The Practice of Statistics: The Real World is an Idea Whose Time Has Come

William G. Hunter

The American Statistician, Vol. 35, No. 2. (May, 1981), pp. 72-76.

Stable URL:
http://links.jstor.org/sici?sici=0003-1305%28198105%2935%3A2%3C72%3ATPOSTR%3E2.0.CO%3B2-W

The American Statistician is currently published by American Statistical Association.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/astata.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org
Fri Jan 12 14:11:17 2007

The Practice of Statistics: The Real World Is an Idea Whose Time Has Come

WILLIAM G. HUNTER*

This article identifies three consulting roles—helper, leader, and colleague; recommends that statistical consultants ask more questions about basic mechanisms; suggests that statistical consultants regard their primary responsibility as providing guidance about the scientific method itself; discusses the need for continuing education to help improve consulting skills; and describes Wisconsin's Master's degree examination, which was designed to help students become effective practicing statisticians.

KEY WORDS: Consulting; University education; Continuing education; Basic mechanisms; Master's degree examination.

1. INTRODUCTION

Statistical education in the United States has traditionally stressed theory. Emphasis has been given to deductive reasoning, and students have learned to value, for example, the beauty of a proof. In recent years, however, growing attention has been focused on the practice of statistics. What is the nature of statisticians' work? When hiring statisticians, what do employers look for? What can be done by universities to help students become effective practicing statisticians? In some other professions extensive continuing education programs are available, even mandatory. Do statisticians need a wider range of continuing education programs?

The training of doctors and other professionals is certainly concerned with theory, but it does not ignore practice. The activity, not the degree, defines the profession. The activity that defines a statistician is working on problems concerning the collection and analysis of data, either as a consultant or a technician.

2. THREE ROLES PLAYED BY CONSULTANTS

Consulting can be usefully discussed in terms of Figure 1. It shows three roles the statistician can play: helper, leader, and colleague. Let us consider each of these in turn.

2.1. Statistician As Helper

In this situation, an active client seeks statistical assistance from a passive statistician. Without becoming involved to any important degree in the subject matter itself, the statistician attempts to do what the client requests. In this helper role, the statistician typically asks few questions and tries to get on with the job, as defined by the client, as quickly as possible. (In the extreme, the client may even prescribe the statistical technique to be used, but then the statistician is a technician rather than a consultant.)

Sometimes consultants willingly accept this role, which is appropriate in certain circumstances. More often, however, because of lack of time, they reluctantly acquiesce. In any event, I believe that statistical consultants should generally try to move to the right in Figure 1.

2.2. Statistician As Leader

In the leader role the statistician is active, the client passive. This role is sometimes inappropriately thrust upon the statistician when the client, offering up a complex set of data, asks the statistician to figure out what it all means. In abdicating any responsibility for the analysis, however, the client is misguided because analyzing a body of data means trying to extract all the useful information from it. Since the client will typically know much more about the data than the statistician, a better, more insightful, analysis will be possible if the client is actively involved.

Marquardt (1979) mentions two leader roles, one "bad" and one "good." He states that failure is almost guaranteed if the statistician moves into the role of the executive running the show, and he points out that sometimes the statistician can profitably take the initiative in promoting the use of statistical methods in new areas. In undertaking such a project, one places considerable demands on oneself because what is required is the expert blending of many technical and psychological ingredients that include a knowledge of statistics and an understanding of people. Varying amounts of tact, persistence, patience, force, and logic may be needed at specific times. With respect to such projects, I have two pieces of advice. (a) Realize that in some situations a direct approach may not be the best way to achieve change. Ask yourself if it is important that you get recognition for the deed, or is it more important that you get the personal satisfaction

* William G. Hunter is Professor, Department of Statistics and Engineering Experiment Station, University of Wisconsin, 1210 West Dayton Street, Madison, WI 53706. The author is grateful to many friends who shared their ideas on the practice of statistics, especially J.R. Boen, C. Daniel, G.J. Hahn, W.J. Hill, C.D. Hendrix, R. Hooke, B.L. Joiner, D.W. Marquardt, J.R. Rosenblatt, and R.D. Snee.

of seeing that the deed is done, even if others get the credit. (b) Don't go overboard in proceeding by indirection, because what you intend as diplomacy may be perceived by others as manipulation. Of course, points (a) and (b) cut against each other; it is a matter of judgment just how they should be balanced in a given campaign. The more sensitive you are to the basic gratifications you seek from the client and vice versa, the more informed your judgment will be and the more successful you will be as a practicing statistician. (This point applies not only to the "selling" role we are now discussing, but to all work of a consulting statistician.)

Although this role of statistical salesperson is rarely studied, it is one that statisticians are often asked to play. It is more obviously part of the job for the first statistician in the first years at a new location, but it is also part of the job of many other statisticians. For this reason, there should be more discussion and serious study given to the relative success of various "selling" techniques.

2.3. Statistician As Colleague

In terms of Figure 1, nonroutine projects are most successful when both the client and the statistician contribute actively and, in particular, when the statistician is a fully functioning colleague. Hooke (1980), for example, describes the statistician's consummate role as that of giving advice about scientific method itself, rather than assisting with mundane technical chores such as checking calculations for an analysis of variance (see also Marquardt 1979 and Mosteller 1979).

The two roles of helper and leader are characterized by one-way communication—the helper receives, the leader transmits. The role of the colleague, on the other hand, necessarily involves two-way communication, and hence makes possible what, for me, is the joy of being a statistician: working on and learning about many different problems, and sharing with clients the excitement of solving these problems.

W.J. Hill has suggested that a third dimension might be added to Figure 1 to bring in the passive and active aspects of what Daniel (1969) has called the organizational situation in which the work is done. One can then further refine the roles identified above. Perhaps for the passive organization the four roles could be called nonfunctionary, crusader, helper, and colleague (proceeding in standard Yates order); and, for the active organization, the corresponding four roles could be called teacher, leader, data blesser, and collaborator.

3. THE CONSULTING PROCESS

Often the problem as initially posed by the client is not the real one; for example, crucial elements of the problem may not be volunteered. Therefore, it is of utmost importance that in each new situation the consultant try to discover what the real problem is. To

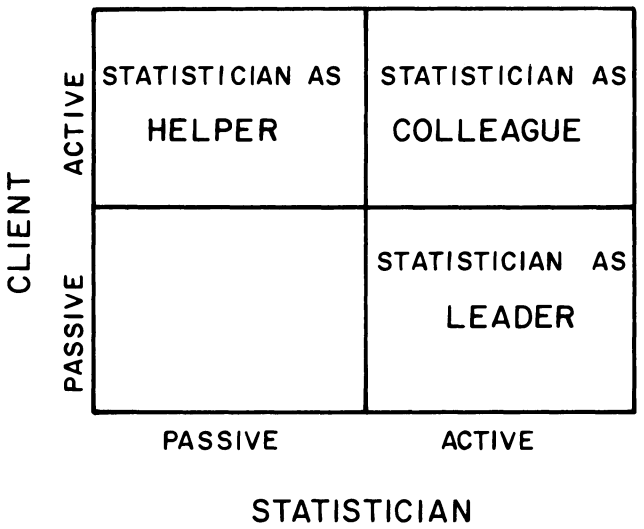

Figure 1. Three Roles of a Statistical Consultant: Helper, Leader, and Colleague

avoid the mistake of solving the wrong problem, thorough probing is essential (see Kimball 1957 for a discussion of errors of the third kind). Be curious. Ask lots of questions. If the client indicates, verbally or otherwise, that you are treading on sensitive ground, proceed with due regard for the client's feelings.

At the outset the most important question for the statistician to ask is: What is the objective of this investigation? I remember asking that question of two investigators who had been working together for some time on a particular project. After realizing that they disagreed on the answer, they started to thrash the whole thing out. A lively 45-minute discussion ensued. I listened to this discussion, but did not participate in it. When it ended, they agreed on what it was they were about. They thereupon said that I had been most helpful, and we said goodbye. Frequently the objectives of an investigation gradually change over time. Therefore, this aspect of a project always needs to be monitored.

Other useful questions to ask are: How were these data collected? In what order? On what days? By whom? How? How does the equipment work? What does it look like? May I see it work? Do any other relevant data exist? How much theory is known about the phenomenon being studied? To ensure that the consultant understands the problem—whether it be one of design or analysis—it is useful for the statistician to explain how she or he perceives the problem, asking the client to correct any misconceptions (Box, Hunter, and Hunter 1978, p. 14). The essential purpose of asking such questions was succinctly stated by Fisher (1934–1935, p. 13) as follows:

> It is a statistical commonplace that the interpretation of a body of data requires a knowledge of how it is obtained. Equally, it is usually understood that the conclusions drawn from experimental results must rest on a detailed knowledge of the experimental procedure actually employed.

Figure 2 depicts statistical consulting as a five-step process similar to that used in solving problems in

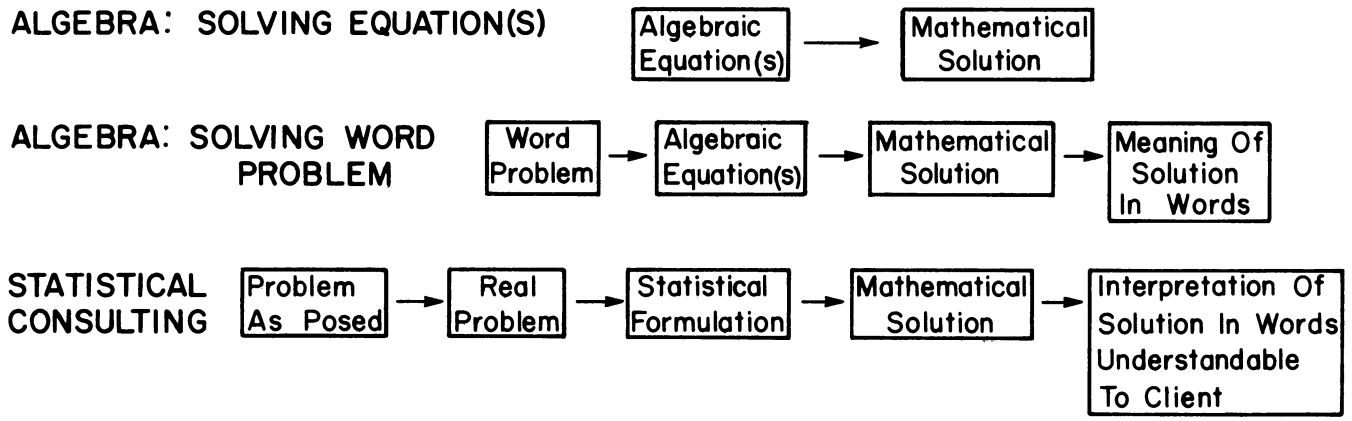

Figure 2. Statistical Consulting As a Five-Step Process: Analogy to Solving Word Problems in Algebra

algebra. Although this figure captures many aspects of statistical consulting, it misses some. For example, the main jobs of statisticians, deciding how best to collect the necessary data for a particular study and analyzing the data once they become available, are both iterative processes. Thus Figure 2 should not be interpreted to mean that statistical consulting is a one-shot process, with no looping, backtracking, or trial and error. Nor should it be interpreted to mean that problems always originate with the client. As pointed out by Marquardt (1979), the statistician may sometimes take the initiative in identifying a worthy project and convincing the client to undertake it.

4. WHAT ARE EMPLOYERS LOOKING FOR?

In helping prepare the report by Snee et al. (1980), I telephoned 11 individuals, primarily from industry, to ask what criteria would be most important if they were hiring a statistician. In somewhat abbreviated form, but in approximately their own words, here are the 11 responses they gave:

1. Experience dealing with real live problems is much needed. The textbook stuff that people get is generally all right but if people have experience with real problems, we find they can contribute much sooner.
2. We want someone who is energetic and willing to work. Likes working with people. Willing to get more training and, if necessary, to learn things out of textbooks. Statistics is a bridge technology and by itself does nothing; therefore, the person should be able to pull things together and not to be too narrow.
3. Some exposure to applications in addition to theory, for example, real examples of surveys that illustrate special problems relating to local conditions and other constraints.
4. Experience in one physical or engineering science. Personable. The candidate should not regard locking the door and writing papers as his or her top priority. Gift for finding out what a problem is and translating it into statistics. Self-confidence to use simple and straightforward techniques. Flair for graphical techniques.
5. Knowledge of a physical science, for example, chemistry or engineering. Ability to consult (ask questions, listen, observe, be aware of interpersonal transactions). Communication skills. Ability to work effectively on teams.
6. Personality is most important. Candidate should be a self-starter, be able to work independently without much supervision, and have basic statistical skills with a little experience, perhaps in a statistical laboratory. After all, we want somebody to *do* something.
7. Communicating and relating with people are more important than knowing lots of math. Can think statistics but does not talk statistics; doesn't talk down to nonstatisticians.
8. Has writing skills. Some potential for eventually writing self-contained reports. Ability to coordinate three or four large projects; to guide experimenters and to listen to them. Can come up with plans, for example, to monitor studies that are done piecemeal.
9. Good theoretical background in statistics—basic knowledge. Has an understanding of experimental science, at least a minor. Can see the scientific problem rather than the mathematical one. Open-minded. Communication skills with nonstatisticians are very important. Interest in hearing what other person has to say, not in jamming the problem into a set mold.
10. Has a degree in physics, chemistry, or chemical engineering. Appropriate experience or Master's degrees in statistics. Hands-on experience with project work.
11. Willing to get hands dirty. Hands-on experience with data and projects. Able to get along with people. Technical background and training from a good school. Ability in consulting, in particular, in asking the right questions and drawing people out. Willingness to do a lot of things that are not statistics.

5. PREPARATION OF PRACTICING STATISTICIANS

Given that employers, when evaluating statisticians, weigh many factors that have nothing to do with mathematics, what can universities do to help prepare students become effective practicing statisticians? The report by Snee et al. (1980) discusses the prep-

aration of students for careers in industry. Most of its recommendations also apply to careers in government and academia.

The Master's degree program at Wisconsin culminates in an examination in which candidates are given challenging problems similar to those that may later be encountered on the job. (These problems have involved, for example, the safety of nuclear power plants, the incidence of cancer, the growth of corn, the rate of chemical reactions, and the effectiveness of teaching methods. They have been somewhat messy and nonstandard.) One week later, written reports are submitted. Individual oral examinations follow in which candidates are questioned about their reports. To pass, the candidate must not only produce sensible solutions to the problems but also convince the examining committee that he or she can satisfactorily explain statistical ideas orally and in writing to nonstatisticians. Effective practicing statisticians, of course, must have a firm grasp of basic theory; the oral examination also provides an opportunity to assess this aspect of a candidate's abilities. G.E.P. Box, J.H. Klotz, and I proposed this form of examination in September 1971. Our department has used it since then with favorable results.

The examining committee once expressed disappointment that the students did not make a serious attempt to learn about the technical background of the problems they attempted to solve. Specifically, one problem concerned three different methods of measuring peak areas under curves produced by gas chromatographs; these were identified in the problem statement as "general, triangulation, and planimetric" methods. In the oral examination students were unable to explain, even in the most cursory way, what these methods were. It is true that in the problem statement the students were not told to learn about these methods (just as they will almost surely not be given similar instructions when they are confronted with problems to solve after graduation). Nevertheless, such self-education is an integral part of being an effective problem solver. (See Marquardt 1979 for some words of advice along these lines, and Moore and Stuttard 1979 for an amusing illustration of the kind of embarrassing mistake that can be made when basic subject matter information is overlooked. For further discussion of the preparation of practicing statisticians and the nature of the field, see Boen and Fryd 1978, Box 1976, 1979, and Hunter 1981.)

6. SCIENTIFIC LOGIC AND BASIC MECHANISMS

In an excellent article, Hooke (1980) has argued that scientific logic should be the business of statisticians. I agree with him. Let us briefly consider some aspects of this idea. There may be important information external to the particular data being analyzed that will have a profound influence on the final conclusions reached by their clients. This external information may be extensive and vaguely incorporated into what is called experience or intuition, and it may be nonnumeric. These persons may not explain voluntarily (nor, at times, will they even be able to explain if asked) exactly how this external information helps shape their final conclusions or exactly what is included in this external information. Nevertheless, the incorporation of such information can be legitimate and valuable, even though it is not readily encompassed in one of our schemes of statistical inference (see Kruskal 1968 and 1979).

Successful statistical consultants help their clients create useful models (hypotheses, theories) and isolate pivotal issues, the resolution of which will lead to further advances. They ask questions such as these: What mechanism is operating here? Can you think of a plausible reason why the results came out as they did? Can you think of alternative plausible (or semiplausible) reasons? Such questions help clarify the thinking of both the client and the statistician. Statisticians should be continually learning, especially in fields in which they regularly consult, and failing to ask questions about the theory underlying the data is to miss an opportunity to learn. With deeper understanding, a statistician may be able to suggest more effective ways to collect and analyze data. Furthermore, in failing to ask questions about basic mechanisms, statisticians miss some of the fun that is easily within their grasp. It is often fascinating to hear investigators explain or speculate about the detailed processes that govern the behavior of the biological or physical system or the human enterprise that is being studied. Obviously, too much questioning may waste time, but, as I see it, statisticians err in the opposite direction. They do not ask enough questions of this kind. Their training is a handicap in this regard because it is devoid of mechanisms. The stuff of statistical examples in lectures, texts, and articles is almost exclusively of the simple empirical variety (linear models). Consequently, statisticians overlook basic mechanisms and underestimate how much more incisive and effective they could be if they had a deeper understanding of what is really going on. I firmly believe that clients are well served by statisticians who have a healthy curiosity about underlying mechanisms.

7. CONCLUSION

In a talk I gave at one of the seminars organized by the graduate students at the University of Wisconsin, I mentioned the acquisition of interpersonal communication skills for effective consulting, but the students were not interested in exploring this topic. Although students at Wisconsin are required to take a one-semester course in consulting, and are encouraged to get further consulting experience, they seemed to say that spending time to learn more about the deeper psychological aspects of consulting at this stage in their careers would be akin to reading Dr. Spock before the baby arrives; it just would not be the best time to

learn such things. One experienced statistician commented on this point as follows:

> The students may be entirely right. On the other hand, their response may simply mean that their academic experience has not managed to develop their sensitivities in the subject area. After all, insensitivity in this area is a pervasive theme of all technical academic culture everywhere.

But practicing statisticians, who spend an important part of their professional lives in consultant-client relationships, have not shown much interest in improving their consulting skills either. This fact is particularly puzzling to me. Why isn't there more demand for programs in this vital area? Naturally there is the question of how much can be taught in a formal way. As one statistician wrote to me,

> I confess to some skepticism about the possibility of teaching people to have a genuine care for the other fellow's problem: The statisticians who have it are quickly recognized by clients; those who don't invariably give themselves away sooner or later. But aspiring practitioners may be helped to learn successful attitudes if the profession continually makes known its support for collaborative work in the application of statistical methods.

In summary, here are a few key things I now believe about statistical consulting:

1. Statistical consulting is an important part of most of our graduates' jobs.
2. Many consulting experiences are unsatisfactory to some extent because of interpersonal or communication problems.
3. Important interpersonal and communication skills can be identified and learned.
4. An introduction to the general ideas of good consulting can be presented to graduate students; these students can get some limited first-hand experience doing consulting work.
5. The time for some serious training is after they have been out in the trenches for awhile (two years?). If there is no formal training, simply swapping stories and other anecdotal information can be valuable.
6. Effective training in statistical consulting, which, to my knowledge, has been made available in very few organizations, does not consist of lectures alone, but rather involves considerable role playing and other forms of active participation.
7. Large numbers of practicing statisticians may be ready to participate in high quality, well-organized training programs of this kind.
8. Such a development could have a tremendous impact by making it possible for statisticians to serve much more effectively the needs of clients in particular and society in general.

In universities the best way for faculty members to teach students how to be good practicing statisticians is by example. As far as role models are concerned, it seems to have escaped recognition that Fisher spent much of his time and derived great satisfaction helping other scientists with their work. (See Box 1978, p. 245 for an illustration of how he generously gave of his time when a scientific problem was presented by a young unknown army sergeant.) It is a shame that almost all statistics graduate students now learn something about Fisher's theory but nothing about his practice. Generally speaking, universities should be giving more attention to the practice of statistics. In particular, students should be given more encouragement and opportunity to

- grapple with real data,
- listen to accounts of how practicing statisticians go about their work (including lectures by experienced consultants),
- observe practicing statisticians while they work,
- try their hand, under adequate supervision, helping clients solve live problems, and
- read about the practice of statistics.

Likewise, practicing statisticians should be given more encouragement and opportunity, through continuing education, to learn more theory and to improve their consulting skills.

[*Received March 1979. Revised September 1980.*]

REFERENCES

BOEN, J.R., and FRYD, D. (1978), "Six-State Transactional Analysis in Statistical Consultation," *The American Statistician*, 32, 58–60.

BOX, G.E.P. (1976), "Science and Statistics," *Journal of the American Statistical Association*, 71, 791–799.

——— (1979), "Some Problems of Statistics and Everyday Life," *Journal of the American Statistical Association*, 74, 1–4.

BOX, G.E.P., HUNTER, W.G., and HUNTER, J.S. (1978), *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*, New York: John Wiley.

BOX, J.F. (1978), *R.A. Fisher, The Life of a Scientist*, New York: John Wiley.

DANIEL, C. (1969), "Some General Remarks on Consulting in Statistics," *Technometrics*, 11, 241–245.

FISHER, R.A. (1934–1935), "The Effect of Methods of Ascertainment Upon the Estimation of Frequencies," *Annals of Eugenics*, 6, 13–25.

HOOKE, R. (1980), "Getting People to Use Statistics Properly," *The American Statistician*, 34, 39–42.

HUNTER, W.G. (1981), "Six Statistical Tales," *The Statistician*, in press.

KIMBALL, A.W. (1957), "Errors of the Third Kind in Statistical Consulting," *Journal of the American Statistical Association*, 57, 133–142.

KRUSKAL, W.H. (1978), "Statistics: The Field," in *The International Encyclopedia of Statistics*, New York: Macmillan.

——— (1979), Comment on "Field Experimentation in Weather Modification" by Braham, *Journal of the American Statistical Association*, 74, 84–86.

MARQUARDT, D.W. (1979), "Statistical Consulting in Industry," *The American Statistician*, 33, 102–107.

MOORE, B.R., and STUTTARD, S. (1979), "Dr. Guthrie and *Felix domesticus*, Or: Tripping Over the Cat," *Science*, 205, 1031–1033.

MOSTELLER, F. (1979), Comment on "Field Experimentation in Weather Modification" by Braham, *Journal of the American Statistical Association*, 74, 88–90.

SNEE, R.D., BOARDMAN, T.J., HAHN, G.J., HILL, W.J., HOCKING, R.R., HUNTER, W.G., LAWTON, W.H., OTT, R.L., and STRAWDERMAN, W.E. (1980), "Preparing Statisticians for Careers in Industry," *The American Statistician*, 34, 65–75.