Training Statisticians to be Consultants

John H Schuenemeyer - Newark, Delaware, USA

1. Introduction

Training of students to be consultants should begin with two premises:

(i) statistics is about data; and
(ii) every statistician is a consultant.

These premises set the tone for the training of statisticians. The need to gain insight into data should drive the understanding and development of statistics. A statistician should be a problem solver, a numerical Sherlock Holmes, with the desire to work with others in a cooperative spirit. Statistics provides the framework for the difficult task of making decisions when there is uncertainty. Through statistics we have a marvellous opportunity to serve, to learn about our world, and to educate.

In 1979 a report was prepared by the American Statistical Association's Committee on Training of Statisticians for Industry (Snee et al., 1980). It listed the skills needed by industrial consultants, including a broad knowledge of statistics, the ability to listen, to ask questions, to work on multiple projects, to meet deadlines, and to communicate effectively. Two important themes of that report are the need for students to work on real problems and for faculty to have a positive attitude to this work.

A few universities have had long-standing programmes to train students in statistical consulting. Many more programmes have been initiated in recent years. These include formal courses in statistical consulting, client problem seminars, internships with government and industry, and consulting in university based statistical laboratories. It is important to recognise that the training of statistical consultants depends upon an atmosphere created by the faculty in all courses, and on other faculty-student interactions. My comments concern the training of statistics' graduate students.

2. Statistical consulting process

Before reviewing specific training programmes for statistical consultants, we need to understand the consulting process, including the tasks consultants perform. As

part of her doctoral dissertation, Sarah Tung (1989), and Tung and Schuenemeyer (1990), investigated the consulting process. It is no surprise that this process is complicated with many feedback loops, false starts, and other detours. Some of the factors which contribute to this complexity are:

(i) difficulty in establishing the client's wants and needs;
(ii) training and experience of the consultant;
(iii) individual style of the consultant;
(iv) most problems do not have a unique solution;
(v) knowledge of the client;
(vi) resources available to client and consultant; and
(vii) the state of the problem (no design, data collected, etc.).

During our research, we (Tung and Schuenemeyer) interviewed six experienced consulting statisticians from industry, government, and academia to determine what they do and how they do it. Summaries of some of these videotaped interviews are presented in Tung (1989). Good consultants used the following strategies:

(i) Be interested in the client and his or her problem. Listen! Put the client at ease. Find out what the client wants. (Occasionally it may be a paper, a dissertation, or a satisfied boss as opposed to a good design or analysis.) You may not be able or wish to satisfy these wants but understanding them will facilitate communication. Ascertain the client's background.
(ii) Gain insight into the problem area. Draw pictures. Restate the problem in your own words. Use, insofar as possible, the client's terminology. Avoid statistical jargon.
(iii) Know when to say no. Know your limitations and determine those of the client's experiment. A corollary to this is, "only promise what you are prepared to deliver".
(iv) Develop a check list.
(v) Identify the experimental unit. Ask the client what he or she did with a single subject.
(vi) Encourage the client to plot the data.
(vii) Propose the simplest model or analysis possible, consistent with the complexity of the problem and other restrictions. The consultant should ensure that the client's objectives are consistent with the experiment.
(viii) Attempt to make the design robust. Be concerned about (possible) missing data and departures from assumptions.
(ix) Summarise the results of the meeting. Try to ensure that the client understands your advice. This is difficult to do in a non-threatening way. Decide who will do what and when.

The interviewed consultants stated that helping the client to define the problem, identifying the experimental unit, and ensuring that the client understands the advice, are among the most difficult tasks.

3. Teaching statistical consulting

Having identified some of the important parts of the consulting process, we will examine the tools used to teach consulting and see how they match the components described by experienced consultants. For this discussion, I rely mostly upon my experience at Delaware but will mention other activities. Besides directing the Statistical Laboratory at Delaware and coordinating internship programmes, I have been instrumental in the development of what now is the American Statistical Association Subsection on Statistical Consulting Education. I have also consulted widely with government and industry on a variety of problems.

One must recognise that *statistics is something you DO*. In order to train statistical consultants, we need to have them do consulting in a supervised environment. Among the approaches to this are:

(i) *A Statistical Laboratory*: At the University of Delaware we have such an institution. It was begun in 1984 and consists of a 1/4 time director (in theory), and two advanced graduate students. Each graduate student spends 20 hours per week in the laboratory. We also have an advisory committee of about 15 faculty and staff. These are statisticians and others with statistical expertise. We serve graduate students, faculty, and staff within the university, and some outside clients. Funding comes from the Department of Mathematical Sciences (our home base), the Provost's Office, faculty with grants (a very small amount), and outside clients. Service to the university community is free except for those with grants who are expected to pay.

At the beginning of the semester, I meet with the student consultants to discuss consulting procedures and operating procedures of the laboratory. Students are given consulting check lists, client information forms, and articles and references to material on consulting. I recommend that they read Jim Boen and Doug Zahn's excellent book on the human side of statistical consulting (Boen and Zahn, 1982). In addition, I try to allay the fears that all new consultants have.

The following example illustrates the operation of our Statistical Laboratory. A chemical engineering graduate student calls the secretary for an appointment. If this is her first visit for a particular problem, we request a that a brief written problem statement be submitted three days prior to the appointment. (There is not always a high correlation between this document and the problem the client presents, but it helps.) We also request that the faculty advisor accompany a student client on the first visit which I try to attend. During this visit, normally scheduled for an hour, we try to understand the problem. If it is straightforward we may propose a solution. Otherwise, we schedule a second appointment and then investigate the problem. For the first several new client visits, I lead the discussion. After this I let the student take the lead and try to speak only when asked. The role of statistical consulting units has been discussed by Carter, Scheaffer, and Marks (1986).

We evaluate the consulting services in several ways. We give each client an evaluation form and study the completed forms. I also meet weekly (for about an hour) with the graduate student consultants to examine the status of cases, make sure there are no unusual time delays, and provide help as needed. We also videotape a few consulting sessions and review the tapes with the students. We have done this on a limited basis because of space and time constraints. A couple of times we reviewed tapes of consulting sessions with a psychologist in the Department of Psychology. He provided

some interesting insights into the process, but these consultations are difficult to arrange and time consuming. Doug Zahn (at Florida State University) has worked with Dan Boroto (a Psychologist at Florida State University) and has considerable experience in this area (Zahn and Boroto, 1981).

(ii) *Statistical Laboratory review sessions*: This is a Statistical Laboratory function but I list it separately because of its importance in training consultants. Each week the students or I select an interesting client problem for presentation at this informal seminar which normally lasts for one and a half hours. When possible, the client makes a short presentation. Otherwise, the graduate student consultant makes the presentation. The student consultant then leads the informal discussion which focuses upon solutions to the client's problem. Faculty and graduate students participate. One problem is getting new graduate students, who incorrectly feel faculty know all the answers, to participate.

This session has several benefits. Students (and faculty) learn what questions to ask and how to ask them. For example, in attempting to ascertain randomisation, a student might ask a client, "Have you randomised the experimental units?" The client, who may not know what the terms randomisation or the experimental unit mean will either answer with a bewildered look or in the affirmative (without knowing why). As we know, a better way to ask the question, after first identifying the experimental unit, would be, "Please tell me how you conducted the experiment?". After a few more non-technical sounding and non-threatening questions, the consultant should be in a position to assess the randomisation. In the course of this informal discussion, students learn how statisticians approach problems. They see statistics used in real life and gain insight into what compromises one can make when model assumptions are not met, and the proper procedures to use. Sometimes we interest other students in working with us on a client's problem. A few problems have been the basis or inspiration for masters theses.

(iii) *Statistical internship programme*: We are in our sixth year of an internship programme with the Quality Management and Technology Center, DuPont. (This center evolved from the Applied Statistics Group and is headed by Donald W Marquardt, who has been very supportive of our efforts.) We also have internship agreements with Hercules and the Delaware State Department of Public Instruction. The programme was begun for graduate students working on masters' degrees but many doctoral students in statistics now participate. Students normally spend one year at the university taking courses in mathematical and applied statistics including a quality control course. In the middle of their first year, students may apply for internship positions. A committee of full time and adjunct faculty review applications and the decision to accept a student is made jointly. The student begins his or her internship in June and works at the sponsoring agency until the following February (approximately 7 months). (Three months, the normal summer break, is too short a time period to benefit either the student or sponsor.) The student spends the next semester on campus writing a project report and masters' students complete course work. During the internship, the student will be assigned to one industrial (or government) statistician as his or her primary supervisor but will work with many statisticians on a variety of projects. The student will take internal courses, visit field sites, and participate in consultation sessions with clients. Many students have been able to take the lead in these sessions by the end of their internships. They prepare written reports and conclude their internship with a

formal oral presentation to management. This is an anxiety causing but beneficial experience.

Student interns learn:

(a) how statistics are applied in the real world;
(b) to work on multiple problems;
(c) to work under time constraints;
(d) different statistical styles;
(e) to communicate with non statisticians; and
(f) that statistics is an exciting discipline and their interest in learning is heightened.

I emphasise that the students in these programmes are not observers or gofers (people who do trivial tasks). They DO statistics under the supervision of practicing statisticians.

How to start an internship is the subject of a separate paper, however, many opportunities in industry and government exist, especially with world-wide concerns about quality and productivity and the environment. Also, to be successful it is important to create a win-win-win situation. That is, the student, the university, and the sponsoring agency must benefit.

(iv) *The adjunct faculty*: These faculty help to teach consulting in a variety of ways. For example, Dr S Mike Free, Jr, an independent statistical consultant, regularly participates in our Statistical Laboratory weekly review sessions and provides insights from his many years of industrial experience. Others, including Dr James Lucas, at DuPont, jointly supervise dissertations in design and quality control.

(v) *A course in statistical consulting*: This year we plan to begin a formal course in consulting but we already have many of the components in place. Most of our students gain consulting experience as Statistical Laboratory consultants or through internship programmes, however, our goal is to ensure that all students learn consulting. Typically, such courses involve observing and reporting on several actual consulting sessions, and being a consultant in a mock session. These sessions are often videotaped and critiqued by the class.

In some courses students (often working in teams) analyse and report on realistic problems. They learn what compromises can reasonably be made when model assumptions are not completely satisfied. Students need practice with realistic problems in a context-free environment, that is when an appropriate method is not suggested by the material being studied.

Students in successful courses develop problem solving and interpersonal skills. One example of an apparently successful consulting course is at The University of Reading, UK (Pike, Coe, and New, 1990). Other examples of successful courses are at Florida State (McCulloch et al., 1985) and Queen's University, Ontario (Smith, 1986).

(vi) *Consulting (or perhaps problem solving) as part of the total educational process*: I firmly believe that training consultants cannot be left to a single course or internship experience. Oral and written communication skills should be stressed in almost every class. Far too often these skills are woefully lacking. In the United States, we often find bright students who appear before a PhD qualifying examination committee and can barely recall their own names. Also, given the ready access to inexpensive and powerful

computing and quality software, there is no excuse for not introducing larger and more realistic data sets in courses. If this were done, students would gain a greater insight into the advantages, limitations, and assumptions required by statistical methods and would be better prepared to solve real world problems. Very often new consultants have no idea how to manage large messy data sets.

4. Critique

To conclude this presentation, I am going to discuss the results of a questionnaire (available from the author) that I gave to current graduate students with Statistical Laboratory consulting or internship experience and recent graduates who also had these experiences. Eight of eighteen current and former students responded. Many of the required skills mentioned by the respondents were expected, such as: a good knowledge of statistics, communication skills, and a good knowledge of statistical programming languages, like SAS and Minitab. Also mentioned were: the desire to be a good consultant, develop effective problem solving skills, and know one's limitations. Among the more difficult tasks that these consultants faced were:

(i) working on multiple problems;
(ii) working with large messy data sets;
(iii) understanding the robustness of model assumptions;
(iv) working with clients who think they know statistics;
(v) dealing with clients "who only want an answer";
(vi) dealing with impatient clients;
(vii) keeping up with new techniques;
(viii) helping the client define his or her problem; and
(ix) being constrained by lack of time or other resources.

The experiences students thought were beneficial included the Statistical Laboratory consulting, laboratory review sessions, internship programmes, working with other students and professionals on projects, and class projects which involved analysis of realistic data sets and written reports. Experience which helped to foster communication skills included internship programmes, Statistical Laboratory consulting experiences, and teaching assistantship duties (for oral communication). Many students did not feel their communication skills were developed sufficiently in the statistics programme. We need to do more in this area. Many of the students would like to see themselves videotaped and critiqued in mock and real consulting sessions. They would also like more opportunities to see experienced consultants at work.

I stated earlier the difficult problem of trying to ensure that the client understands the consultant's advice. One solution to this problem is to have the student consultant prepare a brief written report for the client.

Some of these new consultants mentioned the difficulty of keeping up with new techniques. We require students to read journals in their research area but we need to encourage students to read the general statistical literature.

More needs to be done but I believe the steps taken at Delaware and other universities in recent years contribute toward the training of more effective statisticians.

References

Boen, J and Zahn, D A (1982) *The Human Side of Statistical Consulting*. Lifetime Learning Publications, London.

Carter, R L , Scheaffer, R L and Marks, R G (1986) The role of consulting units in statistics departments. *American Statistician* **40**, 260-264.

McCulloch, C E, Boroto, D R, Meeter, D, Polland, R and Zahn, D A (1985) An expanded approach to educating statistical consultants. *American Statistician* **39**, 159-167.

Pike, D, Coe, R and New, L (1990) Teaching the skills of consultation, collaboration and communication. *The Statistical Consultant* **7**, 3-5. American Statistical Association, Subsection on Statistical Consulting Education.

Smith, J T (1986) First experiences with a statistical consulting course. *Queen's Mathematical Preprint #1986-12*. Queen's University at Kingston, Ontario, Canada.

Snee, R D et al. (1980) Preparing statisticians for careers in industry: Report of the ASA Section on Training of Statisticians for Industry. *American Statistician* **34**, 65-80.

Tung, S T Young (1989) *An expert system for ANOVA based on a study of the statistical consulting process*. PhD dissertation, University of Delaware, Newark, DE, USA.

Tung, S T Young, and Schuenemeyer, J H (1990) An expert system for statistical consulting. *Journal of Applied Statistics*. (Accepted for publication)

Zahn, D A and Boroto, D R (1981) An exercise for training statistical consultants. *Proceedings of the Section on Statistical Education, American Statistical Association*, 67-72.