REVIEW OF STATISTICAL LITERACY COURSE DESIGN AT AUGSBURG: 2002-2003

Julie Naylor, Consultant Bill Jasperson, Augsburg College
2211 Riverside Drive. Minneapolis, MN 55454

Abstract: Statistical literacy at Augsburg College involves courses in Business and in General Studies that involve the study of statistics as evidence in arguments. The goals of these courses are presented by Dr. Milo Schield along with the selection and organization of related topics in the teaching materials by chapter in the appendix. During the 2002-2003 academic year, the authors observed Dr. Schield's teaching of a one-semester course in statistical literacy for Business majors and studied the teaching materials and tests being used. This paper is based on their review of the course goals, teaching materials, tests, homework, web tutorials and student evaluations. Recommendations for improvement are presented and are being incorporated into the course.

Keywords: Education

INTRODUCTION

Statistical Literacy has been taught in the department of Business Administration at Augsburg College since 1994.[1] The materials used in this course have been generated by Milo Schield. The materials used in the 2002-2003 school year are the basis of this review. As a part of the course evaluation and review, the authors attended all the classes, helped grade homework and tests, and proposed test questions and changes in the materials.

The purpose of this paper is twofold: (1) to summarize the goals of this statistical literacy text along with the associated principles and concepts in each chapter as taught in the 2002-2003 academic year and (2) to summarize the comments of the reviewers on these goals, concepts and materials for teaching these methods. Appendix A contains a summary of the course taught to Business majors in Spring 2003.

Because the students in the class were business majors, observational studies and confounding were the primary focus, rather than experiments and chance, and the expectations of this course were to help students understand how one can make strong arguments using observational data. The course was not a standard statistics course in which the students are asked to create statistics and interpret them. The course is presented largely in nonmathematical terms.

[1] In 1998, Augsburg added a statistical literacy course (GST 200) for majors in the humanities and liberal arts. In 2001, Augsburg College received a major grant from the W. M. Keck Foundation to develop teaching materials in statistical literacy that would be "useful to students and usable by other faculty."

1.1 INTRODUCTORY MATERIAL (CHAPTERS 1 AND 2)

The stated goals of Chapter 1 are (1) to distinguish association from causation, and to recognize that (2) association is not necessarily proof of direct causation, (3) association is often our only evidence of causation, and (4) association is contextual.

An important concept is the difference between association and causation and the danger possible in assigning causation to associations. The class emphasized the concept that associations may result from confounding, bias, or chance, and not only as a result of causation.

The idea of confounding is presented very thoroughly in this course. However, even though there is a more lengthy discussion of bias in the text, the in-class discussion of bias was brief and inadequate for students to develop an understanding of the topic. Chance is only introduced with regard to large data sets. The effect of chance was not discussed, except that in large data sets chance occurrences are uncommon. The role of chance in the strength of statistical arguments is disregarded.

The course and text spend very little time on collection of data and the associated problems. However, if students are expected to be able to evaluate the strength of arguments, they must be taught about the process of collecting ~~collection of the~~ reliable data. Students must be taught that arguments based on faulty data will be faulty without regard to bias, chance or confounding. The course and the text present mainly US Census data where the data sets are complete populations. The students are therefore limited in their exposure to different types of data collection. Most of the data sets students will be exposed to following this class will not be entire population surveys.

Likewise, the subject of experiments vs. observational studies is presented but not completely enough to allow students to get an understanding of the differences and their consequences in terms of association and causation.

The problems with the introductory material are present in both the lecture and the text. Specifically, at the end of the presentation of the material, students did not have a clear concept of bias, its importance, and precautions one should take to eliminate or reduce bias. Dr. Schield discusses data from "well designed observational studies" but does not indicate how to produce or recognize such a study.

1.2 COUNTS, RATES, PERCENTAGES, AND THEIR COMPARISONS (CHAPTERS 3 AND 4)

The stated goals of Chapter 3 are to distinguish numerator (part) from denominator (whole) in ratios such as rates and percentages, to learn the grammar for constructing and reading statements involving such ratios, and to read tables containing such ratios. The stated goals of Chapter 4 are to form arithmetic comparisons of counts, named ratios (such as rate, percentage and chance), and of ratios using *likely.* See Schield (2000). Dr. Schield (see appendix) believes that this requires a mastery of the grammar of comparisons for the signs that indicate the base and the kind of comparison.

In both the course and the text, considerable time is spent on interpreting data presented in tables. Many students, at the beginning of class, have a difficult time properly describing the part-whole relationships in even simple count based tables. The margin value rules presented in class and in the text seem to be understandable to students and allow them to more easily determine the proper part-whole relationships. The graphical interpretations, such as Venn Diagrams (see appendix), also seem to be effective for the students to determine the proper relationships.

The course successfully teaches students that taking the size of a group into account (rates, ratios, percentages) make for stronger arguments than just simple counts.

Time is also spent on comparing simple differences as well as differences in ratios and percentages. Emphasis is placed on understanding these concepts, representing them graphically, and learning the language used to express them. Graphical interpretations presented in the book are well thought out and useful for the comparisons as well as the ratios.

Sentence templates exhibiting multiple forms of expression are provided for students. Because students are allowed to use these templates while taking exams, students were not required to learn the structures and their differences. In all, too much time was spent on this aspect of the material.

The course and the text are lacking in ~~much~~ discussion of non-count-based data. Non-count-based data is widely presented in the literature but it doesn't fit into the part-whole classification. Ignoring non-count-based data leaves the students with an incomplete understanding of all types of tables.

Likewise, little attention is spent on presentation of data in the form of graphs and in interpreting graphs, both in class and in the material. Data is presented in graphical form frequently and this is a deficit in the course material.

1.3 CONFOUNDING (CHAPTER 5)

Confounding can occur in any set of observational data and Dr. Schield centers the course on evaluating and judging (and "taking into account") confounding. This section is the core of the current course.

In the Fall semester, Dr. Schield talked mostly about using relative risk (relative difference) to estimate which variable might be the dominant one (see example in Appendix A on hospitals). The use of multiple regression (via Minitab) was also used to explore the importance of various variables. During the second semester of observation, multiple regression was eliminated, relative risk was covered briefly, and more time was spent on the standardization process described in Appendix A.

We found that this graphical presentation is interesting, useful and effective in showing how arguments might be strengthened or weakened or even reversed due to confounding factors. The graphical representation is a good illustration of Simpson's Paradox, a concept students traditionally have difficulty understanding. During the second semester presentation, the materials were clarified so that students were able to duplicate analyses presented in class on homework assignments and on exams. While students were able to reproduce the graphical methods, it was unclear whether they understood the meaning of what they produced.

1.4 MEASUREMENTS (CHAPTER 6)

This section contains the material that is traditionally taught in the first part of beginning statistics courses. These include: types of data, frequency distributions, measures of location, central tendency, and spread, percentiles, and the empirical rule of the normal distribution.

The topics are adequately but non-uniformly covered in the text. However, because of the large amount of time spent on the grammar of tables, The topics were gone over too quickly in class for the students to come away with ~~have~~ a satisfactory understanding of the material. We are not convinced that the concept of a distribution and the relevance of the percentiles are sufficiently presented so that they are understood by the students. This is the case in both the class and in the text.

1.5 CHANCE AND PROBABILITY (CHAPTER 7)

Chance wasn't covered at all in the first term. In the second term, chance was covered to give a background for the presentation of confidence intervals and statistical significance.

Tree diagrams are used briefly to present the concept of conditional probability and independence. These two necessary concepts should be covered in association with table interpretation where the students could have a concrete example of conditionality. The probability rules weren't covered sufficiently to allow students to understand conditional probability from this direction.

1.6 BINOMIAL DISTRIBUTION, ESTIMATION, AND SIGNIFICANCE (CHAPTERS 12,13)

The binomial distribution is covered very briefly because it is applicable to estimating the 95% confidence interval for proportions. (±1/sqrt(n)). As this is the only use of the binomial distribution in this class, it would be more useful to minimize the emphasis on the binomial distribution rather than spend time on the details which will not help the students in their understanding of confidence intervals and significance.

In the class, significance of the difference between proportions is determined by whether the two confidence intervals for the proportions overlap or not. This is most likely good enough for a basic understanding of what significance means. There are limitations to this method, including the fact that students are presented only with the idea of significance, not any methods to quantify it and that they are only presented with what is analogous to a two-sided test. It is unclear if the students understood significance as presented in these terms.

2.1 COMMENTS ON THE TEXT

The text materials improved substantially between the first and second semesters we sat in on the course. The text still needs to be more balanced in terms of the depth of some of the material presented. Some topics that many would consider very important are underrepresented, while other material is covered in so much detail as to be confusing, particularly for a student's first exposure to the subject.

There are few worked examples in the text although there are a considerable number of questions/problems interspersed with the text. However, these problems deal mostly with very specific concepts. There are very few problems that are complex enough to allow students to integrate concepts in a single question or problem.

The text is presently not adequate for the use by teachers other than Dr. Schield.

2.2 CHANGE IN THINKING ABOUT STATISTICAL LITERACY

Both of us came to this class trained in traditional statistics and each of us has biases in terms of what students should come out of a one semester course knowing. Dr. Schield presents a strong argument that material he presents is as important as many of the more traditional materials. A question that each of us has asked ourselves during this review is what group of people is this material most useful for. The course we sat in on was a basic statistics course for business students and it is also presented in a general studies course. Specifically, we asked whether this course sufficient for the business majors coming out of Augsburg College.

We don't have any definitive answers. Some of the material is useful for typical business students, though we would change the mix slightly. New and non-traditional material, such as the graphical analyses were very illuminating in understanding the subtlety of how confounding variables can affect interpretation of data.

One of our major concerns is the focus placed primarily on count-based and binary data, making the material strictly confined in both the text and class. Another concern is that the large amount of new material presented in a first course doesn't allow adequate time for the more traditional concepts. These concepts are presented quickly and with inadequate foundation. Therefore, few students left the course with an understanding of these concepts, which we believe are still necessary to achieve statistical literacy.

2.3 STUDENT REACTION

The relatively small (20 students) class size allows Dr. Schield to involve the students in the material and the learning process. The students are engaged in the learning process. The students' reactions to the course were largely favorable. In reviews, most students felt that they left the class with knowledge and skills that would be of use to them in their professional and personal lives.

3.1 CONCLUSION

We see this review as a two part process. Our first goal was to decide if the materials were adequate to teach to Dr. Schield's definition of statistical literacy. The materials which Dr. Schield has developed are a measurable first step to achieve this goal. However, they are not adequate for anyone other than Dr. Schield to teach from at this time. They lack uniformity in detail and examples and exercises that test integration and understanding of the overall concepts.

Our second goal was to decide if the topics included in these materials would adequately produce what we would believe to be statistically literate students. The answer to this depends on what one defines as statistical literacy. Dr. Schield makes a convincing case for his definition, "the use of statistics for making strong arguments." He is making progress in trying to better define topics that address his definition.

We believe that there are topics that would need to be added to the material to achieve our definition of statistical literacy, including many of the topics that are traditionally thought to be necessary. We question the ability to teach, in the span of one course, Dr. Schield's methodology to a level of real understanding and cover the traditional topics.

The difficulty in teaching Dr. Schield's materials and those we believe also need to be included may lead to a failure of the materials to be used by a more general audience. It is unclear if the materials will fill a useful niche in the current education marketplace for statistics.

APPENDIX A: COURSE DESIGN

1.1 STATISTICAL LITERACY

Based on the importance of statistical literacy, the W. M. Keck Foundation made a grant to Augsburg College to develop appropriate teaching materials that would be "useful to students and usable by faculty." This is a summary of the teaching materials used at Augsburg College in spring 2003 in teaching introductory statistics (BUS 379) to Business Majors. Dr. Schield, the project director, is the author of these materials. His approach to statistical literacy is presented in Schield 1998a, 1998b, 1999 and 2000a.

1.2 GUIDING PRINCIPLES

"As evidence in arguments, statistics have three sources of weakness: chance (random variation), bias (systematic error) and confounding (influence of associated factors). (1) Chance is dominant in small, well-designed, random samples. (2) Bias is dominant in studies that are poorly designed or poorly conducted... (3) Confounding is dominant in large well-designed observational studies." Schield (2002c) brochure

Because the students were business majors, observational studies and confounding were the primary focus rather than experiments and chance.

1.3 TABLE OF CONTENTS

The following are the chapters of the statistical literacy textbook for the 2002-2003 school year.

Table 1: Table of Contents:

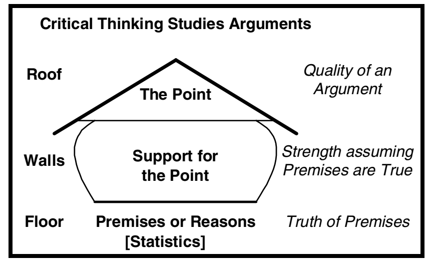
Ch	Topic	Pages
1	Stat Lit & Critical Thinking	12
2	Reasoning with Statistics	13
3	Describing Rates & Percents	66
4	Comparing Count-Based Data	72
5	Confounding & Standardizing	44
6	Interpreting Measurements	60
7	Chance and Probability	24
12	Discrete Random Variables	40
13	Estimation & Stat. Significance	21

2.1 CHAPTER 1

The primary goals of chapter 1 are (1) to distinguish association from causation, and to recognize that (2) association is not necessarily proof of direct causation, (3) association is often out only evidence of causation, and (4) association is contextual: it can become stronger or weaker after taking into account the influence of a related factor (a confounder).

Secondary goals include (1) recognizing the kinds of arguments that use statistics as evidence and (2) visualizing an argument using the house metaphor

Figure 1.

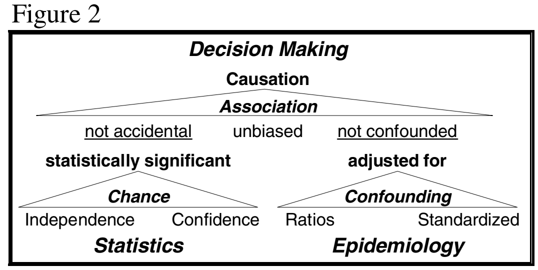


In the house metaphor, the point of the roof indicates the point of the argument; the most disputable the point, the more weight the roof needs to support. The strength of the foundation indicates the truth of the premises. The strength of the walls indicate the strength of the premises in supporting the point of the inductive argument.

2.2 CHAPTER 2

Statistical data gives stronger support in arguments when it is not accidental (statistically significant), unbiased, and not confounded.

Figure 2

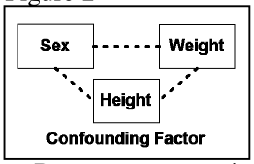


The overarching goal is for students to be able to see all three elements in essays using data to argue about causes.

- Understanding chance involves distinguishing sample from population, random from representative, and statistic from parameter, and understanding the larger the size of random sample, the smaller the margin of error due to chance.
- Understanding bias involves distinguishing researcher bias (mental) from data bias (systematic error in obtaining or presenting the data), and understanding that bias cannot be overcome by obtaining a larger sample.
- Understanding confounding involves distinguishing experiments from observational studies, controlled studies from uncontrolled studies, and *control of* from *control for.* It involves understanding how randomized experiments statistically control for confounders which may be unknown, unmeasured or unmeasurable.

The influence of a confounder in observational studies is presented graphically.

Figure 2



But arguments using statistics obtained from observational studies can be strengthened. Another goal is for students to learn the various techniques for taking into account the influence of a confounder.

Table 2

Confounding factor	Technique
Existence of confounder	Separation/matching
Size of group or factor	Use ratios: rates, %
Size of relevant comparable	Comparisons
binary confounder on ratios	Standardize ratios
mean & SD on scores	Normalize: z-scores
Binary confounder on means	Standardize means

Each technique is studied in a subsequent chapter.

2.3 CHAPTER 3

To take into account the size of the group or the size of a closely related factor, we form ratios. The goals of chapter 3 are to distinguish numerator (part) from denominator (whole) in ratios such as rates and percentages, to learn the grammar for constructing and reading statements involving such ratios, and to read tables containing such ratios. See Schield (2000 and 2001.

Table 3.

Percentage of infants with low birth weights				
MOTHERS	All	< 15	15-17	18-19
All	7.6	14.2	10.6	9.1
Smoker	12.2	15.2	12.0	10.9
Nonsmoker	6.8	14.1	10.3	8.7

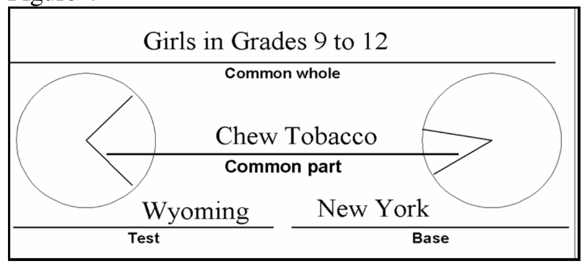
In this table, students should be able to identify what is part and whole in describing these percentages.

2.4 CHAPTER 4

To take into account the size of a relevant comparable, we form comparisons. The goals of chapter 4 are to form arithmetic comparisons of counts, named ratios (such as rate, percentage and chance), and of ratios using *likely*. See Schield (2000). This requires a mastery of the grammar of comparisons for the signs that indicate the base and the kind of comparison. Common-part comparisons must be distinguished from distinct-part comparisons. In making common-part comparisons, a goal is to fill in a graphic to distinguish test and base from common-part and common whole.

Among girls in grades 9-12, those in Wyoming are 10 times as likely to chew tobacco as those in New York.

Figure 4



2.5 CHAPTER 5

This chapter introduces some of the common uses and misuses of statistics as evidence in arguments.

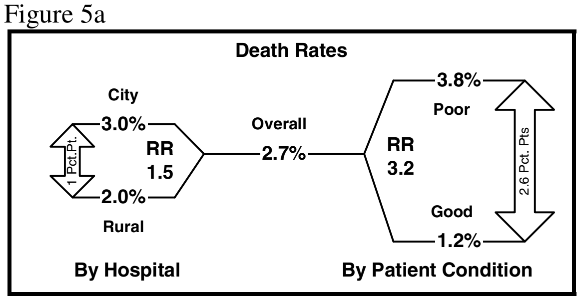
Choosing an appropriate whole is examined. Birth defects may be decreasing a percentage of all births, but can be increasing as a percentage of all infant deaths.

A Bayesian comparison is also examined. See Dawes (1999) and Schield and Burnham (2001). The ability to switch part and whole while maintaining the same numerical relationship can be quite handy.

To take into account the influence of a binary confounder on a ratio, we standardize ratios that are weighted averages if they have different mixes of a binary confounder.

A graphical technique is used to compare the association between two binary factors and a binary outcome of interest.

Figure 5a



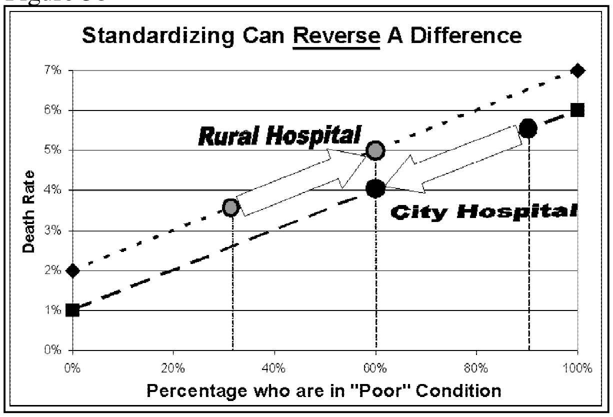
4a. In relation to the variable of interest, the binary variable that has the higher correlation typically has the greater difference and the greater relative risk. (Schield, 1999). The binary variable with the greater correlation has the greater explanatory power and is considered the more important.

4b. Consider an association after taking into account a confounder. If the association is negated, then the association is spurious. If the association is reversed, then that is an example of Simpson's paradox.

4c. An association involving two values of a binary variable can be negated or reversed *only if* a confounding factor is more important. (Schield, 1999)

The third step is to learn how to standardize two ratios so they have the same mix of a confounding factor. The graphical technique highlighted by Wainer (2002) is utilized to do this.

Figure 5b



2.6 CHAPTER 6

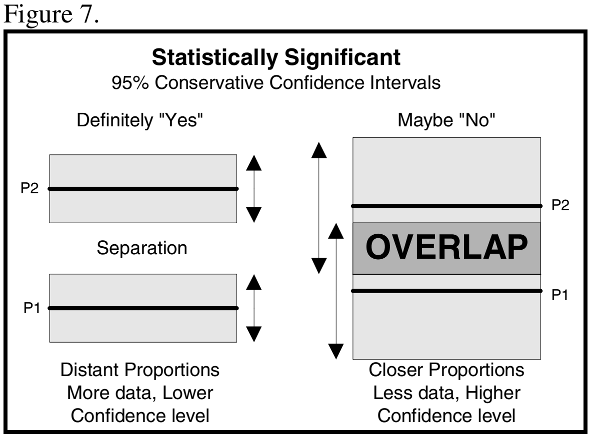
The distinction between quantitative and qualitative is reviewed (6B) along with appropriate types of graphs for each kind of data. Standard measures such as percentile (6E), mode, mean, median (6F), range and standard deviation (6K) are reviewed.

paper.

2.7 CHAPTERS 7,12,13

In the spring, statistical inference was taught using an approach used by Giere (). This approach focuses entirely on binary data (percentages). This simplifies the interpretation of the sampling distribution. This approach also uses the lack of overlap in confidence intervals to present statistical significance. Although this shortcut does not include continuous data and does not discuss hypothesis testing with Type 1 and Type 2 error, it has the advantage of being more quickly presented and more readily understood. See Figure 7.

Figure 7.



REFERENCES

Cohen, Victor (1999). *How to Help Reporters Tell the Truth.* 1999 ASA Proceedings of the Section on Statistical Education

Dawes, Robyn. *Everyday Irrationality*. Westview Press.

Giere, Ronald (). *Understanding Scientific Reasoning*.

Schield, Milo (1998). *Statistical Literacy: Thinking Critically about Statistics.* 1998 Inaugural Issue of "Of Significance" published by APDU: American Public Data Users.

Schield, Milo (1998). *Statistical Literacy and Evidential Statistics.* 1999 ASA Proceedings of the Section on Statistical Education. P. 137

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

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

Schield, Milo (2000b). *Statistical Literacy: Difficulties in Describing and Comparing Rates and Percentages.* 2000 ASA Proceedings of the Section on Statistical Education. P. 176

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

Schield, Milo and Burnham, Thomas (2002). *Algebraic Associations in 2x2 Tables.* 2002 ASA Proceedings of the Section on Statistical Education

Wainer, Howard (2002). "*The BK-Plot: Making Simpson's Paradox Clear to the Masses*" Chance Magazine Vol. 15, No. 3, Summer 2002, pp. 60–62.

Acknowledgments: This research was supported by a grant from the W. M. Keck Foundation to Augsburg College "to support the development of statistical literacy as an interdisciplinary discipline." Thanks to Donald Macnaughton for providing the reference to the article in Scientific American.

Contact: Dr. Naylor is at Jnaylor@mn.rr.com Dr. Jasperson is at jaspers@augsburg.edu Dr. Schield is at Schield@augsburg.edu.