

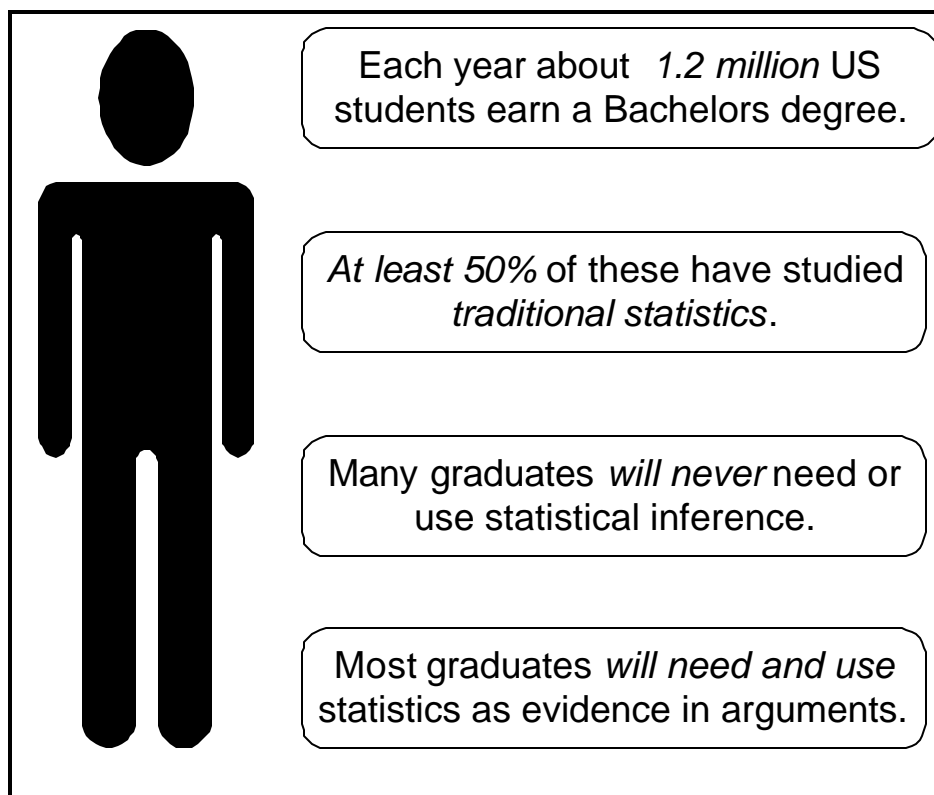
Appendix A

Statistical Literacy

A New Discipline For The 21st Century

BACKGROUND:

Q1. How big is statistics in the US?



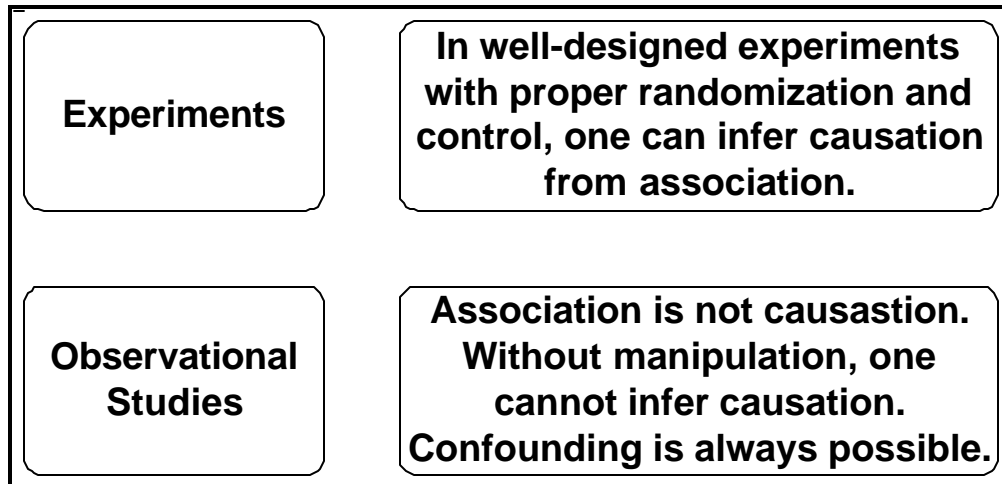
Traditional statistics is big in US colleges. More students at four-year colleges may take traditional statistics than almost any other single course outside of those in English.

Traditional statistics focuses on the influence of chance in obtaining samples from populations. This is the basis for confidence intervals and hypothesis tests. Traditional statistics focuses briefly on causality – but only that associated with experiments where random assignment is required because the subjects are heterogeneous: people, cars, seeds, etc.

But many college graduates will never conduct a survey, calculate a confidence interval or test a hypothesis. Many will never conduct an experiment involving randomization to determine causality. Even as consumers of information, they are more likely to analyze the results of an observational study than of an experiment. In large observational studies, the influence of *chance* is often of less importance in interpreting the results than the influence of *confounding*.

In traditional statistics, data obtained from observational studies is often dismissed because it is subject to confounding due to the lack of control in assigning subjects to the “treatment” or “control” groups involved.

Q2. What does traditional statistics say about “observational” causation?



To acquire knowledge, science uses the scientific method, which typically focuses on experiments. In experiments, the researcher is in control either physically (in setting values to variables) or statistically (in randomly assigning subjects to either the treatment group or to the control group). FDA clinical trials are an excellent example of using experiments to determine whether a drug is the cause of the observed effects.

Randomized trials are the ‘gold standard’ for identifying probabilistic causation because randomization breaks any links between the variables under consideration and all other confounding variables, whether these variables are known or unknown, observable or unobservable. Randomized trials are one of the greatest contributions of statistics to general knowledge.

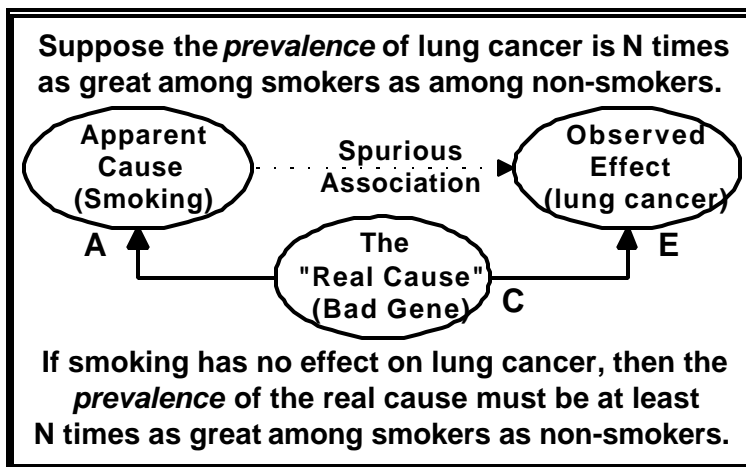
Observational studies are often dismissed by teachers using well-chosen stories such as the correlation between the increase in the number of golf courses and the number of divorces in which either might cause the other, but a common cause (increasing population) is a likely cause of both. Because there is no known test for confounders such as these, traditional statistics discourages the use of “observational data (data obtained from observational studies) to support claims about causation.

But most decisions in ones’ professional, civic or personal life involve data obtained from observational studies. If the study involves a large number of subjects, then the influence of chance is of secondary importance in comparison with the influence of confounders. Our need to make decisions under uncertainty pushes us to use observationally based associations to make judgments about causation.

Q3. How have observational statistics been useful in decision-making?

Observationally based statistics were the basis for asserting that smoking causes cancer.

In 1950, smoking was not generally accepted as a cause of cancer. Experimental data on human beings was not available; the only data was based on observational studies. One possibility was that the association between smoking and lung cancer was spurious. Perhaps the real cause was a common cause: a bad gene that increased people's desire to smoke and their risk of getting lung cancer.



While statistics as a discipline may give little support to arguments about causes based on observational studies, statistics as data can give considerable support.

There are certain mathematical necessary conditions for a confounder to generate a strong association between two variables if the observed association is totally spurious. These Cornfield conditions (Schield, 1999) were used by statisticians using data from observational studies to assert that "Smoking causes cancer" and to rebut the objections of the eminent statistician, Sir Ronald Fisher. In that case, Fisher could not provide data that was strong enough to support the claim that the association between smoking and lung cancer was spurious. On this basis, statisticians said, as statisticians, that, to the best of their knowledge, "Smoking causes cancer." Such statistics never *prove* an observed association is or is not spurious, but they can *provide evidence*.

This minimum necessary effect size might be taught in a traditional statistics course, but it requires a good understanding of conditional reasoning. Statistical educators agree that today's students have a great deal of difficulty with conditional thinking – especially when it involves probability. For example, David Moore, past President of the American Statistical Association, cites student difficulties in thinking conditionally as a principle reason for not teaching Bayesian statistics in an introductory statistics course.

Unfortunately there is no extra time in a traditional statistics course to include these topics without omitting out those topics necessary to derive or explain the role of chance in sampling from populations.

Q4. *Who needs statistical literacy and why?*

There are two groups that need statistical literacy. The first group consists of students who are going to take a traditional course in statistical inference and who need some sort of bridging course. The second group consists of students who are not required to take a traditional course in statistical inference yet who want or need to be able to use statistics as evidence in making informed decisions.

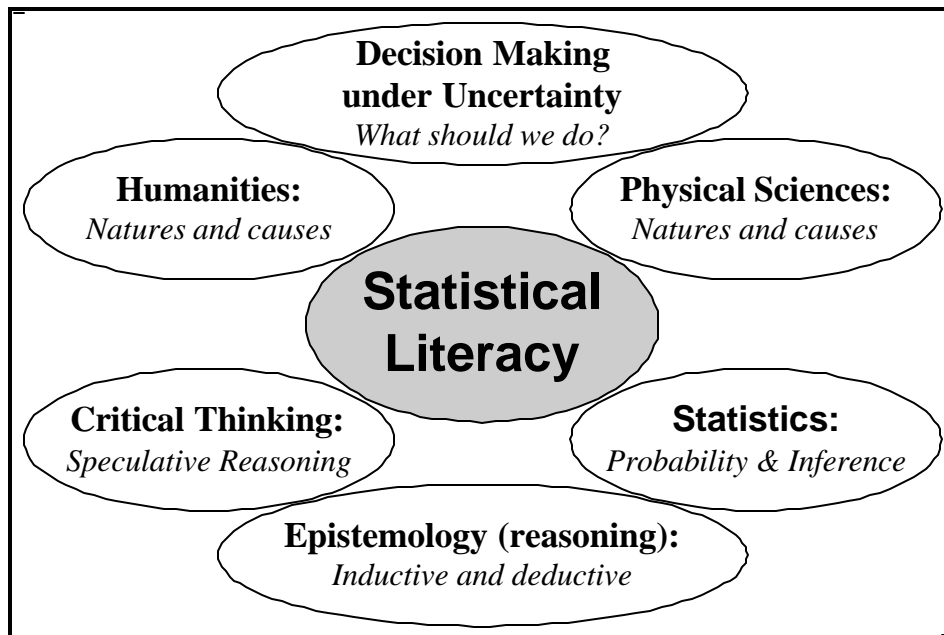
As a bridging course, statistical literacy can prepare students for traditional statistical inference by covering descriptive statistics and modeling. If all students taking traditional statistical inference were required to take statistical literacy first, then those teaching statistical inference could review descriptive statistics and quickly move on to probability thereby freeing up 10% to 20% of the class for a more leisurely pace, for additional topics or for more student activities such as team projects. This would also help teachers in teaching traditional statistical inference to get these students to take a follow-on course in statistics.

As a separate stand-alone course, statistical literacy can prepare students to be knowledgeable consumers of statistical information as used in the context of arguments. Students tend to think of statistics as being like arithmetic where $2+2$ always is 4. But in statistics, having 60% of the market in the Eastern US and 70% in the Western US does not mean having 130% of the market in the entire US. In statistics, if A has a better batting average than B in both the first half of the season and in the last half there is still no guarantee that A will have a better batting average than B has for the entire season – taken as a single unit. In statistics, associations can change magnitude and direction depending on what else one takes into account.

In general terms, anyone who must make decisions under uncertainty with statistics as evidence will have need of statistical literacy sometime. Making decisions under uncertainty should involve each of us both as consumers (in evaluating claims on health supplements, the quality of schools or the quality of hospitals) and as civic members of modern society (in evaluating the effectiveness of different political approaches to social problems).

If our society is to flourish with its members having the mental capacity to evaluate complex arguments, statistical literacy is necessary for this common good.

Q5. What is Statistical Literacy designed to do?



Statistical Literacy is designed to help students analyze arguments involving statistics as evidence concerning the natures and causes of phenomena that are studied in both the humanities and in the sciences.

Statistical Literacy integrates elements from critical thinking and statistics. These elements rest on epistemology: the branch of philosophy that studies knowledge and how we determine whether a statement is true or false. An important tool for epistemology is logic, of which there are two branches: inductive logic and deductive logic. Critical thinking and the humanities deal more with the inductive side; statistics and experimental science deal more with the deductive side. Both sides are important in decision making under uncertainty.

Statistical literacy bridges the gap between the formal sciences (logic and mathematics), the liberal arts (e.g., history, philosophy, politics, rhetoric, communications), the social sciences (e.g., sociology, psychology, political science) and the professions (e.g., education, law, business, social work, and criminal justice).

Some may think that statistical literacy is just a stripped down version of the regular statistical inference course. But statistical literacy is not just a baby-statistics course. Traditional statistics focuses almost entirely on the influence of chance in obtaining sample statistics from populations. Chance based statistics have very little to offer when the data being analyzed is the entire population (e.g., all the crimes reported by county in the US as referenced in "More Guns, Less Crime) – or when the data is a very large sample (e.g., the 12,000 subjects in the National Longitudinal Study on Youth as referenced in *The Bell Curve*).

Q6. What are some concrete examples involving statistical literacy?

#1. A hungry bear was chasing two hunters. The first hunter cried, "It's hopeless. This bear can run twice as fast as we can." The second hunter yelled back, "So what? I don't need to outrun the bear.... I just need to outrun you!"⁴ The first hunter argued that since the statistic was true (twice as fast), the conclusion was true (it's hopeless). The second hunter spotted the flaw in the argument: the *truth* of the statistic (twice as fast) is *different* from the *strength* that statistic gives in supporting the truth of a conclusion (it's hopeless). The second hunter was statistically literate: he knew that the truth of a statistic might have no relation to the strength that statistic gives in an argument.

#2. Suppose data shows that cars with phones have a higher accident rate than cars without. If you want to improve public safety, then it might seem you should support a ban on car phones given this data! Statistical literacy says, "Maybe!" If cars with car phones are driven more miles each year, are driven by younger drivers or are driven in places where accidents are more likely, then we have other plausible factors that could be causing the higher accident rate. Statistical literacy says, "To strengthen an argument for direct causation based on association, one must first take into account plausible confounders that may be more important." In this case, mileage driven, the age of the driver and the relative risk of the route driven are all plausible confounders that may be more important in predicting the accident rate than is the presence of a car phone. It may well be that after taking into account these related confounders, those cars with phones will no longer have a higher accident rate than those without. If so, then the argument in favor of banning car phones is not supported by the data.

#3. Suppose that 90% of heroin addicts first used marijuana. If we want to prevent heroin addiction, then it might seem we should criminalize marijuana. Statistical literacy says this is a weak argument: the size of a ratio, per se, gives little indication of the strength it provides in supporting a claim of causation. To see this, suppose that 99% of heroin addicts first drank milk. If the size of the ratio is to be the measure of support for causation, then we should criminalize milk before we criminalize marijuana. Statistical literacy holds that a comparison of two relevant ratios is better than the size of a single ratio in supporting causation.

#4. For districts of Germany in the 15th century, the higher the percentage of Protestants, the higher the suicide rate. It might seem this association proves that Protestants are more likely to commit suicide than Catholics. Statistical literacy would say this data provides only weak support for this conclusion. To take an association between the properties of groups of subjects and infer that same relationship holds for the properties of individual subjects is not valid. This kind of mistaken inference is known as the Ecological Fallacy. And in this particular case, the inference was false; these German Catholics were more likely to commit suicide as their prevalence in a district decreased.

⁴ Adapted from David Friedman, "Hidden Order: The Economics of Everyday Life", Harper Business.

Q7. Can statistical literacy teach mathematical thinking?

Statistical literacy can teach mathematical thinking by teaching conditional reasoning. Students have considerable difficulty with conditional reasoning. Consider these problems in traditional statistics:

In confidence intervals, many students fail to distinguish

- “the probability a random confidence interval will include the fixed population parameter”
“ $P(\mu_0 - 2\sigma \leq \tilde{x} \leq \mu_0 + 2\sigma) = .95$ ” or “ $P(\tilde{x} - 2s/\sqrt{n} \leq \mu_0 \leq \tilde{x} + 2s/\sqrt{n}) = .95$ ”
from “the probability that a [random] parameter will be in an existing confidence interval”
“ $P(\bar{x}_0 - 2s/\sqrt{n} \leq \tilde{m} \leq \bar{x}_0 + 2s/\sqrt{n}) = .95$ ”

In hypothesis testing, many students fail to distinguish

- “the probability of obtaining the sample statistic (or greater) given the null hypothesis is true”
“ $P(\tilde{x} \geq \mu_0 + 2s/\sqrt{n} \mid H_0: \mu = \mu_0) = .05$ ”
from “the probability that the null hypothesis is true given a particular sample statistic”
“ $P(H_0: \tilde{m} \leq \mu_0 \mid \bar{x}_0 = \mu_0 + 2s/\sqrt{n}) = .05$ ”
- “this outcome is unlikely *if* due to chance,” or “ $P(\text{outcome} \mid \text{chance as cause})$ ”
from “this outcome is unlikely *to be* due to chance” or “ $P(\text{chance as cause} \mid \text{outcome})$ ”
- “rejecting the null hypothesis when the null hypothesis is true”
from “finding the null hypothesis is true when the null has been rejected.”

This confusion on conditionality is a big problem. This confusion was the basis for David Moore's argument that we should not teach Bayesian statistics. This confusion was the basis for the MSMESB⁵ recommendation for a “de-emphasis of mathematical formalism (probability, hypothesis testing...).” As a result, one statistics text eliminated hypothesis testing entirely.

Statistical literacy uses conditional probability (as shown in the reading tables of rates and percentages) to educate college students about conditional reasoning. College students recognize that reading tables of rates and percentages is not rocket science but it is not all that easy either. In reading tables of rates and percentages, they often confuse or reverse part and whole. They do not realize that “the percentage of women who run” is the same as “the percentage of runners among women”. They don't realize that “the rate of death among men” is the same as “the death rate of men.” In learning these distinctions, they learn the skills necessary to handle conditional reasoning. Soon they realize that small differences in syntax can have great differences in semantics.

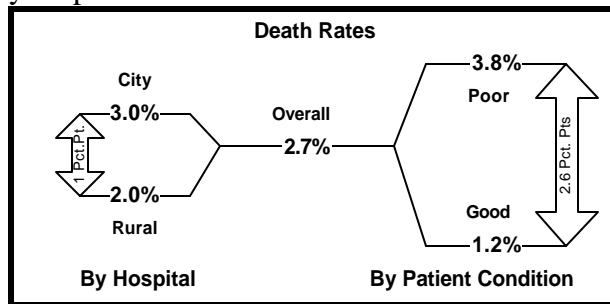
For more on this topic, see Schield (2000), *Statistical Literacy and Mathematical Thinking* and Schield (2000), *Statistical Literacy: Student Difficulties in Describing and Comparing Rates and Percentages*.

⁵ Making Statistics More Effective in Schools of Business.

Q8. What is an important principle of statistical literacy?

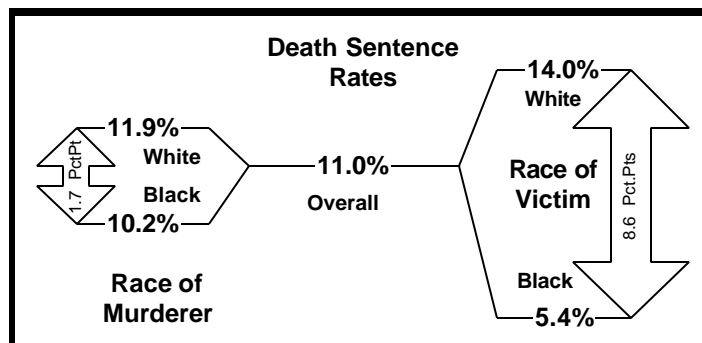
Statisticians have long been aware that a mathematical association can be reversed after taking into account a more important confounding factor. This is well known as Simpson’s Paradox. But by itself, Simpson’s Paradox promotes relativism – the possibility that anything could reverse an observed association. In statistical literacy, students learn that *only* a more important confounder can reverse an association between two factors – and they learn that the size of the difference indicates the strength of the relationship.

Suppose the death rate is higher at a city hospital (3%) than a rural hospital (2%). Does this support the claims that the City hospital is worse than the rural one?



Not necessarily. Given this data, students quickly realize that the condition of the patient (poor versus good) is more closely related to the outcome than is the location of the hospital (city versus rural). If so, then one must first take into account the condition of the patients. A failure to do so may lead to a Simpson’s Paradox reversal. But must one take into account whether a patient is left-handed or right handed? Not unless it is more closely related to the death rate than is the choice of the hospital.

Now suppose we discover that the percentage of convicted murderers who receive the death penalty is higher among whites (12%) than among blacks (10%). Can we conclude our legal system is biased against whites?



After seeing this data, students should realize that the race of the victim is more closely related to the outcome than is the race of the murderer. If so, then one must first take into account the race of the victim. Doing so in this well publicized case did result in a reversal of the association so that given the race of their victims; black murderers were more likely to receive the death penalty than were white murderers. Does this mean we need to take into account the educational background of the murderers? Not unless it is more closely related to the death sentence than is the race of the murderer.

Q9. How is statistical literacy related to quantitative literacy?

Although both a statistical literacy course and a quantitative literacy course may satisfy a skill requirement in quantitative reasoning, they are very different courses.

They may seem quite similar because both focus on numbers and on models. In quantitative literacy, there is often an entire chapter on statistics and chance. Both are mathematical and are typically taught by mathematicians.

But their differences run deep. In quantitative literacy, the result is mathematical and the process is essentially deductive so a given answer is right or wrong, true or false. In statistical literacy, the result is typically not mathematical (although the premises typically are) and the process is essentially inductive so the support one can give to the truth of an answer may be anywhere between right or wrong, true or false. By analogy, quantitative literacy deals in black or white; statistical literacy deals in shades of gray.

Statistical literacy is further from traditional statistics than quantitative literacy is from traditional mathematics.

- Quantitative literacy is an overview of topics in several different mathematics courses. In this sense, quantitative literacy is not a new discipline or a new course per se. It is simply a repackaging of selected topics from existing courses. As such, any college mathematics teacher can readily teach the topics in quantitative literacy.
- Statistical literacy includes some unique topics that are not included in traditional statistics (e.g., describing and comparing rates and percentages, the minimum effect size necessary for Simpson's Paradox) or involves a unique combination of topics from different disciplines (critical thinking and statistics). In view of these unique features, statistical literacy is a new course if not a new discipline. It is not simply a repackaging of courses from existing disciplines. But, few teachers are able to teach the topics in statistical literacy because of the close integration between critical thinking (which analyzes arguments whether inductive or deductive) and traditional statistics (which focuses entirely on arguments involving deductive inference).

Since most teachers of mathematics and statistics are unfamiliar with teaching inductive reasoning, a major goal of this project is to develop materials to help teachers make this transition.

A specific goal of this grant is the development of a Statistical Literacy textbook that is usable by students and useful to teachers from a wide variety of disciplines. It must be a text that statistics teachers find reasonably comfortable to teach or they simply will not teach this material. It must provide a comfortable introduction to the teaching of critical thinking. It must provide relevant examples, exercises and questions involving both the statistics and the critical thinking.

It may take decades for this new discipline to generate a spectrum of texts to handle this blend between traditional statistics and critical thinking, but it is hoped that the text produced under this grant will open the door to new ways of thinking on this most important subject.

Q10. How is statistical literacy related to statistical inference?

Statistical literacy can be viewed as a component of traditional statistical inference when statistical literacy focuses only on descriptive statistics and modeling. From this viewpoint, statistical literacy is just an expansion of an existing part of traditional statistics.

But statistical inference can be viewed as a component of statistical literacy when statistical literacy focuses on analyzing the role of chance in explaining an outcome.

Statistical inference studies the probability of a particular outcome if due to chance. Examples include random generation of outcomes in games of chance and random selection of samples in sampling from populations.

Statistical literacy can study outcomes to see how strongly they support the claim these outcomes are due to chance. In games of chance, statistical literacy can study how strongly the actual outcomes support the claim that the generation process was random (is the coin fair, are the dice fair?, is the dealing of the cards fair?). In sampling from a population, statistical literacy can study how strongly actual samples taken support the claim that the sampling was random.

The relationship between statistical literacy and statistical inference is similar to the relationship between Bayesian statistics and traditional frequentist statistics. Bayesian thinking is more widely accepted in the UK than in the US. For more on this relationship in relation to confidence intervals and hypothesis tests, see Schield (1996) *Using Bayesian Inference in Classical Hypothesis Testing* and Schield (1997) *Interpreting Statistical Confidence*.

David Moore, a past President of the American Statistical Association, argued that the Bayesian approach should not be taught in the introductory course. One reason for this was the difficulty college students have with conditional probability. Dr. Schield agreed with Moore in finding that students have great difficulty with conditional probability [See Schield (1998) *Using Bayesian Strength of Belief to Teach Classical Statistics* as presented at ICOTS-5, Singapore.]. But Schield has argued that we can improve students' ability to think conditionally by teaching statistical literacy [See Schield (2000), *Statistical Literacy and Mathematical Reasoning* as presented at ICME-9 in Tokyo, Japan.]

Once students can handle the subtle distinctions in conditional reasoning, then statistical inference can be included as a component of an extended course in statistical literacy. So the immediate goal of statistical literacy is to give students a solid understanding of conditional probability. Dr. Schield has proposed doing this by teaching students to describe and compare rates and percentages in tables. Once students have a better ability to reason conditionally, then statistical literacy can include both the frequentist and Bayesian views on confidence intervals and hypothesis tests.

In this proposal, the goal is generate a text that covers the foundational aspects of conditional probability and conditional reasoning – not the extended analysis of chance.

Q11. In summary, what is statistical literacy?

**Statistical Literacy
is critical thinking on
statistics as evidence**

Statistical literacy includes many basic skills such as describing and comparing rates and percentages, reading tables and graphs, and describing distributions using percentiles, means, medians and z-scores. Statistical literacy involves some advanced skills such as conditional probability and multivariate modeling using regression. In traditional statistics, the former is often short-changed and the latter is often delayed until a later course. In statistical literacy, multivariate regression is covered in the introductory course. The use of conditional associations and the ability to interpret models is an important part of evaluating arguments about causes such as those presented in “The Bell Curve” or “More Guns, Less Crime.” Arguments such as these that use social statistics to argue public policy will become even more common as more data becomes generally available.

Statistical literacy involves the art of evaluating the strength of an argument. It is not enough to do the math, read the table or model some data. One must be able to interpret the data. One must be able to interpret a statistic as evidence for an explanation or for a prediction if one is to become a discriminating consumer of quantitative data.

The critical thinking part of statistical literacy is at least as important as the statistical component. This blending of two different ‘cultures’, the formal and the informal, is what produces the tension in this discipline. This is what makes statistical literacy more difficult to teach than most other disciplines and more difficult to find teachers who can teach it. But this tension is what students need to experience in trying to integrate their experience in the sciences with the issues they confront in the humanities. Fortunately, there is a young, but vibrant, literature on critical thinking that can be used in building this discipline.

Someday, statistical literacy may be one of the basic skills that are required of all college graduates. Previously, the subject matter was not available in textbook form to accomplish this goal. Today those materials are ready in draft form, reading for testing and further development in our proposed project. Today, the timing is right for statistical literacy to become the premier new discipline of the 21st century.

APPENDIX B

**STATISTICAL LITERACY
TEXTBOOK**

STATISTICAL LITERACY TEXTBOOK

A key element in this proposal is the generation of a textbook on Statistical Literacy that is useful to students and usable by teachers. One does not just whip out a textbook on a new subject. The following pages are taken from the materials currently being used to teach Statistical Literacy (GST 200) at Augsburg College.

Chapter	CHAPTER TITLE	PAGES	TABLES	FIGURES
1	Causality and Statistics	40	1	12
2	Statistical Studies	76	1	25
3	Describing Count-Based Data	83	46	29
4	Comparing Count-Based Data	110	29	23
5	Interpreting Count-Based Data	169	52	21
6	Reading and Interpreting Measurements	125	38	76
7	Longitudinal Graphs	112	0	145
8	Cross-sectional Graphs	110	18	69
9	Linear Models – Single Factor	97	1	77
10	Linear Models – Multiple Factors	83	2	26
	TOTAL	1,005	188	503

These statistics indicate that this book is well along. Some parts have been tested for over four years in the classroom; other parts are newer.

Chapters 3 and 4 are original. They contain the grammar for describing and comparing rates and percentages. Chapter 5 contains a new technique for easily comparing the effect size of a potential confounder with that of a given factor.

The following pages contain

- I. One-page table of contents for each of these chapters.
- II. Table of contents, index, list of tables and list of figures for each of these chapters.

Together these provide some evidence for the level of commitment and experience involved in preparing to achieve the goal of preparing an introductory textbook in statistical literacy that is usable by students and useful to teachers.

March 14, 2001

Appendix C

**STATISTICAL LITERACY
PUBLICATIONS
BY DR. SCHIELD**

March 14, 2001

The following are publications by Dr. Schield on Statistical Literacy.
These are presented as background material on the relevant elements of statistical literacy.

Statistical Literacy: An overview:

Statistical Literacy: Thinking Critically About Statistics 1999 APDU Of Significance
www.augsburg.edu/ppages/~schiold/984StatisticalLiteracy6.pdf

Statistical Literacy and Mathematical Thinking, 2000 ICME-9 Tokyo
www.augsburg.edu/ppages/~schiold/ICME9.pdf

Statistical Literacy and Evidential Statistics, 1998 JSM ASA
www.augsburg.edu/ppages/~schiold/98ASAFNL.pdf

Statistical Literacy and Descriptive Statistics:

Statistical Literacy: Student Difficulties in Describing and Comparing Rates and Percentages,
2000 JSM ASA
www.augsburg.edu/ppages/~schiold/JSM2000.pdf

Simpson's Paradox and Cornfield's Conditions, 1999 JSM ASA
www.augsburg.edu/ppages/~schiold/99ASA.pdf

Common Errors in Forming Arithmetic Comparisons, 1999 APDU Of Significance
www.augsburg.edu/ppages/~schiold/984OfSigCompare3.pdf

Statistical Literacy and Inferential Statistics:

Using Bayesian Strength of Belief to Teach Classical Statistics, 1998 ICOTS-5, Singapore
www.augsburg.edu/ppages/~schiold/98ICOTS5.pdf

Interpreting Statistical Confidence, 1997 JSM American Statistical Association
www.augsburg.edu/ppages/~schiold/97ASA.pdf

Integrating Bayesian Inference and Classical Hypothesis Testing, 1996 JSM ASA
www.augsburg.edu/ppages/~schiold/96ASA.pdf