COMMENTARIES

Commentaries are informative essays dealing with viewpoints of statistical practice, statistical education, and other topics considered to be of general interest to the broad readership of *The American Statistician*. Commentaries are similar in spirit to Letters to the Editor, but they involve longer discussions of background, issues, and perspectives. All commentaries will be refereed for their merit and compatibility with these criteria.

Embracing the "Wider View" of Statistics

C. J. Wild

The interconnected themes of quality and the marketing of the discipline of statistics are explored. An understanding of statistics as the study of the process of scientific inquiry is advocated as a consciously targeted market position. Because it reaches such a high proportion of the managers and decision makers of the future, the introductory university or college statistics course is highlighted as a potent marketing opportunity for enhancing the long-term health of statistics. Attention is given to teaching students to think statistically, to become educated consumers of statistical expertise, and to communicate well with nonstatisticians.

KEY WORDS: Defining statistics; Marketing statistics; Quality; Statistical thinking; Teaching statistics.

1. THE "WIDER VIEW"

1.1 Introduction

The ideas for this article stemmed from two different projects, namely the writing of an introductory statistics textbook with G. A. F. Seber and participation in a team trying to implement total quality management (TQM) in parts of our teaching program. A central precept of TQM is that every product, service, or task has a customer and that one should strive to meet and even exceed the needs of that customer. The customers for statistics courses are many and varied. They include not only the students, the other courses for which the particular course is required, and the discipline itself, but society as a whole as a consequence of government funding. The students occupy the unusual position of being both customers and the product!

Apart from the other university courses that require parts of the statistics program, it is very difficult to ascertain what the customers think their needs are: They tend to be very vaguely defined, often coming down to some idea of general numeracy, and on the part of students, empolyability. Occasionally, particular customers express needs that are far too specialized to be accommodated in broadly based courses. To a large extent, we have to gauge and balance the needs of our customers ourselves. We need to give careful consideration to the aims and objectives of every part of our programs. If we have not thought through very carefully what we are trying to achieve, we are in no position to assess the quality of what we provide. Unfortunately, so much of what we do is not thought through from a careful consideration of customers, aims, and objectives; it just grows in an ad hoc way over the years, building on what has been done before. This has certainly been my experience and all too often my own practice. At its most basic, considering aims and objectives requires thinking about the nature of statistics itself. And yet the nature of statistics as a discipline is something that has never been properly resolved, agreed on, and then communicated.

At a TQM seminar late in 1991, the human resources manager of an international hotel chain was overheard talking to a local businessman. The manager, a recent convert to TQM, was saying that he would just have to bite the bullet and try to come to grips with statistics, although he did not know if he could because it was all so unutterably boring. The businessman, who had a sideline in software, was obviously the consummate salesman. He whipped out a floppy disk and said, "*You don't need statistics, you need this*. It does all the statistics for you. Produces all the control charts, the Pareto diagrams ...." The hotel manager was obviously relieved. He said how great it was that he would not have to do the grinding out of the means and standard deviations and so forth that both of them understood statistics to be. In our view, these people were missing the point that it is not a matter of needing a package to understand and monitor processes *instead of statistics* but that trying to understand such processes through the most effective means available *is statistics*. This is not, however, a viewpoint that is consistently stressed by statisticians.

In fact, statisticians have failed to communicate any coherent picture of the nature and scope of the discipline. "Statistics has an identity problem. It is not visible as a discipline" (Minton 1983, p. 286). "How do nonstatisticians view statistics as a discipline? Can the answer be not at all or almost that?" (Bradley 1982, p. 2). "The profession of statistics has adopted too narrow a definition of itself. As a consequence, both statistics and statisticians play too narrow a role in policy formation and execution" (Mosteller 1988, p. 93). Marquardt (1987, p. 1) lamented

C. J. Wild is Associate Professor, Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand. The author wishes to thank his colleagues at the University of Auckland and Jock MacKay of the University of Waterloo for the years of random conversation that gave rise to many of the ideas in this article and to thank an associate editor for some helpful comments.

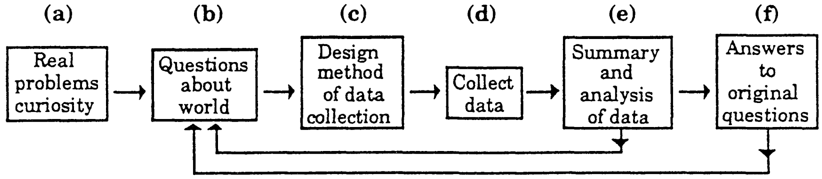

Figure 1. The Investigative Process.

the "lack of visibility of statistics as a discipline" and its "lack of influence." "A sense of dissatisfaction exists in the statistics profession stemming from a consensus that statisticians in all working environments are undervalued and underutilized" (Boroto and Zahn 1989, p. 71). "... we need to broaden both our own view of the field of statistics and the missions of statisticians and the view held by the general public" (Mosteller 1988, p. 99). Sentiments like these have been a recurrent theme in many of the American Statistical Association's Presidential Addresses of the past 10 years (see also Cockerill and Fried 1991).

Marquardt (1987) described three views of statisticians. In the narrow view, statisticians are seen as compilers of data. The middle view sees statisticians as specialists or consultants for defined problems. But in the wider view, the view that Marquardt (1987) advocates, statisticians are seen as "systems guides for, tough, fuzzy issues wherever the collection and interpretation of data is involved" (p. 2). "Statistics is the discipline responsible for studying the scientific method with the greatest intensity and for providing in-depth expertise to other disciplines" (p. 4). It produces "purveyors of the scientific method" (p. 6). Boroto and Zahn (1989) went even farther. Their ideal is the "master statistician" who is "some combination of purveyor of the scientific method, practitioner of the scientific method, and entrepreneur" (p. 72).

So how should one define the subject of statistics in a way that emphasizes the wider view?

1.2 Focusing on the Investigative Process as a Whole

David Finney once described statistics as "making sense of figures." To Marquardt (1987), the subject matter of statistics is the scientific method. Finney's phrase appeals because of its simplicity and directness. It fails to encapsulate statistics, however, because it includes no reference to how the figures were obtained in the first place, or why they were obtained. Marquardt's definition is suitably broad, but the language has the negative connotations of laboratories, test tubes and so on for students who are studying subjects that they do not think of as "sciences."

The definition of statistics that I use is essentially obtained by adding to Finney's phrase (with a consequent lessening of that expression's simple appeal). Statistics is "concerned with finding out about the real world by collecting, and then making sense of, data." I like to combine this definition of statistics with a discussion of the flow chart given in Figure 1, which describes the process of inquiry. Statistics is to be thought of as the kernel of this process—the common core that is left when one strips away the contexts of particular investigations.

The process is driven either by real problems or by simple curiosity about the world. These problems give rise to questions that might be answered with data. This part of the process is probably not statistics at all in the lexicon of those who take a narrower view than Marquardt does, because it is driven primarily by the priorities of the subject matter under study. Most of us know from experience that there is a role for statisticians in the question-generation part of the process. Our personal role may just have been to force subject-matter investigators to clarify their own thinking and make sure that the questions being asked are precise enough that they might feasibly be answered by using the available sources of information. However, much more must be made of it than this in our teaching. Investigation, including the generation of questions, is not some arcane scholarly activity found only in subject-matter research and best locked up in chemistry, economics, sociology, and so forth. It is a general life skill. And not only is question generation arguably the most important part of the investigative process, the bubbling up of questions from an awakened curiosity provides much of the excitement of investigation. If that is left out of statistical education, then training in statistics will be poorer and much less valuable than it should be.

The two most important feedback loops in the process have been represented in Figure 1: that in which the data suggest more questions and that in which the answers to the original questions suggest more questions. One might argue for more feedback arrows, but I do not want to complicate the diagram to the extent that the principal messages are lost to students. Following question generation, the remainder of the elements of the flow chart are indisputably parts of statistics. There is nothing new about this chart. George Box has been using diagrams similar to this for decades.

The process of investigation as a whole should be the heart of any statistics program, particularly of the basic introductory course Stat101. It is something one should never lose sight of and should always come back to.

2. EMBRACING THE WIDER VIEW

2.1 Market Position

It must be astounding to members of other professional bodies that a president of the largest and most influential professional body for statisticians should have to argue impassionedly, with members of his own profession about the nature of his subject at such a basic level as Marquardt (1987) was compelled to do. It is scarcely surprising that the public at large has little idea of who statisticians are and how valuable their skills might be to society if the profession as a whole has no clear idea of what it stands for. If we are going to present a coherent image to the public we must first, through our professional bodies and bearing in mind what the subject has encompassed in the past, make a decision about what we want statistics to be. This vision of statistics should be able to be expressed in very simple terms.

In setting up their discussion, Boroto and Zahn (1989) expressed doubt that statisticians demonstrate a particular attachment to the scientific method. They pointed to the failure of statisticians to *use* the scientific method in the production and assessment of the services they provide and to holes in the standard curriculum with regard to things like research design, measurement theory and technology, and open dialogue. Individual statisticians may not be ideal purveyors of the scientific method. Ordinary mortals working in any academic discipline seldom cover the full extent of that discipline, and statistics courses may not (yet) adequately cover all aspects of the scientific method. The wider view, however, is something we should *embrace*. We should consciously decide to *make* statistics the study of the process of scientific inquiry. In other words, that is where we should aim to position ourselves in the academic marketplace. Almost all of the elements are already there, and it is wholly consistent with the work and professional interests of the giants who built the subject. It is just a matter of nudging statistics a little farther "back toward its roots in scientific inference" (David Moore as quoted by Cobb [1991]).

The hotel manager and the businessman who sold computer software did not perceive statistics as being about investigation at all. To them, statistics was about calculations, and worse than that, about long, tedious calculations done almost entirely by hand. I have noted the complaints about the lack of visibility and influence of statistics and the general undervaluing of the expertise of statisticians being voiced by leading members of the profession. But when statistics becomes clearly embedded in people's minds as being concerned with *investigation* rather than simply *calculation* (or worse still, mere cataloguing of data), there can be no room for doubts about the relevance of the subject. By embracing the investigative process as a whole as the focus of the subject, we are securing it to something that will always be a central human concern. The *t*-test may come and go, but statistics (and therefore statisticians!) goes on forever.

Having embraced the wider view of statistics, we then have to change public perceptions. How can we encourage this change in the public perception? Although many approaches are needed, the potential power of the introductory statistics course and its textbooks for changing perceptions in society should be emphasized. Most of the definitions of a discipline that are ever read are given in introductory textbooks. There are few other settings in which any attempt is made to describe a subject as a whole rather than focusing on specialized corners of that subject. As a group, writers of introductory books therefore have a good deal of power in the establishment of subject definitions. But of course the right subject definition is not enough. The books and courses then have to embody the definition they promulgate. Cockerill and Fried (1991, p. 174) noted that statistics "is one of the most widely required background courses in a broad variety of university undergraduate and graduate programs." Here at the University of Auckland (New Zealand), we teach such courses to one in every three of the students who enter the university gates. If those students go away with their misapprehensions about the nature of statistics intact, it is plainly obvious with whom the fault lies.

2.2 Marketing, Quality, and the Rule of 100

An engineer involved in quality improvement in a manufacturing industry recently visited our department. The role of experimental design in his plant came up in conversation. The engineer claimed to be an advocate of statistically designed experiments but stated that situations requiring them in his company turned up at most twice a year. This was obviously insufficient for the company to need an on-site specialist trained in experimental design. Statisticians often claim that industry does not recognize the potential of designed experiments. How do we overcome this? A knee-jerk reaction would be to say that we need to teach more people going into industry how to design experiments. This story points to a different reality. If the management of the manufacturing plant believe that they have only two situations a year that call for a designed experiment, what they need on site may not be someone who can design experiments but several people who can recognize those two situations when and where they arise and call in an outside consultant.

We have to take a step beyond teaching people how to design experiments, for example, to teaching *mainstream* engineers, business students, and so on, what experimental design can do for them and how to recognize where it should be used. With people on site who recognize where experimentation is required, the aforementioned manufacturing plant may begin to realize that there are considerably more than two situations a year that require experimentation and hire a specialist.

Publications such as *The American Statistician* sometimes contain articles that discuss technical aspects of teaching experimental design or other aspects of statistics to a target group such as engineers. In the future, let them actively seek articles with titles such as "How to Convince Engineers of the Need for Designed Experiments." This has to be done for many aspects of statistics. Our natural talents for self-promotion are so poor that ideas and materials produced by those rare statisticians with a talent for marketing must be quickly and widely disseminated.

The quality-improvement movement has been advancing its cause by collecting and publicizing case studies and success stories. We need to adopt this strategy much more widely. We need repositories of success stories from many target areas so that we can show people what statistics can do for them. One might argue that many textbooks already have case studies. What I am calling for here is different: Case studies given in statistics textbooks show statistics at work in different application areas to give statistics students a feeling that what they are learning is relevant and useful. The case studies tend to stress technical statistical content, and the priorities are statistical. Case studies and success stories for the selling of statistics must be written not only in language the target audience understands but in terms of the priorities of that audience, not the priorities of statisticians. Better still would be testimonials from satisfied clients. As Roberts (1987, p. 270), wrote, "there are few testimonials from businessmen about the value of statistics." This is not something that is beneath us. If we actually believe that what we teach has value in the real world, we are short-changing society by not doing it.

What is our "rule of 100"? For every theoretical statistician that society needs, it needs 100 applied statisticians; for each of those applied statisticians, it needs 100 nonstatisticians who can recognize where the applied statisticians' skills are required. We can look at the rule of 100 another way. To generate a new job for one statistical theorist, we have to generate new jobs for 100 practitioners. To generate each practitioner's job, we need to convince 100 others that they have problems that require the involvement of statisticians. One might argue about the power of 10, but it is hard to argue with the pattern. We need to recognize this reality in designing educational strategies and in designing marketing strategies.

"Our current marketing approach in the field of statistics is dooming us to perpetual status as 'hangers-on' in the mainstream of society" (Marquardt 1987, p. 3). Saying that selling our subject is something that we are not particularly good at grossly understates the reality. Otherwise, statistics would not have a reputation for being boring, if (at best) marginally worthy. Presidential Addresses and Editor's Invited Columns in *The American Statistician* would not need to bemoan the lack of appreciation and visibility of statistics and statisticians. In marketing, we need all the help we can get. Adopting a simple, coherent view of the subject must make the marketing job easier. Making that coherent view the wider view as discussed before would make the job easier still because it addresses the basic human need to find out more about the world.

Minton (1983) and Marquardt (1987) make good strategic suggestions for improving the marketing of statistics, including the fostering of bachelors-level degree programs in statistics. Many actions are needed to change perceptions in the short term. But one of the most potent marketing weapons for long-term change is Stat101, particularly in its service-course incarnations. At Ohio State University, every undergraduate is required to take an introductory course in statistics. Although we do not do nearly so well at Auckland, we now reach one in every three undergraduates. Included in that one in three is 80% of all business students. These courses are now reaching a substantial proportion of the people who will become the managers and decision makers of our society. No other discipline has an opportunity for marketing itself like this.

Unfortunately, this enormous marketing opportunity has largely been squandered. These courses have acquired a reputation for being difficult, boring, and not very useful, leaving students with a view of statistics as "an unimportant, mechanical, uninteresting, difficult required subject, an abstruse corner of the subject XYZ" (Minton 1983, p. 285). Experienced faculty often regard the avoidance of such courses as, in the words of an Associate Editor, a coveted privilege of rank. The courses are often taught by nonstatisticians, but even when statisticians do teach them, they have often not done any better.

University statisticians who turn their noses up at involvement in the provision of service courses (thus abandoning their teaching to others), and those who give them very low priority, are doing a great disservice to statistics. If, however, we take the service commitment seriously and address the quality issues that arise both from believing that the subject matter of statistics is the investigative process and from focusing on the needs of customers, then because of the large captive audience, the marketing of statistics and statisticians comes for free. One important illustration of this idea follows.

Any serious consideration of the needs of the range of customers for Stat101 will reveal that in addition to mastering a range of basic technical skills, students should be taught to recognize when situations or problems require a statistical approach, even though the solution of those problems may be beyond their technical capabilities. (They also have to be taught to recognize the limits of their own technical expertise.) *We* have to teach nonstatisticians to recognize where statistical expertise is required. No one else will. We teach students how to solve simple statistical problems, but how often do we make any serious effort to teach them to recognize situations that call for statistical expertise that is beyond the technical content of the course? And when issues that fall outside the technical content of the course cannot be avoided, the standard response is to provide references to higher-level books and papers that most of the students will never be able to read and understand anyway. In the book Seber and I are completing, we sometimes supplement the references with a sprinkling of boxed slogans like this one from a discussion of regression diagnostics:

> ***If you see any of these problems, you are out of your depth and should consult a statistician.***

This is not blatant advertising at its most crass. It is the best advice we can give and as such is a valid response to the real needs of a customer.

Most existing service courses try to teach students to be low-level statistical practitioners. By teaching students in service courses to recognize problems that require a statistical approach and to recognize the limits of their own technical expertise, we will also produce people who will be educated consumers of the technical skills of statisticians. We will create an increased demand for advanced statistical expertise. I reiterate that the people taking these courses will make up a substantial proportion of the managers and decision makers of the future. By moving to address a real need of our customers, the improved marketing of statistics and statisticians comes for free.

2.3 Equipping the Troops

An issue that is related to convincing people of the potential benefits of statistical services is the reversing of the widespread reputation of statistical service courses as being boring and largely irrelevant. Much has been written about the ideal statistics teacher at the university or college level. He or she should be actively involved in both statistical research and applied statistical practice and be able to bring the fruits of this experience into the classroom to give the course the relevance and gritty realism it requires. In addition, he or she should have the time to run courses with heavy project requirements.

Bailar (1988) wrote, "Statisticians often lament that they have no subject-matter area of their own—that they work in everyone else's area (engineering, medicine, agriculture, etc.)" (p. 8). In terms of making Stat101 or any applied statistics course interesting, being able to work in everyone else's backyard should be seen as conferring a huge advantage that no other subject has. We *should* be able to make sitting in our courses interesting by touching on some of the most fascinating parts of other subjects. (You cannot go too far wrong with sex, drugs, and rock and roll!) They only excuse we need is the presence of data.

The reality is usually far from these ideals. The majority of teachers of statistical service courses are overworked, and many are underqualified (e.g., inexperienced graduate students). They cannot afford to have their teaching commitments eat up large proportions of their research time. Their own personal experiences may only rarely be able to be used directly. For example, the data structures used may be too complex, or perhaps they usually work with medical applications, but this term they have to teach business students. Their department may be far too short-staffed to make much project work feasible. Looking for sufficient recent, exciting, relevant, real-world examples takes enormous effort and more time and general knowledge than most teachers have.

Having a few extremely talented and committed individuals producing excellent Stat101 courses will not accomplish the transformation of society's perceptions about statistics and statisticians that is so badly needed. We have to improve the average quality of all Stat101 courses. We have to tap the resources of the profession as a whole to *make it easy* for an overworked teacher to teach a better Stat101, *without having to make a substantially greater time commitment*, by making available interesting stories and applications that can essentially be simply pulled off the shelf. The Annenberg Foundation's video series *Against all Odds* (COMAP 1989) and books such as that of Tanur et al. (1987) and Brook, Arnold, Pringle, and Hassard (1986) have made an excellent start. Some textbooks have included excellent case studies. But stories in book form are too static. A story that is breaking on today's news is better than one from last month, which in turn is better than one from last year. Certainly there are some evergreen stories with a very long lifetime of potential use, but the teachers of Stat101 need a ready supply of topical stories.

An Associate Editor further advocates shattering existing dominant ecducational paradigms, exploiting learning technology, and making active learning the norm in Stat101. This brings with it the additional need to disseminate to statistics teachers descriptions of activities that give students experience of the statistical process and complement the few better-known ones such as Deming's red bead experiment (Deming 1986) and the beer game cited by Senge (1990). Exciting developments such as the series of physical experiments used in Stat231 at the University of Waterloo (Ontario, Canada) and Math103 at University of Adelaide (Australia) deserve publicity. Experiments such as these have the most impact when performed by the students themselves but can still be used as classroom demonstrations when necessary.

Statistical societies should realize that improving the *average standard* of Stat101 is crucial for the future health and reputation of our discipline and the demand for its practitioners. They should actively encourage the publication, within the pages of their prestigious general journals, of suitable enrichment stories and activities and of articles about how to convince target groups of the benefits that statistics can confer on them. This would send a powerful signal to statisticians that the production of such materials is a valued professional activity, one at least as important as producing a simpler proof or finding a better way to present a theoretical concept. For the stories, the only criteria for publication should be that they can be used in teaching Stat101, and that they have a real potential to excite students. This policy would not only stimulate such activities, it would make the results more readily accessible, and with a quick publication policy stories need not be more than two or three months old when the readership gets them.

3. SOME IMPLICATIONS FOR TEACHING

3.1 Introduction

This section discusses some of the implications for the way we teach, which follow from viewing statistics as being concerned with the investigative process as a whole and from a desire to meed the needs of customers. No attempt has been made to be comprehensive. Many of the ideas originated from thinking about improving the quality of Stat101 but carry over to any course in applied statistics. They also grew out of the environment in Auckland, where we are short-staffed but teach large numbers of students.

There are some specific requirements from client courses that use techniques and ideas from Stat101 that help define its content. Beyond those requirements however, there can be a lot of flexibility. Most of the Stat101 students will not take more statistics courses, so very little in the course can be justified in terms of laying the groundwork for future study. For most, this is the one chance we get to make an impact. We have to prepare a package that has a good chance of being useful in the student's life and future career. Boroto and Zahn (1989) wrote, "The content of statistics makes a contribution to the community only when applied to a consumer's problem" (p. 72). A package that meets the requirement of usefulness in the life of the average student goes a long way toward meeting the needs of society as a whole. Because statistics is concerned with investigations that use data, however, another important goal should be laying the groundwork both for improving the quality of investigations that go on in society and for increasing the frequency with which such investigations are carried out. Perhaps a few more of the many political and commercial decisions that affect us all will then be based on good solid information. As Snee

(1988, p. 31) wrote, "In God we trust, others must have data."

These are serious considerations, not just for Stat101, but for a statistics program as a whole. With obvious approval of the sentiment, Bailar (1988), talked about letters she received "from faculty members who deplored the fact that the colleges and universities were not teaching statistical thinking. They were teaching mechanical manipulations" (p. 7). In trying to think through the implications of trying to improve both the quality of investigations and their frequency, the process has been broken down under four headings: (a) mental habits, which tries to address aspects of the idea of statistical thinking and bridging the gap between technical statistical knowledge and statistical thinking; (b) data collection; (c) data analysis; (d) communication. Much has been written about (b) and (c). In the remainder of this article, I concentrate principally on (a) and (d).

Mental habits. At conferences, I have listened to several panel discussions about statistical consulting in which the discussants were asked to list the most important skills they brought to an investigation. What skills do the consultants list? They never say something such as, "I'm a real whiz at repeated measures." It is much more common to hear them quote skills such as "asking the right questions," "making investigators confront their own assumptions," "problem-recognition skills," and even "simple practical common sense." These characteristics concern mental habits (rather than technical skills), which play a large part in statistical thinking. Where did the consultants pick up these mental habits? Their statistical training may have sowed some seeds, but it is likely that they are the result of years of experience and the sharing of experiences and horror stories with other statisticians. It would be very beneficial to society as a whole if these mental habits could be instilled early and widely, beginning with Stat101.

Most crucial to the investigative process is the habit of curiosity, of the continual bubbling up of questions. It is so basic and so infrequently touched on in statistics courses that a slogan such as "*Questions are more important than answers*" may be needed to emphasize it. After all, the generation of questions precedes the provision of answers. When an interesting or important question is posed, one can always look around for a specialist to help provide the answers, but if the question is never posed, then nothing will be learned. A member of a collaborative team who answers only technical statistical questions that have been postulated by others, no matter how difficult those questions are, tends to be looked down on as a mere technician and not embraced as a full and equal member of the team. Moreover, he or she will miss out on most of the fun and excitement of the investigation. As future participants in the investigative process, we should encourage the mental habits of curiosity in our students. The dimensions of statistical thinking that draw on mental habits are not things we can teach directly. We can talk about their importance, but we can only hope to instill them tangentially and through repetition. Some suggestions for doing this are given in Section 3.3.

In its most important applications, the inquiry process is aimed at learning, certainly, but learning for a purpose. Therefore, as an Associate Editor points out, statisticians need to be involved in investigating, improving, and teaching strategies for active learning. For example, Deming's Plan, Do, Study, Act (PDSA) cycle is one means of incorporating methods of learning about the real world and then improving as a result.

Data collection. Students should come out of their contact with statistics knowing about the various types of statistical studies: surveys, observational studies, and experiments, their roles, strengths and weaknesses. They should have an appreciation of the ideas of stratification and blocking and know that varying one factor at a time is a bad strategy. Boroto and Zahn (1989) identified measurement theory as an important but neglected area. For most students, contact with statistics stops at Stat101, so all of these ideas must be in Stat101 even though we do not have the time (and the students do not have the background) to teach the details.

Data analysis. The purpose of data analysis is to facilitate human understanding of data. This provides an overriding criterion for judging course contents, not just in Stat101 but in all of applied statistics.

Advancing technology has made enormous changes in statistical *practice*, and we are on the edge of even more fundamental changes. We seem to be headed for an era in which the primary language for promoting the human understanding of data will be sophisticated computer graphics rather than mathematics and in which computer-intensive resampling procedures will to a large extent replace the assumption-ridden mathematical models that we have grown up with (see Efron and Tibshirani 1991).

Anything that helps human beings understand numerical information, wherever it is developed, is the province of statistics, to be adopted and popularized by statisticians and to displace less useful parts of mainstream statistical curricula. This includes much of the visually stunning dynamic graphical work being produced by the renaissance teams at the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign. A statistics that is focused on facilitating the *human understanding* of the masses of data that the computer revolution is spawning becomes the natural inheritor of that revolution. If the current means are allowed to obscure the ends (human understanding), however, statistics risks becoming a historical anachronism whose true role is taken over by other groups, for example, those springing up largely within computer science departments working in the subarea of "computational science" which they call "scientific visualization." Because it is concerned with the use of computer graphics to "communicate" numerical data and mathematical models to human beings, this is surely a subset of statistics.

If in the past our collective focus had been on promoting the human understanding of data, the historical debates about statistical packages would never have arisen. Of course, statisticians should teach people how to use packages as they can help people to understand data more quickly. Luddite reactions about the dangers of misuse are irrelevant. All powerful tools can be misused. Our job is to teach students to use the best and most powerful tools, to use them properly and to the best effect. It is almost unbelievable that as late as 1990, Dallal (1990) had to defend teaching the use of statistical packages.

Acceptance of the idea that the purpose of data analysis is to facilitate the human understanding of data leads to a recognition that some areas statisticians have neglected are of central importance. For example, the serious deficiencies that Mosteller (1988) pointed out in research on the relationships between graphic presentation and human information processing (see Kosslyn 1985) and the lack of attention given to the presentation of data tables (Ehrenberg 1978, 1981; Mosteller 1988) lie close to the heart of statistics.

Communication. Communication with others is an essential part of the investigative process, and yet Bailar (1988) stated the following: "Communication about statistics is rarely taught at universities. But we are the losers—as individuals and as a profession—when people brush us off as unintelligible, narrow specialist" (p. 6). Many statisticians "are ineffective at presenting that work (i.e., technical) to a nontechnical audience. Statisticians must learn better communications skills" (Bailar 1988, p. 6). "The statistician must be fluent in the language of the customer ... an *effective* statistician is essentially a skilled translator" who must be in constant "dialogue (rather than monologue)" with the customer (Boroto and Zahn 1989, p. 71). Improving communication skills is again something that can be tackled tangentially. Some ideas for doing so are discussed in Section 3.4.

The hotel manager and the businessman in Section 1.1 thought that statistics was about the hand calculation of means and standard deviations. We emphasize nuts and bolts in statistics courses to such an extent that everyone loses sight of the real point of it all, namely the collection and making sense of data *to answer questions about the real world.* Moore (1990), talking about recommendations for a school curriculum that places a great deal of emphasis on data analysis, wrote, "it is easy to view statistics in particular as a collection of specific skills (or even a bag of tricks)" (p. 95). The purpose of the elements in the "bag of tricks" has to be explained in terms of how they fit into the larger picture that is the investigative process. And although most applied courses will continue to be about techniques (designing studies and analysing data), if we wish to develop statistical thinking, ways have to be found to continually remind students of other elements of that wider context.

Teachers like to present the world in nice self-contained chunks, but the investigative process is anything but self-contained. We have to forego some of the comfort of a neat and tidy self-contained world in which answers are right or wrong. It might be nice, but it is not statistics. The mindset we need to encourage is one of continually probing. We *should* raise issues that we cannot deal with when they are important practically. Students must also be taught about practical issues on the periphery of the topics they are learning about so that they can recognize the practical limits of their own knowledge. For example, a discussion of such features as nonconstant variance and influential observations and diagnostics which reveal them belongs with any introduction to regression, even though one may have no intention of teaching techniques for handling these problems. We should not teach least squares without teaching the limitations of least squares. And we should not be coy about reminding students that when one has a statistical problem that is beyond one's capabilities, one should take it to a professional statistician.

3.3 Keeping the Context With the Data

We cannot convince anyone that statistics is about the use of data to answer questions about the real world without using real data in our teaching. There have been enormous improvements in this regard in elementary texts since the appearance of path-breaking books in the late 1970s. Now, no self-respecting textbook can get by without being laced with examples, exercises, and case studies that are based on real data. Indeed, after reading the vivid prose of Cobb (1987), it is difficult to make up any example without suffering from near-terminal embarrassment.

The main reason for using real data appears to have been to convince the reader that statistics is relevant to the world in a way that is far more compelling than simply by telling them so. There can be side benefits in retaining context, too. Most people like gossip; they like to hear interesting little stories about the world. In contrast, most people find statistical calculations boring. By using real data on subjects with gossip value, we can reward students for the effort they have put into calculation or even for their attendance at class by their learning something new and interesting about the world.

In my experience, however, the real data tend to be used simply to illustrate or give practice in the techniques that have just been taught. But as has been seen, teaching techniques is not enough. We need to teach statistical thinking, including the mental habits that phrase implies, such as the "inquisitive, scientific attitude" that Bradley (1982) appealed for.

Moore (1990) described data as "numbers with a context." The ideas that follow all hinge on keeping considerably more of the context with the data than is usually done and then using that context to help instill elements of "statistical thinking" and to give constant reminders of the investigative process as a whole. We need to instill mental *habits*, and habits are acquired by doing something over and over again until it becomes automatic. One basic idea is to make a practice of adding standard questions to the end of every data-analysis assignment. For example, to encourage an inquisitive frame of mind and a skeptical attitude to sources of information, we might ask a pair of questions such as "What further investigations would you like to see done that are (a) prompted by weaknesses in the data or data collection or (b) prompted by the subject-matter context of the data?" Questions such as this should *always* be present because our purpose is the formation of a mental habit, not the learning of a technical skill.

Generation of subject-matter questions requires subject-matter knowledge. This is another reason why it is so important to try to give some of the context of the investigation along with the data in data-analysis problems. We do not want to develop graduates who keep what they know about psychology, for example, and what they know about statistics in separate mental compartments. We want to produce people who are accustomed to bringing everything they know to bear on the problem at hand. When it comes to generating subject-matter questions, we are not concerned with the quality of the questions the students ask; that depends on their knowledge and background, and we are not in the business of training subject-matter researchers. The instructions we give to the markers of such assignment questions usually consists of "give credit for anything that sounds halfway sensible." It is the exercise of thinking about the problem that is really important. Enough background, however, must be given (or be part of the general knowledge of most students) to enable students to ask questions. My hope is that if we force them to do it often enough, the habit will stick.

The context of a particular data set can also provide the opportunities to make students think about types of measurement. For example, one can ask what types of measurement should be made if the objectives of the study are subtly modified, or one can present several different types of closely related measurement and ask what question each type of measurement is best suited to addressing (e.g., traffic death rates per 100,000 of population, or per 1,000 cars, or per 1,000,000 kilometers traveled).

Last but by no means least, the context of the data is necessary in giving students practice in communicating the results of an analysis.

3.4 English In, English Out

I have often been horrified by the lack of skills in problem recognition that ex-students demonstrate when they come back and ask how to approach a set of data of a type that they have seen and analyzed several times before. One of the problems is simply that they have not seen enough examples. Some people require large numbers of examples before they begin to recognize patterns or structures. Another major problem seems to be the types of assignment problems we tend to set in skills-based courses. We ask for what we want done in too much step-by-step detail. I am sure that we are not alone in this.

The questions asked in real life tend to be asked at a much higher level and in nontechnical language. A statistician is not told what model to use, what variables to include, how to treat each variable, to test this, estimate that, and check for some particular diagnostic features. Statisticians must generate all of this for themselves from their own knowledge and from a dialogue with their clients. To get closer to this situation, even in Stat101, we should progress toward using much less explicit questions, which are couched in plain English, as soon as possible after introducing a new idea. The teaching materials should show students how to recognize data structures, and they should also contain strategies for approaching data structures. Students should be forced to go back to their textbooks and lecture notes for ideas on approaching problems and not have these ideas supplied with the problem itself.

Another unfortunate feature of real life is that the information one is provided with does not consist of exactly those salient pieces necessary to solve the problem. There is normally either too little information, in which case one has to extract the remainder in dialogue with a client or other sources, or there is too much information, most of it extraneous. That creates a huge gulf to be bridged by a student who has been accustomed to a very controlled environment. We need to give some experience that is intermediate between these extremes. We have not tried out strategies for giving too little information at Auckland, although we probably should. (Students tend to regard being given too little information to complete an exercise as conclusive proof of carelessness and incompetence on the part of the lecturer—perhaps because it usually is.) But we have tried several ways of providing too much information. Sometimes, we simply provide more data than the question requires. At the suggestion of one of our quality-team members, we sometimes provide entire newspaper and magazine articles. The questions then require students to read the story and abstract the relevant pieces of information needed to answer the questions. Practices such as keeping lots of context with data and using entire articles have meant that the reading-comprehension requirements of our courses are becoming quite demanding. Some students do not like this: "I took this course because I thought it would be math like at high school." In addition, it can make life considerably more difficult for foreign students who are not native English speakers. Because we believe the objectives are sufficiently important, however, we are prepared to put up with a little flack.

A careful and sophisticated analysis of the data is often quite useless if the statistician cannot communicate the essential features of the data to a client for whom statistics is an entirely foreign language. To perpare for this, students should *always* be made to give a summary of what they have found in nontechnical language. This completes the cycle. They should acquire the skills that enable them to turn verbal questions into verbal answers, hence the slogan "English in, English out."

Translating the results of an analysis into plain English is not an easy skill. We cannot simply fob students off with instructions to express things in their own words. In our textbooks and our teaching, we should make an attempt to give students some of the phrases that will help in doing it. There are some problems here, and the following provides an example.

Few statistically unsophisticated clients want to be forced to absorb a precise understanding of the meaning of a *p value* or of the exact nature of a confidence interval. That is the client's right; consider how long it takes undergraduate students to understand these things. Clients need only have an understanding deep enough to permit them to *use the results* of the analysis *for their own purposes*. In communicating the results of an analysis, as in everything else, the needs of the customer are paramount. Unfortunately, any imprecise phrases a textbook author suggests for translating *p values* and confidence intervals for the layperson are inevitably wrong. That author is almost certain to be jumped on by colleagues and reviewers, and perhaps this is one of the reasons we do not do it. But what is the alternative? Either express in your own words or total silence. Either way, we end up putting the onus of communication onto people who are less experienced than ourselves and who are even less likely to make a reasonable job of it.

In communicating ideas to the layperson, we have to accept the fact that we will often not be able to communicate that idea perfectly but may have to settle for communicating the flavor of the idea. In addition to useful phrases, we must prepare our students for a dialogue in which the client wants to probe more deeply into the meaning of the phrase, a little like peeling an onion at the center of which is the exact nature of the idea. For example, in my book with Seber every confidence interval is re-expressed verbally with a phrase such as "with 95% confidence, the risk of death for smokers is somewhere between twice and four times the risk for nonsmokers." The "with 95% confidence" conveys some feeling of certainty, and many people would not dig beyond it. The student, however, also has to be prepared to answer the person who says, "What do you mean by 'with 95% confidence'?" "Well, I used a method that correctly brackets the true increase in risk for 95% of studies like this one, but I can't be sure it gave the correct answer for this study." "I'm not interested in any other studies, what does this one say?. . ."

Another change we have been making over recent years is to examine discursive material and verbal discussions of concepts in assignments and examinations. We do it in assignments (e.g., "In plain English, try to explain to a nonstatistician. . ..") to give the students practice in actually trying to express themselves (see also Radke-Sharp 1991). We include it in final examinations in part to send a strong signal to students that this is important, that calculations are not the only things that really matter if the students want to pass this course.

Most of the ideas in this section have been things one can do with large classes when one is woefully understaffed (as we are). In better-staffed environments, many of the aforementioned issues can be attacked by using project work, oral examinations, and small-group seminars. But these approaches are still a complement to rather than a replacement for the ideas presented here. Establishing habits and communication skills is not something one can do with one or two projects or seminars. The element that would be missing is constant, habit-forming repetition.

[Received August 1992. Revised August 1993.]

REFERENCES

Bailar, B. A. (1988), "Statistical Practice and Research: The Essential Interactions," *Journal of the American Statistical Association*, 83, 1–8.

Boroto, D. R., and Zahn, D. A. (1989), "Promoting Statistics: On Becoming Valued and Utilized," *The American Statistician*, 43, 71–72.

Bradley, R. A. (1982), "The Future of Statistics as a Discipline," *Journal of the American Statistical Association*, 77, 1–10.

Brook, R. J., Arnold, G. C., Pringle, R. M., and Hassard, T. M. (eds.) (1986), *The Fascination of Statistics*, New York: Marcel Dekker.

Cobb, G. W. (1987), "Introductory Textbooks: A Framework for Evaluation," *Journal of the American Statistical Association*, 82, 321–339.

——— "Teaching Statistics: More Data, Less Lecturing," *Amstat News*, Dec. 1991, p. 1.

Cockerill, R., and Fried, B. (1991), "Increasing Public Awareness of Statistics as a Science and Profession—Reinforcing the Message in Universities, " *The American Statistician*, 45, 174–178.

COMAP (Consortium For Mathematics and Its Applications) (1989), Against All Odds: Inside Statistics (Video), Santa Barbara, CA: Annenberg/CPB collection.

Dallal, G. E. (1990), "Statistical Computing Packages: Dare We Abandon Their Teaching to Others?" *The American Statistician*, 44, 265–266.

Deming, W. E. (1986), *Out of the Crisis*, Cambridge, MA: MIT Center for Advanced Engineering Study.

Efron, B., and Tibshirani, R. (1991), "Statistical Data Analysis in the Computer Age," *Science*, 253, 390–395.

Ehrenberg, A. S. C. (1978), *Data Reduction: Analysing and Interpreting Statistical Data* (rev. ed.), New York: John Wiley.

——— (1981), "The Problem of Numeracy," *The American Statistician*, 35, 67–71.

Kosslyn, S. M. (1985), "Graphics and Human Information Processing," *Journal of the American Statistical Association*, 80, 499–512.

Marquardt, D. W. (1987), "The Importance of Statisticians," *Journal of the American Statistical Association*, 82, 1–7.

Minton, P. D. (1983), "The Visibility of Statistics as a Discipline," *The American Statistician*, 37, 284–289.

Moore, D. A. (1990), "Uncertainty," in *On the Shoulders of Giants: A New Approach to Quantitative Literacy*, ed. L. A. Steen, Washington, DC: National Academy of Sciences, pp. 95–137.

Mosteller, F. (1988), "Broadening the Scope of Statistics and Statistical Education," *The American Statistician*, 42, 93–99.

Radke-Sharp, N. (1991), "Writing as a Component of Statistics Education," *The American Statistician*, 45, 292–293.

Roberts, H. V. (1987), "Data Analysis for Managers," *The American Statistician*, 41, 270–278.

Senge, P. M. (1990), *The Fifth Discipline: The Art and Practice of the Learning Organisation*, New York: Doubleday.

Snee, R.D. (1988), "Mathematics is Only One Tool That Statisticians Use. . .," *College Mathematics Journal*, 19, 30–32.

Tanur, J. M., Mosteller, F., Kruskal, W. H., Link, R. F., Pieters, R. S., Rising, G. R. and Lehmann, E. L. (eds.) (1987), *Statistics: A Guide to the Unknown* (2nd ed.), Monterey, CA: Wadsworth & Brooks/Cole.