or none. Their answers generated two scales. (a) The horizontal *frequency scale*: the percentage of all respondents who said "Moderate" or "Extensive". (b) The vertical *emphasis scale*: the percentage who said "Extensive" of those who said "Moderate" or "Extensive."

[6] GMAT questions presume one is familiar with basic statistical concepts such as " percent of", " percent change", "average", etc. (1) Sample questions in Quantitative Problem Solving: 1a Arithmetic: discount on discount (cumulative), 1b. Percents: basis points, percentage point differences, 1c. Geometry: conversion from area to perimeter, 1d. Algebra: roots of equations, 1e. Algebra: solve two simultaneous linear equations. (2) Sample questions in Quantitative Data Sufficiency: 2a. Percentages: commission and net proceeds, 2b. Percentages: salary increase; percent comparison, 2c. Arithmetic:

identifying mileage on trip, 2e. Arithmetic: identifying operator/operation, 2f. Arithmetic: prime numbers, 2g. Geometry. Some questions focus on the distinction between necessary and sufficient.

[7] Note that the new MAA two-semester Curriculum in Business Statistics focuses more on modeling and less on chance and inference. See www.maa.org/pubs/busmath.html or http://business.math.arizona.edu/MBD/mbd.html.

[8] MEMESB members Cryer and Miller (1994) wrote such a textbook but added hypothesis testing in a later edition.

[9] Bayesian thinking focuses on the 'chance' of the alternate hypothesis being true given the data. For example, as the p-value decreases, does the evidence increase for saying the null hypothesis is false and the alternate is true?

[10] For related curriculum materials, see the ASA-NCTM Quantitative Literacy series by Gnanadesikan et al. (1987).

[11] Garfield (2002), Chance (2002), Rumsey (2002), del Mas (2002a and b) and others are working at distinguishing statistical reasoning, statistical thinking and statistical literacy. The March 2003 issue of The American Statistician contained articles by Utts (2003), Gal (2003) and Sowey (2003) on this topic. Gal (2000) and van Groenestijn (2002) have reviews on adult numeracy. For more background, see websites maintained by the IASE International Statistical Literacy Project (http://course1.winona.edu/cblumberg/islplist.htm), by the Statistical Reasoning, Thinking and Learning (SRTL) project (www.stat.auckland.ac.nz/~iase) or by the W. M. Keck Statistical Literacy project (www.Augsburg.edu/StatLit). Note that this statistical literacy curriculum is proposed as a supplement or companion to – not as a replacement of – the introductory curriculum as described by Garfield et al (2002).

[12] Li (2005, p. 221) noted that statistical literacy is described by the China Ministry of Education (MOE, 2001) as follows:

1. Familiarity with using statistical thinking to deal with problems containing data.
2. Appreciating the role statistics plays in decision making by going through the process of collecting, displaying, analyzing data, and making reasonable decisions.
3. Being able to critically read data resources, data analyses, and summarized information.

[13] While some see this percentage table as being illegitimate, this rate-style table is found in government and business publications in the US, the UK and Australia.