THREE KINDS OF STATISTICAL LITERACY: WHAT SHOULD WE TEACH?

Milo Schield
Augsburg College Minneapolis, Minnesota, USA

ABSTRACT:
Statistical literacy is analyzed from three different approaches: chance-based, fallacy-based and correlation-based. The three perspectives are evaluated in relation to the needs of employees, consumers, and citizens. A list of the top 35 statistically based trade books in the US is developed and used as a standard for what materials statistically literate people should be able to understand. The utility of each perspective is evaluated by reference to the best sellers within each category. Recommendations are made for what content should be included in pre-college and college statistical literacy textbooks from each kind of statistical literacy.

STATISTICAL LITERACY

Statistical literacy is a new term and both words (*statistical* and *literacy*) are ambiguous. In today's society *literacy* means functional literacy: the ability to review, interpret, analyze and evaluate written materials (and to detect errors and flaws therein). Anyone who lacks this type of literacy is functionally illiterate as a productive worker, an informed consumer or a responsible citizen. Functional illiteracy is a modern extension of literal literacy: the ability to read and comprehend written materials.

Statistics have two functions depending on whether the data is obtained from a random sample. If a statistic *is not* from a random sample then one can *only do* descriptive statistics and statistical modeling even if the data is the entire population. If a statistic *is* from a random sample then one can *also do* inferential statistics: sampling distributions, confidence intervals and hypothesis tests.

Statistics have two different functions depending on whether the data was obtained in an experiment or an observational study. If a statistic were obtained from an experiment where subjects were randomly assigned to a treatment and a control group, then an unlikely difference in their sample statistics (a difference that was unlikely *if* due to chance) would give strong evidence that the treatment caused the difference. If the subjects were *not* randomly assigned, then an unlikely difference would *not* give strong evidence of direct causation by itself.

To highlight both these differences in *statistical* and to separate out the study of flaws and fallacies, this paper classifies *statistical literacy* as either (1) *chance-based* (random sampling), (2) *fallacy-based* or (3) *correlation-based* (non-random assignment). The *chance-based* form could also be called *inferentially-based* to emphasize the use of classical statistical inference. The *correlation-based* form could also called *observationally-based* to better identify the lack of random assignment but an observational study is often confused with using observational techniques such as a survey. The overlap in these classifications highlights their tensionl

Statistical literacy studies arguments that use statistics as evidence (see Schield, 1999b). *Chance-based* literacy studies deductive arguments predicting variability due to chance. *Correlation-based* literacy studies inductive arguments involving causal explanations or predictions. *Fallacy-based* literacy studies fallacies and weaknesses in statistical arguments.

For an overview of the relation between statistical literacy and scientific literacy, see Ido Gal's forthcoming survey article (2002). For a comparison of the terms *statistical literacy* and *numeracy*, see http://science.ntu.ac.uk/rsscse/activities/terminology.htm.

POPULAR BOOKS THAT ARE STATISTICALLY RELATED

To stay in touch with statistically related situations that are regularly encountered, a list of today's most popular, statistically related books in general circulation was generated.

(1) A list of popular books was generated objectively based on Amazon.com searches using these search terms: statistic, risk, number, chance, numeracy, quantitative literacy, and innumeracy. Books of fiction, books on sports statistics and games of chance were omitted.

(2) A second list of popular books was generated using a subjective criterion: "Does the book feature statistics or is the book's argument strongly dependent upon the use of statistics?"

(3) Highly specialized books were omitted from the two aforementioned lists. Those omitted included books on pure math (e.g., *Zero: The Biography of a Dangerous Idea*), financial statement analysis (e.g., *Barron's Finance and Investment Handbook*), statistical programs (SPSS and Minitab) and advanced statistical analysis (Biometry, Biological Sequence Analysis, Pattern Classification, etc.) along with specialized reference books (DSM-IV). Although extremely popular, books on Six Sigma were omitted because they typically focus more on managerial issues rather than on statistical process control issues.

(4) From the books remaining, the top 35 were selected based on their sales rankings at Amazon.com as of December 2001. (These sales rankings do not reflect all book sales in the US. Textbooks sales are very underrated since they are made directly to bookstores. However, Amazon.com rankings reflect popular sales, and popular sales are relevant to statistical literacy.)

(5) These books were classified as (1) *chance-based* (deductive), (2) *fallacy-based*, or (3) *correlation-based* (inductive). Group (3) was split into (3A) *persuasive* and (3B) *didactic*.

I. CHANCE-BASED STATISTICAL LITERACY

Chance-based statistical literacy is concerned primarily with variation due to chance – random variation. This kind of literacy is closely related to what others have called *statistical reasoning* (Chance, 2000) or *statistical competence* (Rumsey, 2001). Of the top 35 statistically related books, the following focus on chance. [The number is the sales ranking within these 35 books based on sales at Amazon.com. Rank 1 indicates the highest sales.]

2	*Statistics for the Utterly Confused* by Lloyd Jaisingh, 2000
4	*Fooled By Randomness… in the Markets and Life* by Nassim Taleb, 2001
6	*Against the Gods: The Remarkable Story of Risk* by Peter Bernstein, 1998
7	*Elements of Statistical Learning* by Hastie, Tisbshirani and Friedman, 2001
9	*The Cartoon Guide to Statistics* by Larry Gonick and Woolcott Smith, 1994
10	*The Intelligent Asset Allocator:… Maximize Returns and Minimize Risk* by Bernstein, 2000
13	*Asset Allocation: Balancing Financial Risk* by Roger Gibson, 3rd. ed. 2000
18	*Risk Management* by Crouhy et al., 2000
23	*Chance in the House of Fate: A Natural History of Heredity*, Jennifer Ackerman, 2001
24	*Value at Risk: New Benchmark for Managing Financial Risk* by Philippe Jourion, 2000
30	*How to Think About Statistics*, 6th ed. by Harold Phillips, 1999

Although many of these chance-based books involve stocks/finance/risk or biology/genetics/heredity, none focus on general business. Why not? One reason may be the relative unimportance of chance-based issues in general business.

Consider the results of a 1985 survey conducted by Peter Holmes on the "Statistical Needs of Non-Specialist Young Workers." This survey "was not concerned with those working in Operational Research or Statistics departments nor with those who are employed as statistics specialists." It surveyed 155 employees ages 18-25 located at 25 businesses.

The following table shows the percentage of respondents who answered affirmatively:

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

In his summary, Peter Holmes noted "One of the surprising, and perhaps alarming, features of these results is the low position of all the tasks related to probability and probability distributions." In the table above, 17% assign probabilities to events, 13% use probability as a measure of uncertainty, 13% use statistical tests to compare sets of data (perhaps a two-t test) and 6% use a statistical test of significance.

(The low position of probability in this survey would be *alarming* if businesses were overlooking meaningful opportunities to use chance-based tools. It would *not be alarming* if chance weren't especially relevant to most business situations.)

Although the Holmes survey was published in 1985, a recent survey of 150+ Augsburg College alumni who had majored in business gave similar results. Less than 20% of those surveyed used confidence intervals or hypothesis tests in their work. [Personal communication]

Given the lack of usage of chance-based statistics in business (excluding Six Sigma) and given the lack of chance-based books of general interest, it seems that *chance-based* statistical literacy may be quite limited in what it can contribute to general statistical literacy.

II. FALLACY-BASED STATISTICAL LITERACY

Fallacy-based statistical literacy is concerned with mathematical/statistical fallacies or mathematical thinking. Of the top 35 books, the *fallacy-based* include:

1	*Damned Lies and Statistics* by Joel Best, 2000
17	*How To Lie With Statistics* by Darrel Huff, 1954
22	*Innumeracy: Mathematical Illiteracy and Its Consequences* by John Paulos, 1985
28	*A Mathematician Reads the Newspaper* by John Paulos, 1999
32	*Where Mathematics Comes From* by Lakoff, et al., 2001
33	*How the Other Half Thinks: Adventures in Mathematical Reasoning,* Sherman Stein, 2001
35	*The Sum of Our Discontent: Why Numbers Make Us Irrational* by David Boyle, 2001

Huff's classic, *How To Lie With Statistics*, is the trendsetter for the genre by retaining its best-seller status despite having been published almost 50 years ago. Of the 84 stories in Huff's book, 10 involve probability or sampling, 34 involve descriptive statistics (mean, median, etc.), and 39 involve middle school math (rates and percentages). Of these 84 stories, none mention sampling distributions, confidence intervals or hypothesis tests (chance-based statistics) while only 12 mention confounding in observational data or studies (correlation-based statistics).

Given the general popularity of this genre, the analysis of statistical fallacies should be included in teaching statistical literacy. Although books on the importance of identifying motives (Joel Best) or on the subjectivity of mental constructs such as poverty or IQ (David Boyle) are valuable, they are no substitute for books showing how to evaluate arguments that use statistics as evidence.

III. CORRELATION-BASED STATISTICAL LITERACY

Correlation-based statistical literacy is concerned with analyzing arguments involving statistics as evidence where the statistics are obtained from observational studies. These books are divided into two groups: *persuasive books* and *didactic books*.

Group IIIA: *Persuasive books*: these book present arguments using data or present data.

11	*More Guns; Less Crime* by John Lott, 1998 (2nd ed., 2000)
12	*Seven Myths of Gun Control* by Richard Poe and David Horowitz, 2001
14	*It's Getting Better All the Time (100 trends)* by Stephen Moore & Julian Simon, 2000
15	*The Mismeasure of Man* by Stephen Jay Gould, 1981 (Revised, 1996)
21	*The Bell Curve* by Richard Herrnstein and Charles Murray, 1996
25	*The Statistical Abstract of the US, 2000* edited by Glenn King, 2000
26	*The Arithmetic of Life and Death* by George Shaffner, May 2001
29	*The First Measured Century* by Caplow, Hicks and Wattenberg, 2000
31	*Climbing Mount Improbable* by Richard Dawkins, 1997
34	*Race and Culture: A World View* by Thomas Sowell, 1995

All of these books involve statistical data that is based on observational studies – there are no experiments (no random assignments). Some of the books focus primarily on data for an entire population (*More Guns; Less Crime*) while others focus on large samples (e.g., the 12,000 subjects in the National Longitudinal Study of Youth – the basis for *The Bell Curve*). Although the US Statistical Abstract is not persuasive, it provides data for persuasive arguments.

Group IIIB: *Didactic books*: these books focus on the teaching of good practices in using of statistics as evidence in inductive (persuasive) arguments.

3	*Envisioning Information*, by Edward Tufte 1990
5	*The Visual Display of Quantitative Information* by Edward Tufte, 1983
8	*Visual Explanations: Images and Quantities, Evidence and Narrative* by Tufte 1997
16	*Visual & Statistical Thinking: Displays of Evidence for Decision Making* by Tufte 1997
19	*Say It with Presentations* by Gene Zelazny, 1st ed., 1999
20	*Say It with Charts: Executives Guide to Visual Communications*, Zelazny, 4th ed. 2001
27	*Causality: Models, Reasoning and Inference* by Judea Pearl, 1999

Most of the books in the teaching group focus on the visual display of statistical data and are the products of just two authors: Edward Tufte and Gene Zelazny. Only Judea Pearl's book *Causality* focuses on non-visual arguments involving causality.

RECOMMENDATIONS FOR PRE-COLLEGE STATISTICAL LITERACY

There are three topics in the *correlation-based persuasive books* that should be included in a secondary or middle school statistical literacy text. [If these topics were taught with a focus on the grammar, it might decrease the math anxiety that seems to begin with ratios and fractions.]

1. Teach students proper techniques for describing rates and percentages in English (see Schield, 2000b). For example, is "the percentage of men who are runners" the same as "the percentage of men among runners"? Is *percent* the same as *percentage*? Is "death rate of males" different from the "males' rate of death"? Since *unemployment rate* is measured in percent, is it correct usage to say *unemployment percentage*? If 20% plus 30% is 50%, then does a 20% market share in the Eastern US and a 30% market share in the Western US give a 50% market share in the entire US?

2. Teach students how to read and interpret difficult tables and graphs involving rates and percentages (see Schield, 2001). In the 1997 *US Statistical Abstract*, about 20% of the tables involve rates or percentages; college students find many of these quite difficult to read.

3. Teach students how to compare rates and percentages (see Schield 2000b). For example, is going from 5% to 10% a 5% increase? Is "twice as much as" the same as "two times more than "? Is "men are twice as likely as women among smokers" the same as "men are twice as likely as women to be smokers"? Is relative risk the same as the odds ratio?

Some have argued that teaching these topics at post-secondary levels is remedial since the mathematics involved is clearly middle school ratios and fractions. I disagree. The difficulty is not in dividing one number by another. The difficulty is in using the proper grammar to describe the correct parts of these ratios. Furthermore, teaching or reviewing these topics at the college level is certainly appropriate so long as the primary goal is to teach students about conditional probability and mathematical thinking (see Schield, 2000a).

RECOMMENDATIONS FOR COLLEGE-LEVEL STATISTICAL LITERACY

- CHANCE- BASED: The following *chance-based* topics should be included.

1. Teach students that a statistically significant association can become statistically insignificant after taking into account the influence of a related factor.

2. Teach students about the likelihood of rare events in time (100 year floods) or in space (e.g., geographic hot spots) and their counterparts (e.g., lowest rainfall, geographic "cold" spots).

3. Teach students the relation between sample size and confidence intervals for relative risks involving rare outcomes. When are sample sizes of 1,000 or 10,000 necessary?

4. Teach students how to use inferential statistics as guides to action. Consider confidence intervals. For decision-making purposes, should one treat a fixed 95% confidence interval the same way one would treat a 95% chance of drawing a ball from an urn? Schield (1997) argues they should be equivalent with respect to action. Consider hypothesis tests. For decision-making purposes, should we say, "the smaller the p-value, the stronger evidence that the alternate hypothesis is true"? Schield (1996) argues this can true from a Bayesian perspective.

- FALLACY- BASED: Generally speaking, all of the topics in the *fallacy-based statistical literacy* books should be included in a statistical literacy text both as examples of bad practice and as motivators to learn good practice.

- CORRELATION- BASED: The following *correlation-based* topics should be included.

1. Teach conditional probability using tables of rates and percentages. Use *numerator* and *denominator* (*part* and *whole*) to introduce ordered relationships (see Schield, 2000a).

2. Teach the *median overlap* as introduced in *The Bell Curve.* This is a clever method of describing the distance between two groups using percentiles instead of standard deviations.

3. Teach the use of z-scores in graphing the results of logistic regression on the same subjects with different controls – as introduced in *The Bell Curve.* By doing so, one can easily see from their slopes which control is more highly associated with the variable of interest.

4. Teach the benefit of *relative risk* as compared to *prevalence*. For example, a study of suicides in southern California found that widows (women) were a larger percentage of these suicides than were widowers (men). See Cohn (1989). One explanation is that women are less able to withstand the loss of their spouse than are men. Students should be taught that a high prevalence (more widows than widowers among those who commit suicide) is not strong evidence of a high relative risk (higher percentage of suicides among widows than widowers).

5. Teach students that associations obtained in observational studies are *always* vulnerable to confounding – to being changed after taking into account the influence of a related factor. For example, a city-research hospital may have a much higher death rate than a rural hospital. But after taking into account the difference in the health of the patients, the city research hospital may actually have a lower death rate. (The sicker the patient, the more likely they are to go to the city research hospital. The condition of the patient confounds the apparent relation between hospital and death rate.)

6. Teach students how to predict whether an association could be reversed by a confounder. This involves using the Cornfield conditions for a Simpson's paradox reversal of a spurious association (see Schield, 1999a).

7. Teach students to evaluate claims based on relative risk. Based on a relative risk of 1.19, the US Environmental Protection Agency (EPA) claimed that second-hand smoke was the cause of about 3,000 cancer deaths a year and was justified as being classified as a Class A carcinogen. A Federal judge denied the EPA's claim saying, the "EPA adjusted established procedure and scientific norms to validate the agency's public conclusion." The EPA's established procedure had been to treat a relative risk of less than 3 as being spurious. But in the case of second-hand smoke, the EPA presented a relative risk of 1.19 as being real.

8. Teach the difference between association and causation in both syntax (grammar) and in semantics (meaning). Consider these three claims:

A. People who are heavier tend to be taller [than those who are lighter.]
B. As weight increases, height increases. [Weight is positively correlated with height.]
C. If you increase your weight, you can expect to increase your height.

Note that A indicates pure association and C indicates pure manipulative causation. But what is the causal status of B? Most students view it as causal. Technically it is association.

9. Teach students to identify and evaluate cross-level fallacies. For example, in medieval Germany the higher the prevalence of Protestants in a town, the higher the suicide rate. It would be a cross-level fallacy to use this association involving towns to form an association ignoring town and conclude that Protestants were more likely to commit suicide than Catholics. (Actually, the likelihood of Catholics to commit suicide increased as their prevalence in a town decreased.)

10. Teach students to evaluate how compelling a statistic is as *evidence for an action.* Suppose that a majority of those in prison do not have a high-school degree. How strongly does the truth of that statistic support the claim that high-school students should graduate in order to decrease their risk of going to jail?

11. Teach students about the Lieberson (1985) quintet of problems: the selectivity problem due to pseudo-controls, the contamination effect of the non-treatment group by the treatment, asymmetric causation (irreversible processes), the inappropriate use of variation explained as the goal of a good explanation (high R-squared), and the invalid presumption that adding more control variables necessarily takes one closer to the true association.

CONCLUSION:

If people are to become more statistically literate, they need to be taught to use statistics as evidence in the arguments encountered in their daily life as workers, consumers and citizens.

Statistical education must give less emphasis to *chance-based literacy* (random variation) and more emphasis to *correlation-based literacy* (systematic variation). Doing so may decrease the distance between the ASA approach to quantitative literacy (see www.amstat.org/education) and the math-based initiative for quantitative literacy (see Steen, 1997 and 2001).

Teaching statistical literacy with a strong *correlation-based* focus based on everyday issues and arguments should elevate the value – and the appreciation – of statistical education in the general population.

REFERENCES

Chance, Beth (2000). *Components of Statistical Thinking and Implications for Instruction and Assessment.* Annual Meeting of the American Education Research Association.

Cohn, Victor (1989). *News and Numbers*. Iowa State Publishers.

Gal, Ido (2000). *Adult Numeracy Development: Theory, Research and Practice.* Hampton Press.

Holmes, Peter (1985). *Statistical Needs of Non-Specialist Young Workers.* RSS Centre for Statistical Education. (For a copy, see http://science.ntu.ac.uk/rsscse/files/needs.pdf.)

Lieberson, Stanley (1985). *Making It Count: The Improvement of Social Research and Theory.* University of California Press.

Rumsey, Deborah (2001). *Statistical Literacy as a Goal for Introductory Statistics Courses.* ASA Proceedings of the Section on Statistical Education (forthcoming).

Schield, Milo (1996). *Using Bayesian Reasoning in Classical Hypothesis Testing.* ASA 1996 Proceedings of the Section on Statistical Education, p. 274-279.

Schield, Milo (1997). *Interpreting Statistical Confidence*. ASA 1997 Proceedings of the Section on Statistical Education, p. 137-142.

Schield, Milo (1999a). *Simpson's Paradox and Cornfield Conditions.* ASA 1999 Proceedings of the Section on Statistical Education, p. 106-111.

Schield, Milo (1999b). *Statistical Literacy: Thinking Critically about Statistics*. Of Significance Journal. The Association of Public Data Users, Volume 1.

Schield, Milo (2000a). *Statistical Literacy and Mathematical Reasoning*. ICME-9, Tokyo.

Schield, Milo (2000b). *Statistical Literacy: Difficulties in Describing and Comparing Rates and Percentages.* ASA 2000 Proceedings of the Section on Statistical Education, p. 76-81.

Schield, Milo (2001). *Statistical Literacy: Reading Tables of Rates and Percentages.* ASA 2001 Proceedings of the Section on Statistical Education. (forthcoming)

Steen, Lynn (2001). *Mathematics and Democracy: The Case for Quantitative Literacy*. See www.woodrow.org/nced/quantitative_literacy.

Steen, Lynn (1997). *Why Numbers Count: Quantitative Literacy for Tomorrow's America*, College Entrance Examination Board.

ACKNOWLEDGMENTS:

To the W. M. Keck Foundation for their grant to promote Statistical Literacy. To Linda Schield for her editorial support and to Romney Schield and Tom Burnham for their reviews of this paper.
Dr. Schield can be reached at schield@augsburg.edu; his papers are available at www.augsburg.edu/ppages/~schield.