PRINCIPLES OF PROFESSIONAL STATISTICAL PRACTICE

By W. Edwards Deming

Consultant in Statistical Surveys
4924 Butterworth Place
Washington ~~16~~
20016

Reprinted from The Annals of Mathematical Statistics
Vol. 36, No. 6, December 1965, pp. 1883–1900
Printed in U.S.A.

Reprinted from THE ANNALS OF MATHEMATICAL STATISTICS
Vol. 36, No. 6, December 1965
Printed in U.S.A.

SPECIAL PAPER

PRINCIPLES OF PROFESSIONAL STATISTICAL PRACTICE[1]

BY W. EDWARDS DEMING

Washington

I. PURPOSE AND SCOPE

1. Purpose. Application of statistical theory has found wide acceptance and has made notable contributions in agriculture, medicine, industry, accounting, administration, and consumer research. It has made equal contributions to the natural and social sciences. Sampling and survey-design, along with statistical theory of response and non-response, form the structural framework of government statistical series of divers types, and of censuses of population, agriculture, and commerce.

Statistical theory grows year by year more difficult and abstract, more and more a specialism, more and more powerful in application. Statistical theory is transferable. The specialist in statistical methods may find himself applying the same basic theory in a dozen different fields in a week, rotating through the same projects the next week. Or, he may work day after day primarily in a single substantive field.

Either way, the statistician requires certain principles of practice for effective use of statistical knowledge. Knowledge of statistical theory is necessary but not sufficient. Statistical theory does not provide a road-map toward effective use of itself. The purpose of this paper is to propose some principles of practice, and to explain their meaning in some of the situations that the statistician encounters.

The statistician has no magic touch by which he may come in at the stage of tabulation and make something of nothing. Neither will his advice, however wise in the early stages of a study, ensure successful execution and conclusion. Many a study, launched on the ways of elegant statistical design, later boggled in execution, ends up with results to which the theory of probability can contribute little.

Even though carried off with reasonable conformance to specifications, a study may fail through structural deficiencies in the method of investigation (questionnaire, type of test, technique of interviewing), to provide the information needed. The statistician may reduce the risk of this kind of failure by pointing out to his client[2] in the early stages of the study the nature of the contributions that he himself must put into it.

Received 15 April 1965.

[1] Special invited address at the meeting of the Institute of Mathematical Statistics, Boston, 28 August 1958.

[2] The word client will denote the expert or group of experts in a substantive field

The purpose here is not to describe any new statistical techniques, nor old ones. Neither is it to convince anyone that statistical theory is important, nor how to become a successful statistician. Instead, the aim is to find and illustrate principles of practice by which statisticians may make effective use of their knowledge. Statistical practice is mostly a collaborative venture between statistician and experts in subject-matter. Logical principles of statistical practice will give greater effectiveness to the high returns that come from creative imagination on the part of the expert in the subject-matter with a logical and critical mind, and with a spice of ingenuity coupled with an eye for the simple and the humdrum (to borrow words from Bradford Hill [16]).

Previous writings on this subject have helped to clarify the position of statistical inference in legal evidence, and have pointed to the need for principles of practice and for principles for presentation of data [1], [5], [6], [7], [8], [9], [11], [12], [13], [15], [21], [25], [27]. This paper, in recognition of the need, attempts to lay out a road-map to meet it.

The recommendation offered here is simple in principle, viz., to have logical rules concerning the division of responsibilities between the substantive and statistical aspects of a study, and to come to an understanding at the outset of an engagement, with everyone concerned in the study, on just what these principles mean in operation.

What is the logic behind a proper division of responsibilities? How important is this division? How does the statistician work with other experts and hold firm to the responsibilities that he prescribes?

In order to stay within the bounds of my own experience, the suggestions to be offered here will relate specifically to only a portion of theoretical statistics, namely, professional practice in sampling and in the design of experiments. The purpose of this paper will be served if these suggestions provide some guidance in areas in which they are less directly applicable, or even if they only stimulate interchange of ideas that will lead to further work on professional standards.

Incidentally, it seems to me that the salaried statistician should have the same responsibilities and principles of workmanship as the consulting statistician; and that a company has the same obligations either way to the statistician, whether he is on a salary or acts as a consultant. The same principles apply also to the statistician who works as a member of a research team.

Of course, in a small place, the same man must sometimes work both as statistician and as engineer or other expert. The rules of statistical practice nevertheless remain the same. One should, to be effective, depend on statistical theory as a guide to valid and efficient plans for sampling and experimentation, and for safe statistical inferences. Substantive knowledge (engineering, medicine, marketing, psychology, or other) is no substitute for statistical theory.

(medicine, engineering, production, marketing, psychology, agriculture, or other) who are responsible to the man that will pay the bill for the study.

2. What is a professional man? Professional practice stems from an expanding body of theory and from principles of application. A professional man aims at recognition and respect for his practice, not for himself alone, but for his colleagues as well. A professional man takes orders, in technical matters, from standards set by his professional colleagues as unseen judges; never from an administrative superior. His usefulness and his profession will suffer impairment if he yields to convenience or to opportunity. A professional man is one who feels an obligation to provide services that his client may never comprehend or appreciate.

A professional statistician will not follow methods that are indefensible, merely to please someone, nor support inferences based on such methods. He ranks his own name and profession as more important than convenient assent to interpretations not warranted by statistical theory. He can be a trusted and respected public servant.

His career as an expert witness will shatter in shipwreck if he indicates concern over which side of the case the results seem to favor. "As a statistician, I couldn't care less," is the right attitude in a legal case, or in any other report.

He will not deliberately attempt to carry out work that lies beyond his competence, but will, on request, recommend other specialists.

He will, if asked, render all possible assistance to another statistician in order to improve statistical service.

One further point. The practicing statistician of imagination and probity will meet problems that, for best results, require new theory, or even a new line of attack. He will too often not have time to arrive, either by himself or with help, at what he believes to be a better solution than the one that he proceeds to use. In such cases, the practitioner has an obligation to bring the problem to the community of his colleagues, as by publication of a letter to the editor, or by a more formal note or paper.

II. LOGICAL BASIS FOR DIVISION OF RESPONSIBILITIES

3. Limitations of statistical theory. The limitations of statistical theory serve as signposts to guide a logical division of responsibilities between statistician and client. We accordingly digress for a brief review of the power and the limitations of statistical theory.

We note first that statistical inferences (probabilities, tests of significance, estimates, confidence limits, fiducial limits, etc.), calculated by statistical theory from the results of a study will relate only to the material, product, people, business establishments, etc., that the frame was in effect drawn from, and only to the environment of the study, such as the method of investigation, the date, weather, rate of flow and levels of concentration of the components used in tests of a production-process, range of voltage or of other stress specified for the tests (as of electrical equipment).

Empirical investigation consists of observations on material of some kind.

The material may be people; it may be pigs, insects, physical material, or records of transactions. The aim may be enumerative, which leads to the theory for the sampling of finite populations. The aim may be analytic, the study of the causes that make the material what it is, and which will produce other material in the future. A definable lot of material may be divisible into identifiable sampling units. A list of these sampling units, with identification of each unit, constitutes a frame. In some physical and astronomical investigations, the sampling unit is a unit of time. We need not detour here to describe nests of frames for multi-stage sampling. The important point is that without a frame there can be neither a complete coverage of a designated lot of material, nor a sample of any designated part thereof. Stephan introduced the concept of the frame, but without giving it a name [26].

Objective statistical inferences in respect to the frame are the speciality of the statistician. In contrast, generalization to cover material not included in the frame, nor to ranges, conditions, and methods outside the scope of the experiment, however essential for application of the results of a study, are a matter of judgment, and will depend on knowledge of the subject-matter [10].

For example, the universe in a study of consumer behavior might be all the female homemakers in a certain region. The frame therefor might be census blocks, tracts, or other small districts, and the ultimate sampling unit might be a segment of area containing households. The study itself will of course reach only the people that can be found at home in the segments selected for the sample. The client, however, must reach generalizations and take action on his product or system of marketing with respect to all female homemakers, whether they be the kind that are almost always at home and easy to reach, or almost never at home and therefore in part omitted from the study. Moreover, the female homemakers that the client must take action on belong to the future, next year, and the next. The frame only permits study of the past.

For another example, the universe might be the production-process that will turn out next week's product. The frame for study might be part of last week's product. The universe might be traffic in future years, as in studies needed as a basis for estimating possible losses or gains as a result of a merger. The frame for this study might be records of last year's shipments.

Statistical theory alone could not decide, in a study of traffic that a railway is making, whether it would be important to show the movement of (for example) potatoes, separately from other agricultural products in the northwest part of the U. S.: only the man that must use the data can decide that one.

The frame for a comparison of two medical treatments might be patients or other people in Chicago with a specific ailment. A pathologist might, on his own judgment, without further studies, generalize the results to people that have this ailment anywhere in the world. Statistical theory provides no basis for such generalizations.

No knowledge of statistical theory, however profound, provides by itself a basis for deciding whether a proposed frame would be satisfactory for a study. For example, statistical theory would not tell us, in a study of consumer research, whether to include in the frame people that live in trailer-parks. The statistician places the responsibility for the answer to this question where it belongs; viz., with his client: are trailer-parks in your problem? Would they be in it if this were a complete census? If yes, would the cost of including them be worth while?

Statistical theory is indispensable in discussions with the client on what precision to expect in tabulation plans for certain areas or classes or on the question of how much difference the inclusion of (e.g.) people that live in trailer-parks could make in the results. The statistician may be helpful by just happening to know where to look for figures on the magnitude of a specific group of people (though familiarity with sources of information is not the statistician's primary obligation). Or, he may suggest an inexpensive study that will provide rough estimates of the magnitude and importance of questionable classes. He may know by experience that investigation of certain areas or classes encounters considerable difficulty; and he may, to be helpful, raise a question, for the client's decision, on the possibility of omitting them. Such activity, however helpful to the client, is not an assumption of his responsibilities.

Statistical theory will not of itself originate a substantive problem. Statistical theory can not generate ideas for a questionnaire, nor for a test of hardness, nor specify what would be acceptable colors of dishes or of a carpet; nor does statistical theory originate ways to teach field-workers or inspectors how to do their work properly, nor what characteristics of workmanship constitute a meaningful measure of performance. This is so in spite of the fact that statistical theory is essential for reliable comparisons of questionnaires and tests.

Definitions and questions to use in the Current Population Survey, carried out monthly by the Bureau of the Census, by which a man in the sample is declared and recorded as unemployed or under-employed are substantive problems. They belong to the experts in the labor force. In contrast, statistical design to test alternative choices of questions on employment and unemployment, and the design of the sample, and the design and interpretation of controls for supervision, are statistical [14].

Statistical theory takes no cognizance of substantive requirements. The role of the statistician tapers off as the results are extended, projected, or generalized beyond the frame.

4. Contributions of statistical theory to subject-matter. It is necessary, for statistical reliability of results, that the design of a survey or experiment fit into a theoretical model. Part of the statistician's job is to find a suitable model that he can manage, once the client has presented his case and explained why he needs statistical information and how he could use the results of a study. Statistical practice goes further than merely to try to find a suitable model (theory). Part of the job is to adjust the physical conditions to meet the model selected.

Randomness, for example, in sampling a frame, is not just a word to assume: it must be created by use of random numbers for the selection of sampling units. Recalls to find people at home, or tracing illegible or missing information back to source-documents, are specified so as to approach the prescribed probability of selection.

Statistical theory has in this way profoundly influenced the theory of knowledge. It has given form and direction to quantitative studies, by moulding them into the requirements of the theory of estimation, statistical test of hypothesis, and other techniques of inference. The aim is, of course, to yield results that have meaning in terms that man can understand.

Any statistician could offer endless illustrations. A few reminders will suffice here. The two risks that management must accept and specify numerically for the application of the theory of probability (queueing theory), introduced independently by Molina [19], [20], and Erlang [18] to minimize the net economic loss from too much equipment or service, or from too little, is an example. Use of the theory of probability, following Shewhart, to minimize the net economic loss from the risks of blaming a common cause of variation on to a specific possibility (operator, machine, or other local source) and the cost of the converse error, is appreciated the world over [22], [23], [24]. The purchaser's plan of acceptance-sampling forms the only meaningful language in which to communicate the average quality desired of a stream of product.

Such techniques are building a theory of management. The principle of minimax, along with use of biased estimates to minimize the mean-square-error, are examples of more recent contributions to knowledge. The theory of probability is making profound impact on the study of genetics. The average book-dollar of investment, in the evaluation of the current-cost-new of property and equipment, has made impact on accounting methods.

Statistical theory has contributed important refinements to methods of testing in many lines of work (industrial, physical, psychological) by requiring that an acceptable system of tests and criteria must exhibit randomness in the Shewhart sense. It is only with the aid of statistical theory that one may construct and measure reliably and economically the repeatability of a proposed test. In fact, the words repeatable, reproducible, precise, good, reliable, round, hard, etc., as characteristics of performance or of material, have no operationally verifiable meaning except in terms of statistical tests and criteria.

Further, the statistician is often effective in assisting the substantive expert to improve accepted methods of interviewing, testing, coding, and other operations. Tests properly designed will show how alternative test- or field-procedures really perform in practice, so that rational changes or choices may be possible. It sometimes happens, for example, that a time-honored or committee-honored method of investigation, when put to statistically designed tests, shows alarming inherent variances between instruments, between investigators, or even between days for the same investigator. It may show a trend, or heavy intraclass correlation between units, introduced by the investigators. Once detected, such sources of variation may then be corrected or diminished by new rules, which are of course the responsibility of the substantive expert.

Use of a random sample of x and y coordinates at which to test a surface (as of enamel) for control of a process for finishing surfaces, or for evaluation of deterioration, in place of a judgment selection of points or complete reliance on a general view of the whole surface, is a simple example of a contribution to test-methods.

Statistical techniques provide a safe supervisory tool to help to reduce variability in the performance of man and machine. The effectiveness of statistical controls, in all stages of a survey, for improving supervision, to achieve uniformity, to reduce errors, to gain speed, reduce costs, and to improve quality of performance in other ways are well known. A host of references could be given, but two will suffice here [3], [4].

5. Statistical theory as a basis for division of responsibilities. We may now, with this background, see where the statistician's responsibilities lie.

In the first place, his specialized knowledge of statistical theory enables him to see which parts of a problem belong to substantive knowledge (sociology, transportation, chemistry of a manufacturing process, law), and which parts are statistical. The responsibility as logician in a study falls to him by default, and he must accept it. As logician, he will do well to designate, in the planning stages, which decisions will belong to the statistician, and which to the substantive expert.

This matter of defining specifically one's area of responsibility is a unique problem faced by the statistician. Businessmen and lawyers who engage an expert on corporate finance, or an expert on steam power plants, know pretty well what such experts are able to do, and have an idea about how to work with them. It is different with statisticians. Many people are confused between the role of the theoretical statistician (one who uses theory to guide his practice) and the popular idea that a statistician is skilful in compiling tables about people or trade, or is one who prophesies the economic outlook for the coming year, and which way the stock market will drift. They may know little about the contributions that the statistician can make to a study, or how to work with him.

Allocation of responsibilities does not mean impervious compartments in which you do this and I'll do that. It means that there is a logical basis for allocation of responsibilities, and that it is necessary for everyone involved in a study to know in advance what he will be accountable for.

A clear statement of responsibilities will be a joy to the client's lawyer in a legal case, especially at the time of cross-examination. It will show the kind of question that the statistician is answerable for, and what belongs to the substantive experts.

Statisticians have something to learn about professional practice from law and medicine.

6. Assistance to the client to understand the relationship. The statistician must take the time and trouble in the early stages to understand the client's problem, and what it means to him. One way to go about this is to direct the client's thoughts toward possible uses of the results of a statistical study of the entire frame without consideration of sampling, which only confuses the client's thinking in the early stages. As groundwork, there will be, of course, consideration of possible frames. Then follows the question of how the client proposes to investigate any unit, however it be selected from the frame.

Once these matters are cleared, the statistician may outline one or more statistical plans and explain to the client in his own language what they mean in respect to possible choices of frame and environmental conditions, choices of sampling unit, skills, facilities and supervision required. The statistician will urge the client to foresee possible difficulties with definitions, or with the proposed method of investigation. He may illustrate with rough calculations possible levels of precision to be expected from one or more statistical plans that appear to be feasible, along with rudimentary examples of the kinds of tables and inferences that might be forthcoming. These early stages are often the hardest part of the job.

It is good practice in most work to make the point that prediction of the precision or protection that any statistical plan will deliver is dependent on assumptions concerning the numerical values of certain statistical characteristics of the sampling units in the frame, and dependent on the performance of investigators. These assumptions relate to conditions to be encountered; hence the precision predicted may differ somewhat from the actual precision delivered.

The aim in statistical design is to hold accuracy and precision to sensible levels, with an economic balance between all the possible uncertainties that afflict data—built-in deficiencies of definition, errors of response, non-response, sampling variation, difficulties in coding, errors in processing, difficulties of interpretation.

A general theory that would embrace an economic balance of all sources of uncertainty does not exist, or might be unwieldy if it did exist. Hence, to strive toward an economic balance of all uncertainties, the statistician usually ascribes subjective weights to the importance of the various sources of uncertainty, and seeks to allocate effort toward reduction of each source, in accordance with these weights.

Professional statistical practice requires experience, maturity, fortitude, and patience, to protect the client against himself or against his duly appointed experts in subject-matter who may have in mind needless but costly tabulations in fine classes (5 year age-groups, small areas, fine mileage-brackets, fine gradations in voltage, etc.), or unattainable precision in differences between treatments, beyond the capacity of the skills and facilities available, or beyond the meaning inherent in the definitions.

These first steps, which depend heavily on guidance from the statistician, may lead to important modifications of the problem. Advance considerations of cost, and of the limitations of the inferences that appear to be possible, may even lead the client to abandon the study, at least until he feels more confident of his requirements. Protection of the client's bank-account, and deliverance from more serious losses from decisions based on inconclusive or misleading results, or from misinterpretation, is one of the statistician's greatest services.

Joint effort does not imply joint responsibility. Divided responsibility for a decision in a statistical survey is as bad as divided responsibility in any venture—it means that no one is responsible.

Although he acquires superficial knowledge of the subject-matter, the one thing that the statistician contributes to the problem, and which distinguishes him from other experts, is knowledge and ability in statistical theory.

III. SUMMARY STATEMENT OF RECIPROCAL OBLIGATIONS AND RESPONSIBILITIES

7. Responsibilities of the client. The client will assume responsibility for those aspects of the problem that are substantive. Specifically, he will stand the ultimate responsibility for:

a. the type of statistical information to be obtained;

b. the methods of test, examination, questionnaire, or interview, by which to elicit the information from any unit selected from the frame;

c. the decision on whether a proposed frame is satisfactory;

d. approval of the probability model proposed by the statistician (statistical procedures, scope and limitations of the statistical inferences that may be possible from the results);

e. the decision on the classes and areas of tabulation (as these depend on the uses that the client intends to make of the data); the approximate level of statistical precision or protection that would be desirable in view of the purpose of the investigation, skills and time available, and costs.

The client will make proper arrangements for:

f. the actual work of preparing the frame for sampling, such as serializing and identifying sampling units at the various stages;

g. the selection of the sample according to procedures that the statistician will prescribe, and the preparation of these units for investigation;

h. the actual investigation; the training for this work, and the supervision thereof;

i. the rules for coding; the coding itself;

j. the processing, tabulations and computations, following procedures of estimation that the sampling plans prescribe.

The client or his representative has the obligation to report at once any departure from instructions, to permit the statistician to make a decision between a fresh start or an unbiased adjustment. The client will keep a record of the actual performance.

8. Responsibilities of the statistician. The statistician owes an obligation to his own practice to forestall disappointment on the part of the client, who if he fails to understand at the start that he must exercise his own responsibilities in the planning stages, and in execution, may not realize in the end the fullest possibility of the investigation, or may discover too late that certain information that he had expected to get out of the study was not built into the design. The statistician's responsibility may be summarized as follows:

a. to formulate the client's problem in statistical terms (probability model), subject to approval of the client, so that a proposed statistical investigation may be helpful to the purpose;

b. to lay out a logical division of responsibilities for the client, and for the statistician, suitable to the investigation proposed;

c. to explain to the client the advantages and disadvantages of various frames and possible choices of sampling units, and of one or more feasible statistical plans of sampling or experimentation that seem to be feasible (*vide supra*, "Assistance to the client . . .");

d. to explain to the client, in connexion with the frame, that any objective inferences that one may draw by statistical theory from the results of an investigation can only refer to the material or system that the frame was drawn from, and only to the methods, levels, types, and ranges of stress presented for study. It is essential that the client understand the limitations of a proposed study, and of the statistical inferences to be drawn therefrom, so that he may have a chance to modify the content before it is too late;

e. to furnish statistical procedures for the investigation—selection, computation of estimates and standard errors, tests, audits, and controls as seem warranted for detection and evaluation of important possible departures from specifications, variances between investigators, non-response, and other persistent uncertainties not contained in the standard error; to follow the work and to know when to step in (*vide infra* "Necessity for the statistician to keep in touch");

f. to assist the client (on request) with statistical methods of supervision, to help him to improve uniformity of performance of investigators, gain speed, reduce errors, reduce costs, and to produce a better record of just what took place;

g. to prepare a report on the statistical reliability of the results (*vide infra*).

9. The statistician's report or testimony. The statistician's report or testimony will deal with the statistical reliability of the results. The usual content will cover the following points:

a. a statement to explain what aspects of the study his responsibility included, and what it excluded. It will delimit the scope of the report;

b. a description of the frame, the sampling unit, how defined and identified, the material covered, and a statement in respect to conditions of the survey or experiment that might throw light on the usefulness of the results;

c. a statement concerning the effect of any gap between the frame and the universe for important conclusions likely to be drawn from the results. (A good rule is that the statistician should have before him a rough draft of the client's proposed conclusions.)

d. evaluation of the margin of uncertainty, for a specified probability level, attributable to random errors of various kinds, including the uncertainty introduced by sampling, and by small independent random variations in judgment, instruments, coding, transcription, and other processing;

e. evaluation of the possible effects of other relevant sources of variation, examples being differences between investigators, between instruments, between days, between areas;

f. effect of persistent drift and conditioning of instruments and of investigators; changes in technique;

g. non-response and illegible or missing entries;

h. failure to select sampling units according to the procedure prescribed;

i. failure to reach and to cover sampling units that were designated in the sampling table;

j. inclusion of sampling units not designated for the sample but nevertheless covered and included in the results;

k. any other important slips and departures from the procedure prescribed;

l. comparisons with other studies, if any are relevant.

In summary, a statement of statistical reliability attempts to present to the reader a lower limit (or an upper limit) above which (or below which) he may assume rationally, with a stated risk, that the results of bigger samples would fall, if freed of important persistent operational blemishes. It should present any information that might help the reader to form his own opinion concerning the validity of conclusions likely to be drawn from the results.

Evaluation of the statistical reliability of a set of results is not mere calculation of standard errors and confidence limits. The statistician must go far beyond the statistical methods in textbooks. He must evaluate uncertainty in terms of possible uses of the data. Some of this writing is not statistical, but draws on assistance from the expert in the subject-matter.

The aim of a statistical report is to protect the client from seeing merely what he would like to see; to protect him from losses that could come from misuse of the results. A further aim is to forestall unwarranted claims of accuracy that the client's public might otherwise accept.

My own code of professional conduct specifies that the client shall not mention verbally or in print the statistician's participation in a study without approval. Any printed description of the statistical procedures that refer to this participation, or any evaluation of the statistical reliability of the results, must be prepared by the statistician as part of the engagement. If the client prints the statistician's report, he will print it in full.

In my opinion, the statistician should not recommend to the client that he take any specific administrative action or policy. Use of the results that come from a survey or experiment are entirely up to the client. The statistician, if he were to make recommendations for decision, would cease to be a statistician.

Actually, ways in which the results may throw light on foreseeable problems will be settled in advance, in the design, and there should be little need for the client or for anyone else to re-open a question. However, problems sometimes change character with time (as when a competitor of the client suddenly comes out with a new model), and open up considerations of statistical precision and significance of estimates that were not initially in view.

The statistician should reserve the right, I believe, to describe in a professional or scientific meeting the statistical methods that he develops in an engagement. He will not publish actual data or substantive results or other information about the client's business without his permission. In other words, the statistical methods belong to the statistician: the data to the client.

A statistician may at times perform a useful function by examining and reporting on a study in which he did not participate. A professional statistician will not write an opinion on another's procedures or inferences without adequate time for study and evaluation.

IV. SUPPLEMENTAL REMARKS

10. Necessity for the statistician to keep in touch. The statistician, when he enters into a relationship to participate in a study, accepts certain responsibilities. He does not merely draft instructions and wait to be called. The people whom he serves are not statisticians. They can not always know when they are in trouble. The statistician will ask questions and will probe on his own account with the help of proper statistical design, to discover for himself whether the work is proceeding according to the intent of the instructions. He must expect to revise the instructions a number of times in the early stages. He will be accessible by mail, telephone, telegraph, or in person, to answer questions and to listen to suggestions.

He may of course arrange consultations with other statisticians on questions of theory or procedure. He may engage another statistician to take over certain duties. He may employ other skills at suitable levels to carry out independent tests or re-investigation of certain units, to detect difficulties and departures from the prescribed procedure, or errors in transcription or calculation.

It must be firmly understood, however, that consultation or assistance is in no sense a partitioning of responsibility. The obligations of the statistician to his client, once entered into, may not be shared.

11. What is an engagement? Dangers of informal advice. It may seem at first thought that a statistician ought to be willing to give to the world informally and impromptu, if he so desires, the benefit of his knowledge and experience, without discussion or agreement concerning participation and relationships. Anyone who has received aid in sickness from a doctor of medicine who did his best without a chance to make a more thorough examination, or perhaps even by telephone, can appreciate how important the skills of a professional man can be, even under handicap.

On second thought, most statisticians can recall instances in which informal advice backfired. It is the same in any professional line. A statistician who tries to be a good fellow and give advice under adverse circumstances is in practice and has a client, whether he intended it so or not; and he will later on find himself accountable for the advice. It is important to take special precaution under these circumstances to state the basis of understanding for any statements or recommendations, and to make clear that other conditions and circumstances could well lead to different statements.

12. When do the statistician's obligations come to a close? The statistician should make it clear that his name may be identified with a study only so long as he is active in it and accountable for it, as by continued responsibility for evaluation of the statistical reliability of results. This obligation will require regular receipt of reports of statistical controls for the detection of non-sampling errors, and continual assistance on problems that come up, including procedures by which to add to the frame new areas or other new material, and for sampling these additions. A statistical procedure, contrary to popular parlance, is not installed. One may install new furniture, a new carpet, or a new dean, and they stay in place, but not a statistical procedure. It does not stay installed without constant vigil. Experience shows that a statistical procedure deteriorates rapidly when left completely to nonprofessional administration.

A statistician may draw up plans for a continuing study, such as for the annual inventory of materials in process in a group of manufacturing plants, or for a continuing national survey of consumers. He may nurse the job into running order, and conduct it through several performances. Experience shows, however, that if he steps out and leaves the work in nonstatistical hands, he will shortly find it to be unrecognizable. New people come on to the job. They know better than their predecessor how to do the work. Or, they may not be aware that there ever were any rules or instructions, and make up their own.

What is even worse, perhaps, is that people that have been on a job a number of years think that they know it so well that they can't go wrong. This type of fault will be observed, for example, in a national monthly sample in which households are to be re-visited a number of times: when left entirely to nonstatistical administration, it will develop faults. Some interviewers will put down their best guesses about the family, on the basis of the preceding month, without an actual interview. They will forget the exact wording of the question, or may think that they have something better. They will become lax about calling back on people not at home. Some of them may suppose that they are following literally the intent of the instructions, when in fact (as shown by a control), through misunderstanding, they are doing something wrong. Or, they may depart wilfully, thinking that they are thereby improving the design, on the supposition that the statistician did not really understand the circumstances. A common example is to substitute an average-looking sampling unit, when the sampling unit designated is obviously unusual in some respect [2].

In the weighing and testing of physical product, men will in all sincerity substitute their judgment for the use of random numbers. Administration at the top will fail to rotate areas in the manner specified. Such deterioration may be predicted with confidence unless the statistician specifies statistical controls that provide detective devices and feed-back.

13. The single consultation. In my own experience, I have found it wise to avoid a single consultation with a commercial concern unless satisfactory agenda are prepared in advance, and satisfactory arrangements made for absorbing advice offered. This requirement, along with an understanding that there will be a fee for a consultation, acts as a shield against a hapless conference which somebody calls in the hope that something may turn up. It also shows that the statistician, as a professional man, although eager to teach and explain statistical methods, is not on the lookout for chances to demonstrate what he himself might be able to accomplish.

Moreover, what may be intended as a single consultation often ends up with a request for a memorandum. This may be very difficult to write, especially in the absence of adequate time to study the problem. The precautions of informal advice apply here (*vide supra*).

14. The statistician's obligation in a muddle. Suppose that the plans for a study were properly formalized and rehearsed. The field-work is under way. Then the statistician discovers that the client or his duly appointed representatives have disregarded the instructions for the preparation of the frame, or for the selection of the sample, or that the field-work seems to be falling apart. "We went ahead and did so and so before your letter came, confident that you would approve," is a violation of relationship. That the statistician may approve the deed done does not mitigate the violation.

What should the statistician do in a muddle? It may still be possible to repeat some of the work, and thus to salvage the job. It is the client's decision on whether he can stand the expense or the delay. If little can be done to put the study back on the rails, the statistician is technically in the position of trying to do something with a non-probability sample (next section). Morally, he must ask himself how he can balance his responsibilities to his client and to his profession. If it appears to the statistician that there is no chance that the study will yield results that he could take professional responsibility for, then he must make this fact clear to his client, at a sufficiently high management-level. It is a good idea for the statistician to explain at the outset that such situations, while extreme, have been known.

The statistician should do all in his power to help the client to avoid such a catastrophe. There is always the possibility that the statistician himself may be partly to blame for not being sufficiently clear nor firm at the outset concerning his principles of participation, or for not being on hand at the right time to ask questions and to keep himself apprised of what is happening. Unfortunate circumstances may teach the statistician a lesson on participation.

15. Assistance in interpretation of non-probability samples. It may be humiliating, but statisticians must face the fact that many accepted laws of science have come from theory and experimentation without benefit of formal statistical design. Vaccination for prevention of smallpox is one; John Snow's discovery of the source of cholera in London is another [16]. So is the law $F = ma$ in physics; also Hooke's law, Boyle's law, Mendel's findings, Keppler's laws, Darwin's theory of evolution, the Stefan-Boltzmann law of radiation (first empirical, later established by physical theory). All this only means, as everyone knows, that there may well be information in a non-probability sample.

One possible way in which the statistician may render assistance in a non-probability sample is to make rough estimates of precision by pulling out of the data a few replicates, but this is no help in estimating the bias of selection, which is of course the main problem.

Another point is that a plea for assistance in the interpretation of the results of a survey that had not the benefit of statistical design presents a good chance for the statistician to explain how much more effective the study would have been had it started off properly, and had it been carried off in reasonable accord with the statistical specifications.

Perhaps the main contribution that the statistician can make is to advise the experimenter with a non-probability sample against conclusions based on meaningless statistical calculations. The expert in the subject-matter must take the responsibility for the effects of selectivity and confounding of treatments. The statistician may make a positive contribution by encouraging the expert in the subject-matter to draw whatever conclusions he believes to be warranted, but to do so over his own signature, and not to attribute his conclusions to statistical calculations, nor to the kind help of a certain statistician.

V. ACKNOWLEDGMENT

A paper on this subject is too important to entrust to one point of view. Copies of three previous versions went out about 4 years apart to about 50 statisticians for criticism and comment, and I have held innumerable conversations on the matter with statisticians, and with lawyers and other clients. Every line of this manuscript has its origin in experience. The principles expounded here met almost universal acceptance, but with suggestions on emphasis and clarity. There is no difficulty with the logic of division of responsibility. What is at times difficult to comprehend is the consequences of applying the logic. Comments that have come from friends kind enough to read earlier drafts may assist readers of the final version to overcome the same difficulties.

For example, one esteemed friend, head of a statistical organization in a university, said that as a statistician he frequently does the coding in studies that he carries out (surveys on health and sickness), and that the role that I have drawn for the statistician may be too restrictive. In further conversation, however, he agreed firmly with the point of this paper, that although any of us may at times take it on to himself to see that the coding is done (just as any of us may compute, type, draw up forms, or wash the cups after coffee), the fact remains that only a qualified expert in the substantive field should be responsible for the rules for coding, or for the substantive content of the investigation.

Another statistician, eminently distinguished, agreed with the views expressed here but said that he himself takes a deep interest in the questionnaire in a survey. Nothing in this paper precludes the statistician's interest in any aspect of a survey. Although the statistician is responsible for the statistical aspects of a study, he is under no obligation to proceed—in fact, has an obligation to call a halt—if he entertains serious doubts about the questionnaire or the method of investigation, or about the availability of the necessary skills and facilities for carrying out the work.

The only real note of dissent came from an esteemed colleague who in private conversation agreed with the entire content of the paper, and practices it, but had the feeling that a formal statement of responsibilities might alienate the substantive expert, and weaken the bond of dependence. Personally, I have observed the opposite: a careful statement of what is expected of the substantive expert, and what he may expect from the statistician, cements the relationship, with mutual respect.

Another friend agreed with the paper, but thought that the statistician should try to do more than merely use statistical techniques: he should, forsooth, use techniques appropriate to the problem. For example, as he said, in vacuum-zone melting, longitudinal and transverse layers of a sheet of material may exhibit widely different tensile strengths, and the statistician is not at liberty to select samples in one direction or the other willy nilly as takes his fancy.

One could go on and on. In a study of performance and usage of apparatus, the sampling unit must contain not only the main piece that sends or receives signals, but auxiliary apparatus as well. In a study of the psychological or social problems of families, it may not suffice to interview one member. The proper sampling unit may be the family, to complete the story on interaction and stresses. It may be necessary to interview other families in the same community.

Such remarks constitute good advice on how to improve statistical methods. However important, they would be out of scope here. Nothing in this paper relieves the statistician of his obligation to apply statistical methods appropriate to the problem, to the material to be investigated, and to the skills available.

It is a pleasure to mention in particular help from F. F. Stephan, William Kruskal, Frank Hoeber, John Tukey, Morris Hansen, William N. Hurwitz, Samuel Greenhouse, Churchill Eisenhart, Bradford F. Kimball, Nathan Keyfitz, Tore Dalenius, Josephine Cunningham, William G. Cochran, Harold Dodge, W. J. Youden, Bernard Greenberg, A. C. Rosander, Gerald J. Glasser, Leon Pritzker, Leslie Kish, Howard Jones, Edgar King, Richard Bingham, Charles Bicking, Cuthbert Daniel. The report of the Hotelling Committee on the teaching of statistics has provided guiding principles for practice as well as for teaching [17].

An unseen referee made a number of very helpful suggestions in the penultimate version, and suggested references. I express my thanks to him.

REFERENCES

[1] BROWN, THEODORE H. (1952). The statistician and his conscience. *Amer. Statist.* **6** No. 1 14–18.

[2] BUREAU OF THE CENSUS (1954). Measurement of employment and unemployment: Report of Special Advisory Committee on Employment Statistics. Washington.

[3] BUREAU OF THE CENSUS (1964). Evaluation and research program of the censuses of population and housing: 1960 Series ER60. No. 1.

[4] BUREAU OF THE CENSUS (1963). The current population survey and re-interview program. Technical paper No. 6.

[5] BURGESS, R. W. (1947). Do we need a "Bureau of Standards" for statistics? *J. Marketing* **11** 281–2.

[6] CHAMBERS, S. PAUL (1965). Statistics and intellectual integrity, *J. Roy. Statist. Soc. Ser. A* **128** 1–16.

[7] COURT, ANDREW T. (1952). Standards of statistical conduct in business and in government. *Amer. Statist.* **6** 6–14.

[8] DEMING, W. EDWARDS (1954). On the presentation of the results of samples as legal evidence. *J. Amer. Statist. Assoc.* **49** 814–825.

[9] DEMING, W. EDWARDS (1954). On the contributions of standards of sampling to legal evidence and accounting. Current Business Studies, Society of Business Advisory Professions, Graduate School of Business Administration, New York University.

[10] DEMING, W. EDWARDS (1960). *Sample Design in Business Research*, Chapter 3. Wiley, New York.

[11] EISENHART, CHURCHILL (1947). The role of a statistical consultant in a research organization. *Proc. Internat. Statist. Conferences* **3** 309–313.

[12] FREEMAN, WILLIAM W. K. (1952). Discussion of Theodore Brown's paper, (see [1]). *Amer. Statist.* **6** No. 1 18–20.

[13] FREEMAN, WILLIAM W. K. (1963). Training of statisticians in diplomacy to maintain their integrity. *Amer. Statist.* **17** No. 5 16–20.

[14] GORDON, ROBERT A. ET AL. (1962). Measuring employment and unemployment statistics: President's Committee to Appraise Employment and Unemployment Statistics. Superintendent of Documents, Washington.

[15] HANSEN, MORRIS H. (1952). Statistical standards and the Census. *Amer. Statist.* **6** No. 1 7–10.

[16] HILL, A. BRADFORD (1953). Observation and experiment. *New England J. of Med.* **248** 995–1001.

[17] HOTELLING, HAROLD ET AL. (1948). The teaching of statistics (a report of the Institute of Mathematical Statistics Committee on the teaching of statistics). *Ann. Math. Statist.* **19** 95–115.

[18] JENSEN, ARNE ET AL. (1948). The life and works of A. K. Erlang. *Trans. Danish Acad. Tech. Sci.*, No. 2. Copenhagen.

[19] MOLINA, EDWARD C. (1913). Computation formula for the probability of an event happening at least c times in n trials. *Amer. Math. Monthly* **20** 190–192.

[20] MOLINA, EDWARD C. (1925). The theory of probability and some applications to engineering problems *J. Amer. Inst. Electrical Engineers* **44** 1–6.

[21] MORTON, J. E. (1952). Standards of statistical conduct in business and government. *Amer. Statist.* **6** No. 1 6–7.

[22] SHEWHART, WALTER A. (1931). *Economic Control of Quality of Manufactured Product.* Van Nostrand, New York.

[23] SHEWHART, WALTER A. (1939). *Statistical Method from the Viewpoint of Quality Control.* The Graduate School, Department of Agriculture, Washington.

[24] SHEWHART, WALTER A. (1958). Nature and origin of standards of quality. *Bell Syst. Tech. J.* **37** 1–22.

[25] Society of Business Advisory Professions (1957). Report of Committee on standards of probability sampling for legal evidence. Graduate School of Business Administration, New York University.

[26] STEPHAN, FREDERICK FRANKLIN (1936). Practical problems of sampling procedure. *Amer. Sociological Rev.* **1** 569–580.

[27] ZIRKLE, CONWAY (1954). Citation of fraudulent data. *Science* **120** 189–190.