W. M. Keck
Foundation

SITE VISIT
To
AUGSBURG COLLEGE

Augsburg Proposal

"Expanding Statistical Education
to include observational causation"

MAY 4, 2001

TABLE OF CONTENTS

OVERALL PROJECT

1. Dr. Milo Schield, Augsburg
2. Dr. Donald Rubin, Harvard
3. Dr. Judea Pearl, UCLA"
4. Dr. Joan Garfield, Psychology of Education [Project Assessment Evaluator]
5. Dr. Julie Naylor, Statistics. [Assistant Project Manager]
6. Dr. Allan Rossman, Chair of the ASA Section on Statistical Education

THE STATISTICAL LITERACY COMPONENT

8. Dr. Lynn Steen, Former President, American Mathematical Association
9. Dr. Bruce Reichenbach, Author of "Introduction to Critical Thinking"
   Dr. David Kelley, Philosopher, Author of "The Art of Reasoning"
10. Wendy Treadwell, Data Analyst and past-president of APDU
11. Dr. Mark Engebretson, Physics, author of Augsburg's Vision 2004
12. Dr. David Lapakko, Speech and Communication
13. Dr. Bill Jasperson, Space Physics, finance and statistics
14. Dr. Beverly Stratton, Religion
15. John Cerrito, Department Chair, Business Administration
16. Student Comments
17. Dr. W. Phillips Shively, Political Science, University of Minnesota

ASSOCIATED MATERIALS:

ICOTS inviting Milo Schield to give a keynote talk on Statistical Literacy

Article: Need for better Statistical Literacy

Rubin's Course Description: Statistical Inference for Causal Effects

Rubin: Teaching Causal Inference in Experiments and Observational Studies

Article: The dissemination of statistical literacy among citizens and public administration directors

SCHEDULE OF ACTIVITIES

Mercedes V. Talley, Senior Program Officer, W. M. Keck Foundation, and
John Knight, Augsburg Senior Development Officer will attend all sessions.

8:45 TO 9:45 AM. **PROJECT OVERVIEW**
SPEAKERS:

1. Milo Schield, Project Overview
2. President William Frame
3. Milo Schield, Formal Causality and Assessment
4. Joan Garfield, University of Minnesota, Project Assessment Evaluator
5. (Video) Bruce Reichenbach, Philosophy. Author of *Critical Thinking*
6. Wendy Treadwell, Univ. Minn. Past President, Association of Public Data Users
7. (Video) Mark Engebretson, Physics. Co-author of Vision 2004

ATTENDANTS: Magda Zapp and Julie Naylor

10:00 TO 10:40 AM. **IMPORTANCE OF STATISTICAL LITERACY**
SPEAKERS

1. David Lapakko, Department of Communications, teaches Critical Thinking
2. Milo Schield, Department of Business Administration and Project Director
3. Beverly Stratton, Department of Religion (formerly Dept. of Mathematics)
4. William Jaspersen, Department of Space Physics.
5. Boyd Koehler, Associate Librarian; co-teaches Quantitative Journalism
6. Cass Dalglish, English and Journalism. Co-teaches Quantitative Journalism
7. Julie Naylor, Statistics, Assistant Project Director

ATTENDANTS: John Cerrito and Magda Zapp

11:00 to 11:40 AM. **TEACHING STATISTICAL LITERACY**
SPEAKERS:

1. John Cerrito, Chair of Business Administration department
2. Milo Schield, Project Director and Author of *Statistical Literacy*

ATTENDANTS: Julie Naylor, David Lapakko, Magda Zapp

12:00 to 1:00 PM. LUNCH **BIG PICTURE**
SPEAKERS:

1. Milo Schield, Talk on “Statistical Literacy Macro-Issues”
2. Phillip Shively, Professor of Political Science, University of Minnesota.
3. President William Frame

ATTENDANTS: John Cerrito, David Lapakko, Magda Zapp, Julie Naylor

Dr. Milo Schield
Department of Business Administration and MIS at Augsburg College
Project Coordinator

Thank you for selecting our proposal as one of seven finalists and for making a site visit to Augsburg College. I am extremely excited about the potential of this project to introduce fundamental curricular innovation into undergraduate education.

There is a great need among students to be better prepared to use statistics as evidence in arguments – to read and interpret data to make better decisions. Today's students are tomorrows' citizens, voters and policy makers. If they are to be empowered to deal with issues involving statistics, they must be better trained. This proposal is about empowering people so they can think more clearly and communicate more effectively about personal and social decisions.

An underlying cause is the lack of education on inferring causation from association in observational studies. Some students do not take any statistics course because of fear or a perceived lack of value. Others take traditional statistics but are told, "Association is not causation." Since association is sometimes a sign of causation, this claim should be restated, "Association is not always causation." Now the focus changes to questions. How do we know when association is causation? Are there mathematical criteria? Are there ways to tell when the evidence is stronger?

This goal of this project is to expand – if not reform – statistical education to include observational causation – the inference of causation based on observational studies. The goal is to change the way statistics is being taught at the undergraduate level.

This project offers three different remedies: two involving formal mathematical remedies (Dr. Donald Rubin, Harvard, and Dr. Judea Pearl, UCLA) and one involving Statistical Literacy (Dr. Milo Schield, Augsburg College). These remedies are not merely slogans or vague ideas. They have been documented, presented to colleagues, and tested in classrooms. While there are significant differences between the three remedies, they are complimentary and united in seeking to solve the same underlying problem.

This project has a significant focus on disseminating these ideas to statistical educators. The goal is not simply to put these good ideas into teachable materials; the goal is to get statistical educators to understand these ideas and then to use the teaching materials in their classrooms.

This proposal is bold and thus it involves risk. Indeed there are those who believe it is either unnecessary or highly overstated. But I see this disagreement as a sign of curricular innovation involving a fundamental change in content. Although we intend a revolution in statistical education, our approach is that of conversation and dialog rather than confrontation.

Today, we have an opportunity to make a real difference. But making that difference will be impossible without substantial financial support. Although the project is bold in scope, breathtaking in its assertiveness, and exciting in its potential benefits, this project is more than any one institution can fund. External funding is required if this idea is to become a reality in the near future.

Thank you for seeing merit in our first submission by inviting us to make a more detailed proposal. We are prepared to do our part in helping educate students to be better decision makers and citizens in dealing with controversial social issues involving observational statistics.

We look forward to working together in partnership with the W. M. Keck Foundation to introduce fundamental change into the undergraduate curriculum.

Dr. Donald Rubin
Professor and Chair, Department of Statistics
Harvard University

Dear Milo,

I understand that the Keck Foundation will be visiting you regarding your proposal on expanding statistical education to include the teaching causal inference based on observational studies.

From my perspective, this teaching of causal inference is the most interesting topic today in statistical education, certainly so at the undergraduate level. I am extremely enthusiastic about the entire grant.

I'm very much looking forward to my part of the project and the opportunities it offers for creating basic materials for teaching this most fundamentally important topic in statistics. At present, there is not only no generally accessible textbook on causal inference; there aren't even any generally accessible course materials! All existing materials assume a level of mathematical sophistication or philosophical abstraction that is well beyond the vast majority of entering freshman anywhere, including Harvard.

This grant should make dramatic and creative advances into teaching this topic.

It should help students develop their quantitative literacy so they are able to understand daily reports in the newspapers about health, education, policy, etc. regardless of whether the reports are based on experiments or observational studies.

I wish you every success at bringing this grant to fruition!

Best wishes, Don

Dr. Judea Pearl

Joan Garfield

University of Minnesota

My background in statistics education and assessment:
I have been working in the area of statistics education for the past 20 years. Much of my work has been in the area of student assessment. I collaborated on much of this work with Dr. Iddo Gal, formerly of the National Adult Literacy Center at the University of Pennsylvania, and currently at the University of Haifa in Israel. Together, we held an NSF-sponsored conference on the topic of assessment of statistical learning, and co-edited a book (*The Assessment Challenge in Statistics Education*), published in 1997. I have also worked with Dr. Cliff Konold, at the University of Massachusetts, Amherst, to develop instruments to assess statistical learning and reasoning, as part of his NSF-funded ChancePlus project and his current TinkerPlots project. I have just completed my own NSF grant which involved developing tools to teach and assess students' understanding about statistical inference.

My background in statistical literacy:
I have been interested in the topic of statistical literacy for the past decade, beginning with my involvement in the NSF-funded Chance Project at Dartmouth College. This project designed a quantitative literacy course for college students that introduced basic ideas of statistics and probability. We used case studies of statisticians working in different research areas, videos, small group activities, and daily discussions of statistical terms and concepts in the newspaper. I was the project evaluator and developed assessment instruments to evaluate student outcomes in a variety of different "Chance Courses" taught at Dartmouth, Princeton, Grinell, Middlebury, and Spelman College.

During the past three years I have become interested in trying to clarify and distinguish between the different learning outcomes in introductory statistics courses, and have tried to come up with definitions for three of these outcomes: statistical literacy, reasoning, and thinking. I organized a small research forum that was held in Israel in 1999 (SRTL) to gather researchers from around the world to help clarify these definitions and distinctions and to discuss what we have learned from research on these topics and what new research is needed. I am co-chairing a second research forum on this topic to be held this summer in Australia (SRTL-2). I will be reporting on this work in a keynote address at the Sixth International Conference on Teaching Statistics, to be held in South Africa in 2002, which has "Statistical Literacy" as a main theme.

My role in this project:
I am excited to work with Dr. Milo Schield on this proposed project because it gives me the opportunity to work on developing an instrument to assess statistical literacy and to use this instrument in a variety of settings to evaluate the impact of different instructional methods on the development of students' statistical literacy. This builds on some current work I am doing with Dr. Iddo Gal, which is to develop an instrument to assess adult's basic statistical literacy. As an external evaluator, I will be interested in comparing the different outcomes for the proposed courses and helping to determine which methods are most effective in achieving their stated goals.

Final comments:
Statistics educators do not yet have a consensus on a definition of statistical literacy and a model for how it should be taught. Different approaches to developing statistical literacy in a college course are currently being developed and used. However, without appropriate and consistent assessment tools, it is difficult to evaluate the strengths and weaknesses of these different approaches. The strong commitment in this project to have an external evaluation of the statistical literacy courses is an impressive component of the project and one that insures that the results will be carefully examined and utilized.

As a statistical educator, I am not tied to the success of any particular approach in this project. I hope that the W. M. Keck Foundation will fund this project so that the different approaches may be studied and evaluated, and so that the materials, models, and results may be shared with the national and international statistics education community.

Dr**. Julie Naylor,** Assistant Project Coordinator

I received my doctorate in statistics from the University of Minnesota in 1994. During my years in graduate school, I was involved in the teaching of introductory statistics, as well as doing analysis in the field of Epidemiology and other disciplines. Since that time, I have been an independent contractor doing statistical design, statistical analysis and software design. I have also been involved in educational projects, including literacy programs, science programs and school planning and evaluation on a volunteer basis.

With a background in traditional statistics, I find myself disturbed by clients', newspapers', politicians' and others' use of statistics as they pertain to observational data. In the case of observational studies, traditional statistics teaches that association cannot infer causation because there is no deductive proof that the association is not due to chance. Unfortunately, the majority of the population wants answers, and will infer causation whether or not there is any support for it.

Drs. Rubin, Pearl and Schield have developed different ways of addressing the problem of inference of causation from association. Their methods are very different, and each may have benefits and drawbacks. This project as a collaborative study will allow all of these different methods to be evaluated. I believe that it is unlikely that one method will prove to be definitive. My hope is that the strong parts of each could be incorporated into a new method.

Dr. Rubin's work is in formal inference and he has developed the Rubin Causal Model which is an extension of previous work that allows a formal framework of inferring causation beyond randomized experiments and beyond randomization-based inference. Dr. Pearl's work is also in formal inference. He has developed the assumptions and the IC algorithm for structuring candidate causal models to lead to making reliable discriminations between cause and effect.

Both Dr. Rubin's and Dr. Pearl's works require an understanding of the mathematics of traditional statistics. Determining how, and if, these methods can be taught to undergraduates is an important step in teaching the inference of causation from observational studies.

Dr. Schield's work in Statistical Literacy is in informal inference. In Statistical Literacy, students are taught methods to deduce the minimum effect size necessary for a potential confounder to explain an observed association assuming the association is totally spurious. This may lead to more questions than answers, making the students think about their conclusions.

I do not believe that the methods which Dr. Schield is proposing are new to statistics. Trying to have students understand what the statistics really mean and what inferences can be made are goals of most introductory statistics instructors. However, the message seems to get lost in for what is to many students the quagmire of mathematics. Students may have problems getting past the mathematics to what statistics mean. I would see Statistical Literacy as a course either given prior to, or instead of, introductory statistics.

Although I believe that there are some fields which already provide courses dealing with statistical literacy and "observational" causality, I believe most of these courses to be too specific to the particular field. The understanding of statistics in a variety of contexts is an extremely important skill in assessing much of what we read, hear and see about many different fields. Our decisions as to causation from association impacts what foods we might decide to eat, or not to eat, where we may decide to ask our government to spend our tax dollars and even where we decide to live.

I believe that there needs to be a change in statistical education to ensure that all students understand the statistics which inundate our everyday existence. However that will certainly take time since traditional statistics make decisions based on inferences only after much study. Therefore, I have doubts as to whether this project will culminate in any major changes in statistical education in the short term However, the willingness of this project to tackle a difficult problem from different viewpoints gives it a good chance, if the methods prove efficacious, to helping mold statistical education in the future. I believe this project to be a good first step in helping educators identify how to teach students to interpret statistics more accurately

I believe that there needs to be a change in statistical education to ensure that all students understand the statistics which inundate our everyday existence. Although I have doubts as to whether this project will culminate in any major changes in statistical education, I support this project for its willingness to tackle a difficult problem from different viewpoints. I believe this project to be a good first step in helping educators identify how to teach students to interpret statistics more accurately.

Dr. Lynn Steen
Former President of the American Mathematical Association

Dear Milo,

I'm pleased to express my support for your proposal to expand statistical education to include *statistical literacy* – the study of statistics as evidence in arguments about causality

As you know, I have been working for some time with the Woodrow Wilson Foundation (and before that with the College Board) to promote the goal of *quantitative literacy* for undergraduates as something related to, distinct from, yet for many, more important than traditional college mathematics courses such as college algebra or calculus.

I see your notion of *statistical literacy* as standing in relation to traditional statistics as *quantitative literacy* is related to mathematics: they serve different purposes, but in each case the former is typically more useful than the latter for citizens and decision-makers.

A focus on *statistical or quantitative literacy* does nothing to diminish the value of traditional statistics and mathematics for students who need these courses as preparation for future professions. But for the majority who will use their mathematical or statistical knowledge primarily as citizens, homemakers, and parents, the values of *statistical literacy* as expressed in your proposal are of unquestionable importance.

The Keck proposal offers a wonderful opportunity to generate *statistical literacy* teaching materials so it can be taught by other faculty at Augsburg and then at other colleges and universities.

I am pleased to be able to support this proposal and I look forward to following the progress of your work, not least because there will be many ways to tie it into my own work on *quantitative literacy*.

Lynn Steen, Professor of Mathematics at St Olaf College
Former President, Mathematical Association of America

Dr. Bruce Reichenbach
Professor, Department of Philosophy
Augsburg College
Author of "Introduction to Critical Thinking"

Critical thinking is a skill that is both cross-disciplinary and, in some respects, disciplinary. The cross disciplinary dimension can be taught through repeated exposure to ideas and repeated us of the skills learned. Critical thinking is not logic, although logic, both deductive and inductive, involved critical thinking. In looking at arguments, quantitative as well as qualitative elements are important to observe. Students are good with abstract numbers, but often struggle with what are traditionally termed "word problems," that is the application of mathematics to the real world. This is true generally, but specifically the case where reasoning involves numbers, as it does in generalizations and causal reasoning.

Students need to be trained to understand how numbers and statistics function, not simply abstractly, but in concrete ways in ordinary experience. This includes help students to see how statistics can strengthen or confute causal reasoning, which task Milo Schield has undertaken in his carefully developed project.

This project has the advantage of bringing together the liberal arts dimensions in the humanities, with the social science emphasis on statistical reasoning. In this way, it crosses traditional boundaries in helpful ways. This is consistent with the direction Augsburg is going, in emphasizing interdisciplinary and transformative learning.

Augsburg already has courses in traditional statistical inference and in critical thinking. This hopefully attempts to bridge the gap in ways that when publicized will have an impact on programs elsewhere.

I encourage you to carefully consider Milo Schield's project and to fund the creative potential it brings to developing Statistical Literacy.

Dr. David Kelley, Philosopher
Author of "*The Art of Reasoning*"
Guest on an ABC John Stossel video documentary

"Every day, Americans are bombarded by statistical information and claims – from newspapers and TV news, political debates, medical newsletters, stock market analysis, and many other sources. Of particular importance are inferences and arguments about causal connections based on statistical correlation. Yet relatively few people understand the logic of such reasoning, or the pitfalls and fallacies to which it is liable, or the standards that distinguish valid inferences from invalid ones. As a result, they often exaggerate risks and oversimplify complex situations, with harmful consequences for public debate and for their personal well-being.

There are few tasks in education today as urgent as improving the quality of statistical literacy.

It is not necessary that every student learn the techniques of a professional statistician, but it is important that every student know enough to become an intelligent and critical consumer of statistical information."

WENDY TREADWELL

Past President of the Association of Public Data Users

For the past eleven years I have been the coordinator of the Machine Readable Data Center (MRDC) at the University of Minnesota Libraries. In this position I have provided research and instructional support for social science data users within a University setting and provided custom services for the general public through the MRDC's affiliation with the Minnesota Data Center network. Throughout this period it became increasingly clear that the growing group of individuals who needed to use statistical information to support their research, their work, or their social cause, or simply to have a full understanding of public issues, were lacking a basic understanding of statistical information.

As a long-time member and past-President of the Association of Public Data Users, I have watched our members (most of whom are data intermediaries or producers) struggle with the problem of presenting statistics in an understandable manner to the general public. *APDU, as an advocate of broad public access to statistical data,* shares both a concern for the development of a statistically literate populace and a responsibility for helping to ensure that development. Given the promise of increased access to an ever-growing amount of federal data and increased public involvement in the discussion of federal data collection processes we take this responsibility seriously. This is in fact the reason that the premier issue of the publication *Of Significance* focused on statistical literacy.

With the increased use of the Internet and the acknowledgment of the value of information, the discussion of the need for an information literate public has grown. Schools are focusing on teaching students to evaluate the quality of the information they find as well as how to locate it. Colleges and Universities are identifying information literacy as a core competency and academic goal for their students. *Information literacy* is being discussed as a requirement for an informed and active citizenry.

One aspect of information literacy, which is not frequently discussed, is that of *statistical literacy.*

Statistical literacy is broader than specific knowledge of basic statistical methodologies and the ability to perform certain mathematical functions. For the general populace, it is the ability to recognize a 'good' number from a 'bad' one, a questionable citation from one that provides sufficient information for the reader to verify the source. It is a conceptual understanding of probability, of the meaning of statistical variance or standard errors and why they are important, and the implication of using various sampling processes on the accuracy of results.

The proposal before you addresses these issues. In my discussions with other educators, both instructors and librarians, it is clear that they are trying to find both a clear definition of the issue and a structured means of addressing it.

The availability of a well researched textbook and/or instructional package that could be used in a variety of programs and at a variety of instructional levels would be invaluable in advancing the goal of producing an informed and active citizenry.

Wendy L. Treadwell, Data Archivist
Minnesota Population Center. University of Minnesota
537 Heller Hall. 271 19th Avenue South
Minneapolis, MN 55455. wlt@pop.umn.edu

Dr. Mark Engebretson
Professor, Department of Physics

Dear Chris [Chris Kimball, Academic Dean of Augsburg College],

As you may know, Milo Schield has consulted with me from time to time as he has developed his Statistical Literacy Proposal to the Keck Foundation. I'm writing you now to give you my view of the value of this proposal and of its "fit" with my vision of Augsburg.

In a word, I think it's great!

I've had the opportunity to thoroughly review an earlier draft of his proposal, and have taken a quick look at the final, submitted version, especially with regard to the section describing the "fit" to Augsburg's vision and the guiding 2004 principles.

I believe that Milo's project (and the related activities proposed for several others of us) are excellent example of efforts to unite the liberal and the practical by means of truly innovative and significant scholarship and pedagogy. The concept of statistical literacy is also clearly interdisciplinary, crossing the boundaries of many of Augsburg's disciplines.

This project most clearly qualifies as being of great "practical" significance.

Helping students to make sound judgments on the basis of evidence and argument is a goal of most of our faculty, and is a central goal of our curricular emphasis on critical thinking. Improving the quality of decisions made in cases of inadequate or conflicting evidence is a continuing challenge, and this project addresses this concern in a new and important way.

Milo's work appears also to have significant implications in fundamental academic scholarship (the "liberal" side of the fence), in my own specialty of space physics / astronomy as well as in several other areas in the physical and social sciences. As you can well understand, astronomers can't actually perform experiments on stars; all they can do is make careful observations of many of them! The statistical techniques appropriate for controlled experiments aren't really appropriate for estimating the reliability of the inferences we develop from mostly passive observations, hence the interest of my scholarly community and several others in new means of evaluating our data.

As one of the drafters of the Augsburg 2004 document, I believe that what Milo has developed is intellectually exciting, and is very much in line with activities that the College should pursue.

A successful Keck grant proposal will bring added value to the College's academic program, to its faculty, and ultimately to the full range of its mission.

Dr. David Lapakko
Department of Speech and Communication

Professional Background: I am an Associate Professor of Speech Communications at Augsburg. As a high school and college student, I was heavily involved in contest debate. My first teaching job was a high school debate coach. At Augsburg, I have directed the forensics program for ten years. Presently, I teach some sections of Argumentation (SPC 351) within our department, as well as some sections of Critical Thinking (GST 100) in the General Studies program.

Interest in Statistical Literacy: I have nothing against algebra, geometry, calculus and other forms of higher math. But I do believe that most students have a more important and compelling need to understand numbers in more of an "applied" context. Probability and statistical literacy are at the heart of many important "real world" decisions that people must make. For example,

Are cell phones really harmful to human health?
Will carrying a concealed handgun protect me?
What effect does television have on my kids?
What are my odds of contracting the AIDS virus?
Is pornography the "true cause" of sexual assault crimes?
Are drug tests in the workplace reliable?

These and scores of other questions are grounded in statistical data, often generated by social scientists, that need to be understood and interpreted. Traditional math curricula do not typically deal with such issues. That is why a program of "statistical literacy" has a critical place in higher education – a place that I believe has been neglected and devalued over the years.

In short, I enthusiastically support the efforts of Augsburg College to help its students become more adept at seeing the importance of statistical literacy in all facets of their lives.

William H. Jasperson
B.S., M.S., Ph.D. (1972), in Meteorology (University of Wisconsin)
Graduate Minor in Statistics
M.B.A., 1988 (University of Minnesota)

Background:
I have taught courses at Augsburg from 1990 to the present in both the physics and the business departments. Most of the physics courses have been introductory science courses for non-science majors. Business courses I have taught include business statistics, finance, and computers in business.

Comments on Statistical Literacy:
In my experience with both general science and business classes, I continue to be surprised by the lack of numerical skills and the inability of many mainstream liberal arts students to interpret data, tables, and graphs in a useful way.

Statistics provides methods for presenting, describing, and summarizing data in various ways, but the language of statistics that allows interpretation of and possible inference from these data is every bit as important.

I strongly believe that there is a need for increased statistical literacy among the average college student.

Dr. Beverly Stratton
Department of Religion

Dear Professor Schield,

I am excited about your statistical literacy project and potential grant from the Keck Foundation because I think statistical literacy is important for all of our students as citizens. In particular, I'm concerned that students be able to think critically about information, claims, and arguments that include numbers.

As I think about statistical literacy, it builds on the skills of critical thinking that are part of our graduation level critical thinking skill courses and extends them to the quantitative realm. The goal of statistical literacy, however, is not learning how to construct and carry out a research design. We have research methods courses in psychology, sociology, and other disciplines that teach these important skills. Students that are interested in a general introduction to statistics also have several options, perhaps more than we need, sponsored by various departments.

What got me excited about statistical literacy was thinking about our students in Youth and Family Ministry. These are individuals, not unlike the parents and staff members of congregations they will be serving, who may or may not have statistical expertise. Still, as professionals working in family ministry, they need to know where to find credible statistical information and how use it appropriately in creating programs and advocating for children and families. They need to understand statistical information correctly and be able to explain it to others. They need to be able to ask reasonable questions about information, and to evaluate claims and arguments that use statistics. These people need to know what "statistically significant" means, but they do not need to know how to conduct T-tests or how to design a research experiment with statistically significant results. This kind of statistical literacy, particularly in relation to families, should be important to all of our students as citizens and members of families, since we all will feel the impact of social policy decisions that affect families.

As you may know, in the early 1990s I was chair of the skills subcommittee that developed the criteria for skill component courses when our current general education program was put in place. I have also taught in all of the skill areas except quantitative reasoning, which is ironic since I was a mathematics professor at the time. I think it will be important to develop a statistical literacy textbook that can be used by non-experts like me who know just enough about statistics to know that there is a lot we don't know. People from the humanities, as I am now, should be able to teach from a text that makes sense in terms of citizenship and that is not overly complicated in terms of mathematics or research design.

As I mentioned to you, I would be willing to try teaching the GST 200 Quantitative Reasoning course if it can have a focus on families. I'd probably want some coaching from you on the statistical literacy teaching and perhaps on where to find credible quantitative information on families. (I wonder if they teach this at seminaries. I would think it is important for pastors.) In return, you could count on me providing detailed comments on your draft textbook. Working with others, on campus and off, who are also interested in statistical literacy would be a bonus.

Best wishes in the grant proposal process, and tell me if there is anything else that I can do to support your efforts.

Sincerely,

Beverly J. Stratton,
Associate Professor of Religion

John Cerrito

STUDENT COMMENTS:

Emily Wood, Junior in Business Administration.
I believe that Professor Milo Schield's Statistical Literacy class (MIS 379) is a very important class for college students, especially for those going into business. This class helps tremendously in reading tables and charts so that I am able to get the correct information from them. I now realize that before this class, I was reading charts and graphs completely wrong. Now that I have gone through this class, I am confident that I can go out into the world and read any chart or graph correctly. This class is a very important class for all college students.

Monika Morris, Augsburg Marketing major
I am a non-traditional student attending Augsburg College to earn by bachelor's degree in business. My major is Marketing and my minor is in MIS. Before returning to school five years ago, I had not attended school since 1985. I begin at community college where I tested very low in math. I spent a few years taking math classes from beginning Algebra through college Algebra with some statistic classes in the process, eventually working as a math tutor at Metro Community and Technical College. I enjoy working with math and the process involved with manipulating numbers.

In the fall semester of 2000-2001, I enrolled in Professor Schield's Statistical Literacy (MIS 379) class. It was an interesting way to approach quantitative methods. Without attending a class like this before, I do not have a basis for comparison. I found the class helpful in assessing statistics and graphs. I also found the class a bit frustrating because it was more about language than actual calculations. The first obstacle I faced during the semester was mastering the language used to assess the material we went through. Ideally, statisticians would all be using the same language when making graphs and instructors could then easily teach the same language. But unfortunately, there are no hard rules for that as of yet. The second obstacle, which is always a challenge, is working from a manual that is in the process of being written. It is always hard, when the professor is still deciding on the best approach for teaching the material, and often in that situation information is changed as you are going through it. This is the second class I have attended whereby the instructors were in the process of editing their own work.

In general, I enjoyed the class and found it useful. I am currently taking Financial Management, in which I have seen some of the same concepts brought up, and my experience in Statistical Literacy (MIS 379) has helped me to understand those concepts better. I would recommend the class to other students.

Dr. W. Phillip Shively, Professor,
Department of Political Science, University of Minnesota,
Author of *Cross-Level Inference*

I have taught at the University of Minnesota for the last thirty years; before that I was a faculty member at the University of Oregon and Yale University. My research deals with elections, especially with regard to their function as vehicles for representation. I also have taught research methods frequently to our students, although we rely on the department of statistics for formal instruction in data analysis.

Almost all research in political science involves *causal analysis* – do school voucher systems *bring about* greater student success; did Hitler garner electoral support in the 1930s *because of* class divisions in Germany or because of innate anti-Semitism; can Presidents today control American foreign policy less than they could thirty years ago, and why; and so on. In common with other non-experimental fields such as sociology, economics, history, and geography, political scientists have always had to find ways to work with a body of statistical theory and practice that is deliberately not concerned with causal interpretation. We have patched together ways to use statistical techniques that are at heart innocent of causal content, to investigate our causal processes.

A particular problem for many observational social scientists is how to use data that have been grouped, in order to investigate the behavior of individuals. It turns out that *geographic averages* may give a very misleading picture of what is going on causally at the level of individuals. In the classic example, a sociologist showed fifty years ago that for states in the 1930s, % foreign born in the population was strongly and positively related to average literacy in English. If one were to use those data naively, one would conclude that foreign-born immigrants were more literate in English than native-born Americans, which was patently false. So, special techniques are needed to extract from grouped data like these our best inferences of how the variables relate causally at the level of individuals.

Political scientists and sociologists have taken the lead in developing such techniques, largely because statisticians have not seen how this problem fits into their usual pantheon of questions. Gary King at Harvard, Chris Achen at Michigan, and I at Minnesota have all worked extensively on the problem. It is not dealt with in statistics courses at Minnesota, and I doubt if it is dealt with in statistics departments elsewhere. This is <u>not</u> to say that statisticians are wrong-headed or perverse. Any discipline has to be founded on a set of agreed-upon questions and approaches. *But for that very reason, it often takes an outside shock to the system to start a discipline on the sort of self-examination that leads to real breakthroughs.*

For that reason, I strongly support Professor Schield's proposal. It aims to bring the problem of causal inference from observational studies to the attention of statistical researchers and educators *in a very positive way, by showing them how it can fit their disciplinary needs.*

This is the only way reliably to induce constructive change in a discipline.

Especially since the publication of Charles Manski's influential little book, there is plenty for us to work with in doing this. *And, Professor Schield's personal qualities of energy, conviction, and charm make it very likely that the project can succeed.*

The Sixth International Conference on Teaching Statistics

Durban, South Africa

7 - 12 July 2002
http://www.beeri.org.il/icots6/

March 12, 2001

TO: Dr. Milo Schield <schield@augsburg.edu>

Dear Dr. Schield:

On behalf of the ICOTS-6 International Program Committee (IPC), *we are very pleased to invite you to be a speaker at the Sixth International Conference on Teaching Statistics (ICOTS-6). We would like you to present a paper as part of Topic 1 (Statistics Literacy) on "**Three kinds of statistical literacy**". Your talk will be part of Cluster 1.*

However, if you would prefer to modify your topic in this session, please let us know. Specifics about Topic 1 and its clusters can be found at (http://www.beeri.org.il/icots6/) (Click on "Scientific Program", then on "Topic1- Statistics Literacy")"

You are invited to present your current work and not information that has already been published or presented at previous conferences. In sessions that deal with teaching-related topics, you should focus on key issues and evaluative information, rather than course details or syllabus descriptions (which could be provided on a handout). The number of presentations each individual may make is restricted to ensure a balanced program and enable as many people as possible to be involved.

The attached document, "ICOTS6_details" contains important information regarding dates, preparation of manuscripts, forms that you have to fill and send, and more. In the "Announcements" section of the ICOTS-6 Website (http://www.beeri.org.il/icots6/), you will find useful documents to assist you in preparing your manuscript: Guidelines for authors, pre-designed manuscript template (MS WORD), and two sample ICOTS papers.

We draw your attention to the following: (1) As previously announced, by April 1, 2001 you are asked to submit a final Title and an Abstract (up to 500 words, plain text and no references, add full contact information). Please send as plain text inside an e-mail message to Iddo Gal at the address below. (2) We anticipate each talk will be of maximum 25 minutes duration, which includes 5 minutes for questions and answers. Your actual talk should be planned for a maximum of 20 minutes. Please keep this in mind in planning the scope of your paper. (3) Please consult again the Topic 1 abstract on the conference website for the description of the specific vision for talks in your cluster.

We hope you can accept this invitation. Feel free to contact any of us if you have questions.

Yours sincerely,
Topic-1 Co-organizers:

Iddo Gal <iddo@research.haifa.ac.il>
Brian Phillips bphillips@swin.edu.au

International Statistical Institute, 1999

Need for better statistical literacy

There is a growing need for better ***statistical literacy***, seems to be a statement in discussions at the statistical world conference at Finlandia Hall on Tuesday. The ***statistical literacy*** of decision-makers and ordinary citizens alike needs improving.
Statistical literacy is addressed at the conference by Luigi Biggeri, President of the Italian Statistical Society, and Alberto Zuliani, President of the Italian National Statistical Institute (ISTAT). Although the fight against illiteracy has achieved considerable success under the leadership of UNESCO, little or no progress has been made in improving numeracy or ***statistical literacy***. According to Biggeri and Zuliani, numeracy is the weakest component in the campaign against illiteracy despite its apparent centrality in people's everyday lives.

To an ordinary citizen, ***statistical literacy*** is important in two ways. It is a necessity for understanding everyday life, a prerequisite for making rational personal decisions and a means for implementing democratic control in public administration. On the other hand, the globalisation of the economy and the advancements in technology demand from employees ever improving ***statistical literacy***, capability to read statistics and mastery of statistical reasoning. The labour markets offer numerous opportunities for those capable of statistical reasoning.

In public administration, ***statistical literacy*** is particularly necessary for those in management positions. Decentralisation of administrative decision-making and management by results require the development of efficient planning and monitoring systems supported by advanced statistical information systems.

According to Biggeri and Zuliani, ***statistical literacy*** does not mean in-depth understanding of statistics, which only a limited number of experts are capable of. Although knowledge of the basic concepts of statistics and probability are necessary for ***statistical literacy***, understanding the contents of the information is, however, the most crucial element. One must have the ability to recognise the limitations of statistical data and the accuracy with which statistical results can be evaluated.

The presenters regard it especially important for ***statistical literacy*** skills to be improved among teachers, journalists and library officials, who constitute the main intermediary spreaders of statistical information.

Source: http://www.stat.fi/isi99/press_1708.html

DRAFT DESCRIPTION
PROPOSED CORE COURSE
STATISTICAL INFERENCE FOR CAUSAL EFFECTS

Dr. Donald Rubin
Harvard University

Causal inference in statistics concerns addressing such questions as:

- Does a fatty diet increase the risk of heart disease?
- Do private schools do a better job than their public counterparts?
- Does the existence of the QRR improve the quantitative literacy of the undergraduates at Harvard? [QRR is Quantitative Reasoning Requirement]

Such questions dominate many decision-making processes but only rarely are their "answers" based on the careful collection and analysis of empirical data. This course confronts such questions and how to reach inferentially valid answers that summarize uncertainty using formal probabilistic statements.

The course is divided into three parts.

- The first part is devoted to the definition of causal effects: what are they and what kinds of assumptions are needed to make them estimable from data?
- The second part concerns the estimation of causal effects in the straightforward context of randomized experiments, like those used by FDA in the final phases of approving pharmaceuticals, e.g., to control high blood pressure.
- The third part concerns the estimation of causal effects from nonrandomized data, such as observational data on the effects of smoking on disease or of the effects of the alleged misconduct of the tobacco industry on health-care expenditures - two very different causal questions.

There will be two kinds of problem sets, and two kinds of related questions on the examinations, where the mix of the types will depend on the number and backgrounds of the students enrolled.

- The first type consists of relatively straightforward application of statistical ideas in experiments and observational studies--for example, to calculate, on a specific data set, unbiased estimates of causal effects and associated significance levels of null hypotheses and confidence intervals for average effects. The assignments from the early part of the course would use data from a randomized experiment, e.g., comparing two kinds of stents for vascular disease, whereas from the later part of the course would be from an observational study, e.g., concerning the effects of in utero exposure to barbiturates, both real examples of some importance.
- The second type of problem will be more conceptually demanding: a project to formulate the design, analysis, and evaluation of a study to estimate the effects of an intervention, such as the QRR core program itself.

I anticipate that there will be two lectures and one section per week, with a half-dozen problem sets of the first type, which would be discussed in sections, and one major problem set of the second type, which would be due near the end of the course, and both a midterm and a final examination.

TEACHING CAUSAL INFERENCE IN EXPERIMENTS AND OBSERVATIONAL STUDIES

Donald B. Rubin, Harvard University
I Oxford St., 6th fl., Cambridge, MA 02138

Key Words: potential outcomes, Rubin Causal Model (RCM), Fisherian inference, Neymanian inference, Bayesian inference, experiments, observational studies, instrumental variables, noncompliance, statistical education

Inference for causal effects is a critical activity in many branches of science and public policy. The field of statistics is the one field most suited to address such problems, whether from designed experiments or observational studies. Consequently, it is arguably essential that departments of statistics teach courses in causal inference to both graduate and undergraduate students. This presentation will discuss some aspects of such courses based on: a graduate level course taught at Harvard for a half dozen years, sometimes jointly with the Department of Economics (with Professor Guido Imbens, now at UCLA), and current plans for an undergraduate core course at Harvard University. An expanded version of this brief document will outline the courses' contents more completely. Moreover, a textbook by Imbens and Rubin, due to appear in 2000, will cover the basic material needed in both courses.

The current course at Harvard begins with the definition of causal effects through potential outcomes. Causal estimands are comparisons of the outcomes that would have been observed under different exposures of units to treatments. This approach is commonly referred to as 'Rubin's Causal Model - RCM" (Holland, 1986), but the formal notation in the context of randomization-based inference in randomized experiments goes back to Neyman (1923), and the intuitive idea goes back centuries in various literatures; see also Fisher (1918), Tinbergen (1930) and Haavelmo (1944). The label "RCM" arises because of extensions (e.g., Rubin, 1974, 1977, 1978) that allowed a formal framework beyond randomized experiments and beyond randomization-based inference; it also allowed the formal consideration of complications, such as unintended missing data and noncompliance. Because of the importance of making this very intuitive and transparent definition of causal effects central to students' thinking about causality, the course devotes some brief time to the history of this idea, including the history of needed assumptions, such as “no-interference-between-units” (Cox, 1958) and the more encompassing “stable-unit-treatment-value” assumption (Rubin, 1984). Without some such “exclusion restrictions”, to use the language of economists, causal inference is impossible, and understanding this limitation is critical.

This RCM framework is now rather generally accepted in many fields. For example, in psychology, see Wilkinson et al. (1999); in economics see the transition to adopt it reflected by comparing Heckman (1979) to Heckman (1989), and Pratt and Schlaifer (1984) to Pratt and Schlaifer (1988), after discussion by Holland (1989) and Rosenbaum and Rubin (1984), respectively. Also see Baker, (1998), Dempster (1990), Efron and Feldman (1991), Gelman and King (1991), Greenland and Pool (1988), Greenland, Robins, and Pearl (1999), Holland (1988a, b, 1989), Holland and Rubin (1983), Kadane and Seidenfeld (1990), Robins (1987, 1989), Rosenbaum (1987), Smith and Sugden (1988), Sobel (1990, 1995, 1996), Sugden (1988), and their references. A recent article exploring whether the full potential outcomes framework can be avoided when conducting causal inference is Dawid (2000) with discussion.

The essential role of the assignment mechanism is then introduced: without a model for how treatments get assigned to units, formal causal inference, as least using probabilistic statements, is impossible. It is also critical that students appreciate this, especially because most articles purporting to do causal inference in many areas of application of statistics never even consider the assignment mechanism unless the study was randomized. Examples of assignment mechanisms and a classification of them is introduced: classic randomized experiments, unconfounded and regular designs, more general ignorable designs, and nonignorable designs (Rubin, 1976). At this point, we do not NEED any more assumptions to proceed with some forms of causal inference: those based solely on the randomization distribution induced by randomized assignment. Unless students understand how to analyze randomized experiments validly and why their analysis is relatively simple, it is impossible to teach them how to analyze observational data validly for causal effects. Thus, they need to understand very well the basis of causal inference in randomized experiments, and then to use this foundation to learn how to draw causal inferences in nonrandomized studies.

There are two distinct forms of randomization inference, one due to Neyman (1923) and the other due to Fisher (1925). Fisher's is the more direct conceptually and is introduced next. It is closely related to the mathematical idea of proof by contradiction. It basically is a stochastic "proof' by contradiction giving the significance level (or p-value) of the null hypothesis of absolutely no effect whatsoever: the probability of a result (represented by the value of an a priori defined statistic, such as the difference of observed average treatment outcome minus observed average control outcome) this rare or more rare if the null hypothesis were true, where the probability is over the distribution induced by the assignment mechanism. This form of inference is very elegant but very limited: how much can be learned from finding, with high probability, a model that does not fit the data, especially as measured by some possibly arbitrary or mathematically convenient statistic?

Neyman's form of randomization-based inference can be viewed as drawing inferences by evaluating the expectations of statistics over the distribution induced by the assignment mechanism the essential idea is the same as in Neyman's (1934) classic article on randomization-based (now often called 'designed-based") inference in surveys. Typically, an unbiased estimator of the causal estimand is created, and an unbiased, or upwardly biased estimator, of the variance of that unbiased estimator is found (bias and variance both defined with respect to the randomization distribution). Then an appeal is made to the central limit theorem for the normality of the estimator over its randomization distribution, whence a confidence interval for the causal estimand is obtained. This form of inference is less direct than Fisher's; it is really aimed at evaluations of procedures, and only indirectly at inference from the data at hand. Nevertheless, it forms the basis for most of what is done in important areas of application (e.g., the world of pharmaceutical development and approval, the world of randomized clinical trials in medicine), and therefore, once again, it is critical that students understand the framework.

The third form of statistical inference for causal effects is Bayesian, where the model for the assignment mechanism is supplemented with a model for the data (Rubin, 1978). A causal inference is obtained as the posterior distribution of the causal estimand, which follows from the posterior predictive distribution of the unobserved potential outcomes, which in turn follows by Bayes theorem from the observed data and the models for the assignment mechanism and the data. All calculations are described in terms of simulating this posterior predictive distribution - basically by multiply-imputing the missing potential outcomes. This approach opens the door for the much more demanding computational methods used in the latter part of the graduate course, which would be absent from the undergraduate core course. The Bayesian approach, which is the foundation for typical linear and loglinear models, is by far the most intuitive, direct and flexible of the modes of inference for causal effects, but achieves these benefits by postulating a distribution for the data, which the randomization-based approaches avoid. Such a distribution can be very useful, but is like a loaded gun in the hands of a child, fraught with danger for the naive data analyst. This inherent danger is an important message to convey to students, especially those who have been exposed to complicated automatic causal modelling software and so may be more willing to accept and use procedures they do not fully understand.

In practice, one should be willing to use the best features of all approaches. In simple classical randomized experiments with normal-like data, the three approaches give very similar practical answers, but they do not in more difficult cases, where each perspective provides different strengths. At this point, the relationships between these three conceptually cogent formulations and the more common linear equation, regression and path diagram approaches are made. Care must be taken here, especially in an undergraduate course, to avoid confusing students through the introduction of generally inappropriate methods, which can achieve appropriate answers in the simplest settings, but which lead to inappropriate answers in realistic settings.

The course then turns to 'regular" designs, which are like classical randomized experiments except that the probabilities of treatment assignment are allowed to depend on observed covariates and so can vary from unit to unit (e.g., older males have probability .8 of being assigned the new treatment; younger males,, .6; older females, .5; and younger females, .2. Horvitz and Thompson (1952; Cochran, 1963) estimation is considered with known assignment probabilities, and Fisherian and Neymanian inferences are derived and compared with Bayesian model-based estimation for robustness and efficiency. This leads to estimated propensity score methods (Rosenbaum and Rubin, 1983), and their use in combination with models. These techniques are then applied to observational studies.

The key idea is to conceptualize the observational study as if it were a regular design, and therefore an ignorable design, and to use this template to draw inferences. Many illustrative examples are given, using matching, subclassification, and model building (e.g., Rosenbaum and Rubin, 1984, 1985;

Reinisch et al., 1995; Rubin, 1997; Smith, 1997). Using real examples, including ones where there exist nearly parallel randomized and nonrandomized studies (e.g., as in LaLonde, 1986; Dehejia and Wahba, 1999), it becomes clear that in general the combination of both propensity score methods and Bayesian model building is superior to either alone for the objective of achieving essentially the same answer from an observational study as from the parallel randomized experiment. In the graduate course, some attention is paid to theoretical results on propensity scores, such as in Rubin and Thomas (1992a,b) and on the combination of propensity score methods and modeling (Rubin and Thomas, 2000), but the most important ideas are easily conveyed by the real examples, and so the theory would be avoided in an undergraduate course, although not in the graduate course.

Because the conceptualization of an observational data set as arising from a regular (and thus, ignorable) assignment mechanism is an assumption, it is important to consider deviations from that assumption. Sensitivity of conclusions to the assumption of ignorability is then studied using methods such as those of Rosenbaum and Rubin (1985) and Rosenbaum (1995), which display how point and interval estimates change as a function of assumptions. Bounds on point estimates (e.g., Manski and Nagin, 1998; Horowitz and Manski, 1999) are also considered but usually as a special case of sensitivity analysis. Such analyses are relatively easy to explain to students at all levels. Other issues about ignorability (e.g., tests for it as in Rosenbaum, 1984) can be considered in a graduate course but not in an undergraduate course.

The topic of noncompliance in randomized studies is then introduced because it is a relatively well-understood island between the shores of the perfect randomized experiment and the uncontrolled observational study. For example, a random half of doctors are encouraged to give their patients a flu shot, but some patients in each random group do and do not take the flu shot (Hirano, Imbens, Rubin and Zhou, 1999). Such studies are ignorable for the assigned treatment (e.g., encouragement to take the flu shot) but nonignorable for the received treatment (getting the flu shot), and comprise a more general template for the analysis of observational studies than do ignorable designs. The approach taken to such studies in the course bridges classic econometric and statistical approaches to causal inference as described in Angrist, Imbens, and Rubin (1996). The most direct result is the method-of-moments "instrumental variables' estimate for the "complier average causal effect', and used, for example, by Sommers and Zeger (1991) to estimate the effect of vitamin A on infant survival. The much preferred approach to the analysis of such data, however, is likelihood/Bayesian, as described in Imbens and Rubin (1997), but this requires the introduction of iterative maximization and simulation techniques, such as the EM algorithm (Dempster, Laird and Rubin, 1977) and data augmentation (Tanner and Wong, 1987), which is appropriate for a graduate course, but is too demanding for an undergraduate core course. Severd examples of this kind of approach are given, including actual randomized experiments with noncompliance (e.g., Sommer and Zeger, 1991; Goetghebeur and Molenberghs, 1996; Baker and Lindeman, 1994; Frangakis, Rubin and Zhou, 1999) and observations studies where some version of the instrumenw variables approach seems plausible (e.g., Ettner, 1996).

The graduate course continues with the consideration of how to deal with some common complications with randomized experiments with noncompliance, such as missing outcome data (Frangakis and Rubin, 1999), randomization in clustered groups (Frangakis, Rubin and Zhou, 1999), and missing covariates and multivariate outcomes (Barnard, Frangakis, Hill and Rubin, 2000), which are all beyond the scope of an undergraduate course, as is the topic of multiple treatments in a longitudinal setting (e.g., Robins, 1997).

The final lecture is much like this little article: a review of how to apply fundamentally sound statistical thinking about causal inference to estimate causal effects in experiments and observational studies.

As stated in the abstract, I firmly believe that some such course is arguably an essential ingredient of any statistics departments' offerings, and moreover is an important component of the education of anyone who intends to draw conclusions about causal effects.

REFERENCES

Angrist, J.D., Imbens, G.W. and Rubin, D.B. (1996). "Identification of Causal Effects Using Instrumental Variables.' Journal of the American Statistical Association, 91, 434, as

Applications Invited Discussion Article with discussion and rejoinder, pp. 444-472.
Baker, S.G. (1998). "Analysis of Survival Data From a Randomized Trial With All-or-None Compliance &: Estimating the Cost-Effectiveness of a Cancer Screening Program.' Journal of the American Statistical Association, 93, pp. 929-934.
Baker, S.G. and Lindeman, K.S. (1994). "The Paired Availability Design: A Proposal for Evaluating Epidural Analgesia During Labor.' Statistics in Medicine, 13, pp. 2269-2278.
Bamard, J., Frangakis, C., Hill, J. and Rubin. D.B. (2000). "School Choice in NY City: A Bayesian Analysis of an Imperfect Randomized Experiment." To appear in Case Studies in Bayesian Statistics, V. 5, Kass et alia (eds.). New York: Springer-Verlag.
Cochran, W.G. (1963). SaMRIini! Technigues, 2nd ed. New York: Wiley.
Cox, D.R. (1958). Planning, of Experiments. New York: John Wiley.
Dawid, A.P. (2000). "Causal Inference Without Counterfactuals". To appear in Journal of the American Statistical Association, June 2000. With discussion.
Dehejia, R.H. and Wahba, S. (1999). 'Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs.' To appear in Journal of the American Statistical Association.
Dempster, A.P. (1990). "Causality and Statistics.' Journal of Statistical Planning and Inference, 25, pp. 261-278.
Dempster, A.P., Laird, N. and Rubin, D.B. (1977).
'Maximum Likelihood from Incomplete Data Via the EM Algorithm." The Journal of the Royal Statistical Society, Series B, 39, 1, pp. 1-38. With discussion and reply.
Efron, B. and Feldman, D. (1991). 'Compliance as an Explanatory Variable in Clinical Trials." Journal of the American Statistical Association, 86, pp. 9-17.
Ettner, S.L. (1996). "The Timing of Preventive Services for Women and Children: The Effect of Having a Usual Source of Care.' American Journal of Public Health, 86, 17481754.
Fisher, R.A. (1918). "The Causes of Human Variability." Eugenics Review, 10, pp. 213220.
Fisher, R.A. (1925). Statistical Methods for Research Workers. London: Oliver & Boyd.
Frangakis, C. and Rubin, D.E. (1999). "Addressing Complications of Intention-To-Treat Analysis in the Combined Presence of All-or-None Treatment-Noncompliance and Subsequent Missing Outcomes." Biometrika, 86, 2, pp. 366-379.
Frangakis, C., Rubin, D.B. and Zhou, X.-H. (1999). "The Clustered-Encouragement-Design."
Proceedinjzs of the Biometrics Session of the American Statistical Association, pp. 71-79.
Gelman, A. and King, G. (1991). "Estimating Incumbency Advantage Without Bias.' American Journal of Political Science, 3-4, pp. 1142-1164.
Goetghebeur, E. and Molenberghs, G. (1996). "Causal Inference in a Placebo-Controlled Clinical Trial with Binary Outcome and Ordered Compliance." Journal of the American Statistical Association 91, pp. 928-934.
Greenland, S. and Poole, C. (1988). 'Invariants and Noninvariants in the Concept of Interdependent Effects." Scandinavian Journal of Work and Environmental Health, 14, pp. 125-129.
Greenland, S., Robins, J.M. and Pearl, J. (1999). 'Confounding and Collapsibility in Causal Inference." Statistical Science, 1-4, pp. 29-46.
Haavelmo, T. (1944). 'The Probability Approach in Econometrics.' Econometrica, 15, pp. 413419.
Heckman, J.J. (1979). "Sample Selection Bias as a Specification Error.' Econometrica, 47, pp. 153-161.
Heckman, J.J. (1989). 'Causal Inference and Nonrandom Samples.' Journal of Educational Statistics, 14, pp. 159-168.
Heckman, J.J. (1996). Comment on "Identification of Causal Effects Using Instrumental Variables." Journal of the American Statistical Assoc. 91, pp. 459-462.
Hirano, K., Imbens, G., Rubin, D.B. and Zhou, X.H. (1999). 'Estimating the Effect of an Influenza Vaccine in an Encouragement Design." Revision to appear in Biostatistics.
Holland, P. (1986). 'Statistics and Causal Inference.' Journal of the American Statistical Association, 81, pp. 945-970.
Holland, P.W. (1988a). "Causal Inference, Path Analysis, and Recursive Structural Equation Models.' Sociological Methodology, pp. 449484.
Holland, P.W. (198&b). Comment on "Employment Discrimination and Statistical Science' by A.P. Dempster. Statistical Science, 3, pp. 186-188.

Holland, P.W. (1989). 'It's Very Clear." Comment on 'Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Man over Training' by I Heckman and V. Hotz. Journal of the American Statistical Association, 84, pp. 875-877.
Holland, P.W. and Rubin, D.B. (1983). "On Lord's Paradox.' Principles of Modern Psychological Measurement: A Festschrift for Frederic M.
Lord, Wainer and Messick (eds.), pp. 3-25. Hillsdale, NJ: Earlbaum.
Horowitz, J.L.and Manski, C.F. (1999). 'Nonparametric Analysis of Randomized Experiments with Missing Covariates and Outcome Data.' Journal of the American Statistical Association.
Horvitz, D.G. and Thompson, D.J. (1952). 'A Generalization of Sampling Without replacement From a Finite Population.' Journal of the American Statistical Association, 47, pp. 663-685.
Imbens, G. and Rubin, D.B. (1997). 'Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance." The Annals of Statistics, 25, 1, pp. 305-327.
Kadane, J. B. and Seidenfeld, T . (1990). "Randomization in a Bayesian Perspective." Journal of Statistical Planning and Inference, 25, pp. 329-346.
LaLonde, R. (1986). "Evaluating the Econometric Evaluations of Training Programs." American Economic Review, 76, pp. 604-620.
Manski, C.F. and Nagin, D.S. (1998). "Bounding Disagreements About Treatment Effects: A Study of Sentencing and Recidivism." Sociological Methodology, 28, pp. 99-137.
Neyman, J. (1923). "On the Application of Probability Theory to Agricultural Experiments: Essay on Principles, Section 9." Translated in Statistical Science, 5, pp. 465-480, 1990.
Neyman, J. (1934). "On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection.' Journal of the Roval Statistical Society, Series A, 97, pp. 558-606.
Pratt, J.W. and Schlaifer, R. (1984). "On the Nature and Discovery of Structure." Journal of the American Statistical Association, 79, pp. 9-33. Pratt, J.W. and Schlaifer, R. (1988). "On the Interpretation and Observation of Laws." Journal of Econometrics, 39, pp. 23-52.
Reinisch, J., Sanders, S., Mortensen, E. and Rubin, D. B. (1995). "In Utero Exposure to Phenobarbital and Intelligence Deficits in Adult Men.' The Journal of the American Medical Association, 274, 19, pp. 1518-1525. Robins, J.M. (1987). "A Graphical Approach to the Identification and Estimation of causal Parameters in Mortality Studies With Sustained Exposure Periods." Journal of Chronic Diseases (Suppl. 2), 40, pp. 139S161S.
Robins, J.M. (1989). "The Control of Confounding by Intermediate Variables." Statistics in Medicine, 8, pp. 679-701.
Robins, J.M. (1997)."Causal Inference From Complex Longitudinal Data.' Latent Va6able Modeling and Applications to Causality, Lecture Notes in Statistics, M. Berkane (ed.), 120, pp. 69-117. New York: SpringerVerlag.
Rosenbaum, P.R. (1984). 'From Association to Causation in Observational Studies: The Role of Tests of Strongly Ignorable Treatment Assignment.' Journal of the American Statistical Association, 79, pp. 41-48.
Rosenbaum, P.R. (1987). "The Role of a Second Control Group in an Observational Study (with discussion)." Statistical Science, 2, pp. 292316.
Rosenbaum, P.R. (1995). Observational Studies. New York: Springer-Verlag.
Rosenbaum, P. and Rubin, D.B. (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects.' Biometrika, 70, pp. 41-55.
Rosenbaum, P.R. and Rubin, D.B. (1984). "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score.' Journal of the American Statistical Association, 79, pp. 516-524.
Rosenbaum, P. R. and Rubin, D.B. (1985). "Constructing a Control Group Using Multivariate Matched Sampling Incorporating the Propensity Score.' The American Statistician, 39, pp. 33-38.
Rubin, D.B. (1974). "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. Journal of Educational Psycboloizv, 66, 5, pp. 688-701. Rubin, D.B. (1976). "Inference and Missing Data.' Biometrika, 63, 3, pp. 581-592. With discussion and reply.
Rubin, D.B. (1977). "Assignment to Treatment Group on the Basis of a Covariate." Journal of Educational Statistics, 2, 1, pp. 1-26. Printer's correction note 3, P. 384.
Rubin, D.B. (1978). "Bayesian Inference for Causal Effects: The Role of Randomization.' The Annals of Statistics, 7, 1, pp. 34-58.

Rubin, D. B. (4a4). "William G. Cochran's Contributions to the Design, Analysis, and Evaluation of Observational Studies." W.G. Cochran's Impact on Statistics, Rao and Sedransk (eds.). New York: Wiley, pp. 37-69.

The dissemination of statistical literacy among citizens and public administration directors

Luigi Biggeri, President of the Italian Statistical Society (SIS)
Università di Firenze, Dipartimento di Statistica
Viale Morgagni 59, - 50134, Firenze, Italy. biggeri@ds.unifi.it

Alberto Zuliani, President of the Italian National Statistical Institute (ISTAT)
Istituto Nazionale di Statistica
Via C. Balbo 16, - 00184, Roma, Italy. pres@istat.it

1. Introduction

The dissemination of Statistical Literacy is one of the most important challenges that statisticians have to cope with at the beginning of the 21st Century. The aims of this paper is to show: (i) why there is need for quantitative skills and, in particular, for statistical literacy in the present information age and what is the content of statistical literacy that should be disseminated; (ii) how to prepare a strategic plan, defining actors and actions, to improve statistical literacy and culture in particular among citizens and public administration directors; (iii) the activities implemented or planned in the field by the International Statistical Institute (ISI) and, in Italy, by the Italian Statistical Society (SIS) and Italian National Statistical Institute (ISTAT).

2. Needs for dissemination of statistical literacy

2.1. Preface

The fight against illiteracy in the world – under the leadership of UNESCO – has achieved considerable success. However, it is observed that the lack of literacy is usually accompanied by weaknesses in quantitative skills, while numeracy, and in particular statistical literacy, received little visible attention and are the weakest component in the programme against illiteracy, despite their apparent centrality in people's everyday lives and functioning at work.

As officers of UNESCO have said many times, the concept of literacy should be seen as an evolutionary variable, changing with the changes in economic and social situation of the world. Now, familiarity with quantitative thinking is more and more a requirement of contemporary society, while the need is much less appreciated. It is evident that in our information society –always more complex and segmented and in rapid transformation- we are living in a kind of "cage of statistical information", that is very important because it helps us to determine and define "everyday life", the quotidian issues of race, gender, economics, politics, and education, that shape and create our view. The government, the public administration units, the politicians, the trade unions, the managers and also the citizens need to use much of this statistical information, giving correct interpretation to the data and using statistical reasoning and methods for the analysis of phenomenon and, above all, to make their decisions. It is therefore evident that the diffusion of statistical literacy and culture for all is an important challenge that the academic and official statisticians have to face, and to create a future in which statistical information is truly at the service of Society and to have a Society which is more aware of its own "condition" and less subject to the fluctuations of figures.

2.2. The need for statistical literacy of the citizens

In today's life, for a general citizen at least, statistical literacy is necessary: the capacity of applying statistical reasoning and interpreting statistical information is a stronger and stronger request of the Society. The globalisation of the economy, the advancements in the technology, in particular in computers (with the capability to analyse big data bases) and in the means of communication (Internet, etc.), the increasing collection of statistical information is creating many new jobs for those who can read, follows directions, understand mathematics (basic) and interpret statistical information, that is those who have analytical, quantitative and, in particular, statistical skills. On the other hand, statistical literacy is necessary for informed citizenship: to understand "everyday life", to make rationale decisions, to have knowledge-based social control of public policy and of the activity of public administration. Shortly statistical literacy is becoming fundamental for living in a full democracy. There is no doubt that, in the age of information and computers, if we

really want to render the citizens as independent as possible and free of influence and conditioning, the only real course of action is to have them attain a higher degree of statistical culture (or literacy).

2.3. The need for statistical culture among the staff members of the public administrations
In the organisation and management of public administration units there is now a new model of production that comes from the push towards decentralisation, the development of the autonomy of each unit and the introduction of responsibility of managers for the results obtained. This requires the development of the functions of planning, management control and evaluation at the different levels at which the decision processes are formalised. It easy to show that this asks for pertinent *statistical information systems* and each phase of the process requires appropriate statistical information and indicators and the use of appropriate statistical methods to carry out the requested analysis (through cost, efficiency and effectiveness measures; simulation and impact analysis; users' satisfaction analysis; etc.). For some staff members of public administration there is a clear need for statistical skills, while for the most part, for the general staff and for the directors statistical literacy is enough.

2.4. What does statistical literacy really means?
Statistics begins with the notion that we use data to answer questions and it is important to remember that data are not just a set of numbers; data consist of a set of measurements. The issues of measurement must be a major objective of statistical education and of statistical literacy dissemination. We agree with David Moore's thinking (repeated also in Moore, 1997) that statistical education should focus on data and on statistical reasoning rather than on either the presentation of as many methods as possible or the mathematical theory of inference. Understanding statistical reasoning should be the most important objective of the dissemination of statistical culture: we should devote more emphasis to data and concepts, at the expense of less theory and fewer recipes. However we do not think that we have to disseminate statistics for all, but statistical literacy for all (see the discussion of Moore's paper by Anne Hawkins (Moore, 1997)). Numeracy and statistical literacy include therefore: facility in dealing with numbers and quantitative problems; in understanding basic mathematical ideas and patterns, statistical reasoning and the importance of thinking in terms of probability, the importance of data production and presentation, the omnipresence of variability, and the quantification and explanation of variability.
We think that the knowledge of basic concepts of statistics and probability are important, but understanding the meaning of information is essential (such as, the ability to recognise the nature and limitations of statistical information, to be able to distinguish a 'good' number from a bad one; the knowledge of the implications of using various sampling processes on the accuracy of the results; etc.).

3. The preparation of a strategic plan for dissemination of statistical culture and literacy
We need to have in mind that the diffusion of statistical culture and literacy requires the implementing a pervasive educational process, that involves a shifting in the cultural axis towards a scientific-empirical culture, that is difficult to reach. The educational process should be implemented both in school, with a systematic approach to enhance educators' awareness of the use of simple statistical methods and of statistical data for effective teaching (that should be done across the different disciplines emphasising the *problem-solving approac*h), and in the society.
However, to create and disseminate culture, and in particular statistical culture, is not so easy and does not lead to immediate and tangible results. To obtain good results it is necessary to prepare a long term strategic plan that includes all the detailed objectives and actions to be implemented resolutely.
Because the users of statistical information and methods have different characteristics and needs, and also the dissemination means of the culture are different, there is not a standardized "menu" of statistical culture. Therefore it is important to tailor specific "menu" and specific educational strategies, whether linked to statistical literacy needs or to more professional statistical education.
Therefore it is important to involve all the actors of the mentioned educational process in the preparation of the strategic plan: in particular the scientific statistical societies and the national and local statistical agencies (fostering the interaction between academic and official statisticians), but also the schools, the universities, the mass-media, the professional associations, etc.
In order to prepare a plan for the diffusion of statistical literacy and culture it will be useful to construct two matrices: the first regarding the different groups of recipients of educational actions (policy makers, directors of public administration offices, journalists, citizens, etc.) by their types of statistical needs, that shows the content of the necessary knowledge and educational actions; the second regarding the relationships be-

tween spreaders of statistical culture and literacy and recipients, that shows the diffusion process and the possible actions and the means to be used.
In this way it is possible to target the statistical "menu" to the specific audience. For example, with reference to the public administration units the need for statistical training is quite different for different staff and, above all, we have to distinguish between the preparation or training required of the new employees and the training necessary for the personnel in service (for the directors it will be useful to organise specific seminars and/or short course or workshops). It is also important to evaluate the needs for knowledge in different subject matter disciplines in connection with statistical needs.
In general, from this kind of exercise it is clear that some organisations must do the diffusion of statistical literacy (such as the national statistical agencies, the statistical association and the schools) and that some organisations must be at the same time spreaders and recipients of the statistical literacy educational process. Some of them, that are *intermediary spreader*s, have a very important role in the educational process, at least for the citizens (school teachers, journalists and mass-media, libraries).

4. The actions carried out by ISI and in Italy, by SIS and ISTAT, for dissemination of statistical literacy

We would like to conclude this paper by presenting a short summary of the activities implemented or planned in the field by the ISI and, in Italy, by SIS and ISTAT.

4.1. ISI and statistical literacy

As reported by R. Smulders (1995), the ISI established a Committees on Statistical Education that set up a number of Task Forces on teaching statistics at the school level and on international Conferences on statistical education (the most important of them being the International Conference on Teaching Statistics – ICOTS – by a general consensus of both statisticians and non-statisticians).
As suggested by the Committee, recently the ISI established the International Association for Statistical Education (IASE) recognising the importance of the diffusion of statistical education.
Finally, the ISI launched the World Numeracy Programme (WNP) in order to spread quantitative skills all over the world in areas and populations (especially in developing countries and among the young) which could benefit most from increased knowledge of numbers and their applications, with particular regard to statistics. The programme includes various projects whose implementation requires 7-8 years and for some of them ISI has asked for financial help from UNESCO. The activities are on going, but the development of the projects are meeting various difficulties.

4.2. The diffusion of statistical literacy in Italy

In Italy there is now a very important and fruitful interaction between ISTAT and SIS to improve the diffusion of statistical literacy and culture. In fact in Italy, by law, statistical information is considered a public good and ISTAT has among its duties also the dissemination of statistical culture. SIS, as are other statistical associations, has been engaged for many years in the diffusion of statistical literacy and culture through a specific centre (CIRDIS) established in Perugia and other Italian Universities.
SIS and ISTAT have already organised a First National Conference on the Dissemination of Statistical Culture (in 1997), with the co-operation of the Ministries of Education and of the Public Administration. Now, together with the Ministry of Education, they are organising a specific programme to carry out a "Census of Children" in all of the Italian Elementary and Primary Schools (as it is already done in New Zealand). The project has a short term objective to sensitise the citizens to the population census of 2001 and a long term objective to foster the diffusion of statistical culture and literacy in the schools. The didactic value of the experience that will be carried out (the children, their teachers and families will be involved in the collection and analysis of data) will represent an inductive approach to probability and statistics in most of the disciplines.
Especially ISTAT is also doing a lot of other activities for the diffusion of the statistical literacy and culture in the public administration units: diffusion of statistical information by the Internet and by television; dissemination of statistical information at the local level with specific points for access to the data (including links with the libraries); preparation of popularizing publications for the citizens and for the schools; seminars for journalists and opinion leaders; training for the public administration staff, and in particular for the public statistical offices. The need of a wide and disseminated growth of statistical culture in the civil administration has focused the need for planning training and spreading proposals. Solutions has been pioneered in planning training paths with the aim of strengthening professional skill, in emphasising the role of

statistics and statisticians in planning and controlling within the civil administration defined aims. The present training plan foresees four different education processes for the development of statistical skills, from statistical literacy to advanced statistical skills.
All the experiences have showed some important lessons in order to reach some durable results: first, the main obstacles are the "refusal" to use statistics and the preservative cultural context; second, we have to implement a "politics of alliance" with the intermediary spreaders to getting in touch with the receivers in an appropriate way; third, the implementation of the diffusion of statistical literacy and culture is a gradual conquest, that require thousands of small diffused activities in the territory.

REFERENCES

Alston, S. M. (1989). Census Data and Mathematical Literacy: Educating Society for the 21st Century. Paper presented at the 1989 Joint Statistical Meetings on August 9, Washington, D.C.

Gal, I. (1993). Issues and Challenges in Adult Numeracy. Technical Report 93-15, National Center on Adult Literacy, University of Pennsylvania, Philadelphia

ISI (1995). Education Statistics: an International Compendium, Voorburg, Netherlands.

ISTAT and SIS (1997). Conferenza sulla diffusione della cultura statistica. Roma (in press)

Moore, D. S. (1997). New Pedagogy and New Content: The Case of Statistics (including discussion), International Statistical Review, 65, 2, 123-165.

Perry, M. and Kader, G. (1992). Stat-linc and Stat-maps: Teacher Education and Curriculum Design. Proceedings of the ASA Meetings, Section on Statistical Education, USA

Smulders, R. H. M. (1995). ISI and Education Statistics. In ISI (1995).

Résumé

La diffusion de la culture statistique est une de plus important défies que les statisticians devront débrouiller dans le 21emé siecle. Cet communication montre que le besoin de la culture statistique (de base) est tojuors plus importante dans la societé de l'information pour les citoyens et les directeurs de l'administration public. Il est donc nécessaire de préparer un plan pour l'établissement d'un parcours educatif et de formation pour la diffusion de la culture statistique. Quelque renseignement est aussi fournit sur les activitées développées de part de l'Institute International de Statistique et, en Italie, de la Societé Italienne de Statistique (SIS) et de l'Institute National de Statistique (ISTAT).

Source: http://www.stat.fi/isi99/proceedings/arkisto/varasto/bigg0981.pdf