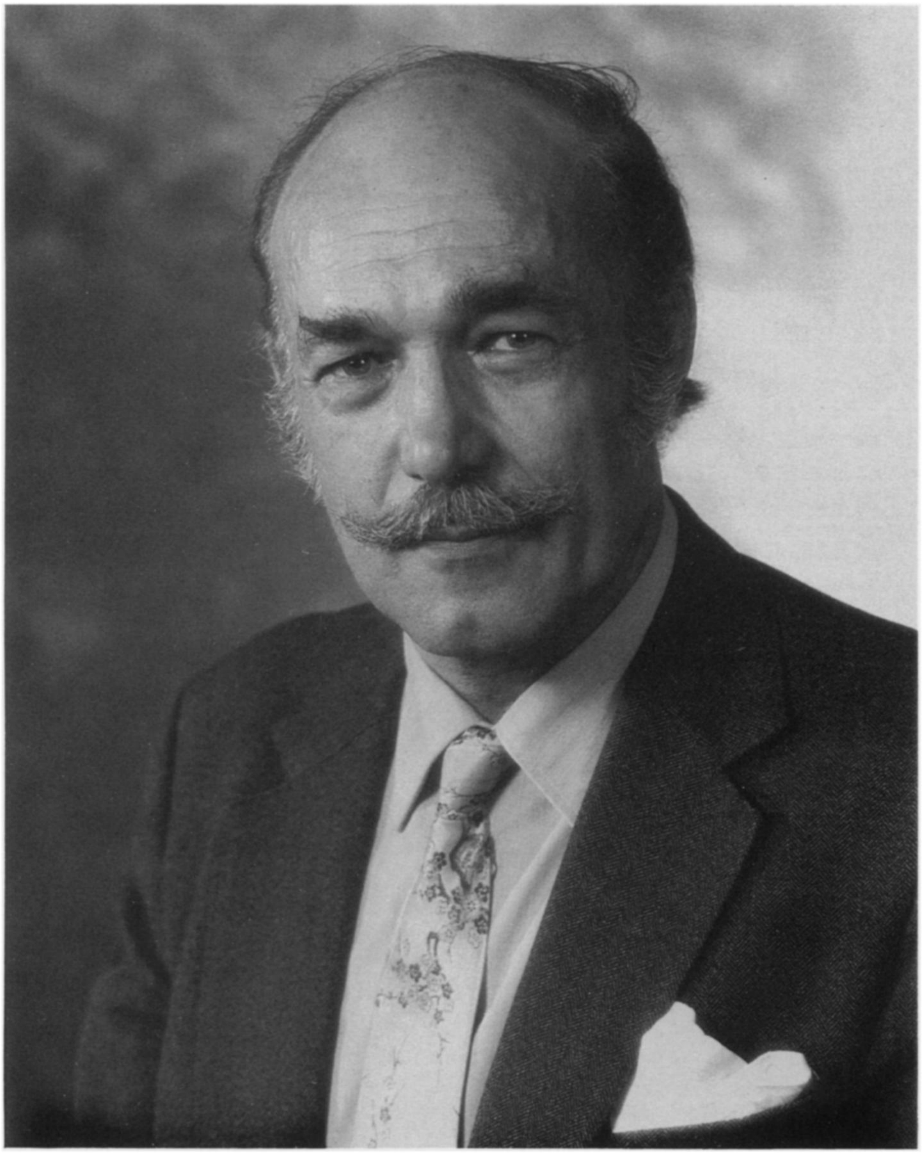
Statistics, Science and Technology

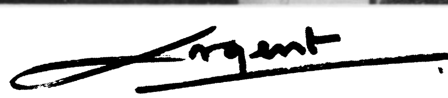
Author(s): J. A. Nelder

Source: *Journal of the Royal Statistical Society. Series A (General)*, 1986, Vol. 149, No. 2 (1986), pp. 109-121

Published by: Wiley for the Royal Statistical Society

Stable URL: http://www.jstor.com/stable/2981525





J. R. Statist. Soc. A (1986)
149, *Part* 2, *pp.* 109–121

Statistics, Science and Technology

By J. A. NELDER

Imperial College, London, U.K. and London Business School, London, U.K.

[*The Address of the President, delivered to the* Royal Statistical Society *on Wednesday, April 16th, 1986*]

SUMMARY

Statistics is seen as being primarily concerned with the theory and practice of the matching of theory to data by research workers. Swings between data-heavy and model-heavy views of statistics are discussed, and also aspects which inhibit communication between statisticians and scientists or technologists, especially the insufficient attention given by statisticians to the problems of combining information from many data sets. Several obstacles to better communication by statisticians are discussed, and also how the current gap between them and scientists/technologists might be bridged.

1. INTRODUCTION

The pride which one may legitimately feel on being elected President of this Society becomes rapidly tempered by a concern that one will have to give a Presidential address; this concern is not lessened by the fact that my immediate predecessors have between them surveyed in considerable breadth the state, scope, and prospects for our subject. They have also done so with an objectivity that I certainly shall not achieve. It is also inevitable that I should stress again some of the points that they have already made, but this is not, I suggest, necessarily a bad thing; the propagation of ideas has a component which consists of keeping on keeping on, however much this may annoy those who were convinced the first time round.

I have chosen to look at the relations between statistics on the one hand and science and technology on the other because I think that these relations are in several ways unsatisfactory, with the result that the ideas of statistics and the work of statisticians are not having the influence that they should on the work of scientists and technologists.

The distinction between science and technology has been much discussed, and Healy (1978) makes many relevant points about the place of statistics in this argument. He emphasises that in technology one is primarily concerned with measuring something, and not with testing theories, the latter being an activity of science. He also points out that in technology a theory may be near enough, while being known to be wrong, and yet still remain very useful. Thus the great technological (not scientific) achievement of putting men on the moon could safely use Newtonian dynamics — they were near enough. To the scientist, however, a theory with known defects is a continual irritant, demanding improvement or perhaps complete respecification.

However, as Kuhn (1970) argues, such complete respecification, involving as it does a revolution or catastrophic discontinuity in scientific thought, is rare and very few individuals achieve it. Mostly scientific activity consists of what he calls normal science, and involves working within an accepted theoretical framework or paradigm. In Kuhn's words (*op. cit.*) 'Mopping-up operations are what engage most scientists throughout their careers'.

I doubt if statistics has much to offer to revolutionary science; indeed I wonder if those who have tried to argue that a theory can be assigned a probability, or that the class of theories can be closed and so be made a suitable subject for probability statements, have really understood the kind of discontinuity involved in a scientific revolution.

The position is very different, however, with both normal science and technology, though some justification may be needed for grouping them together here. I would argue that while

0035–9238/86/149109 $2.00

there are major differences between normal science and technology, for example in the relevance of cost–benefit assessments to their progress, their long-term objectives, and their attitudes to theory (Healy, personal communication), it is nonetheless true that, on the scale of day-to-day activity, the procedures of (normal) scientist and technologist will be found to be very much alike. They will both be working within a given theoretical framework and be concerned, for instance, with estimating quantities defined within that framework, with confirming the estimates of others, and with relating their estimates to those predicted by theory.

At any point in a research programme there are two prior components, *existing theory* and *prior data* from past studies (I shall use *study* to include both experimental and observational collecting of data). These in turn influence the *design* of the new study, which, after its *execution*, yields *new data*. Finally *analysis* tries to integrate the new data with the prior data and the existing theory to produce an updated theory and an extended prior data set.

Of the three basic activities — design, execution and analysis — the statistician has little to say about the second. However, it is most important that (s)he understand its potential complexity, which was well described by my immediate predecessor (Bodmer, 1985a). With design and analysis, however, things are very different, and here the statistician has a role as the person who is interested in the theory and practice of these processes.

I see statistics as giving a central place to the theory and practice of the matching of theory to data. In doing so I do not underestimate the importance of the design phase; good design remains crucial, because without it effective analysis can become impossible. Every statistician has his quota of horror stories about studies where no useful conclusions could be drawn because of poor or non-existent design.

If the statistician can rightly claim to be interested in the theory and practice of crucial activities in science and technology, we need to ask why many scientists and technologists still believe that statistics has little to offer them in their work. In Sections 2–4 I shall look at some of the problems involved.

2. THE TENSION BETWEEN STATISTICAL THEORY AND DATA

The founders of our Society believed that the collection of data was of fundamental importance for casting light on economic and social problems. They saw the collection itself as the vital part, hence the original motto of the Society '*aliis exterendum*' — let others thresh. Here we see a view of statistics as concerned primarily with the collection of *data*, symbolised by the wheatsheaf; however, in 1857 the motto was abandoned, for what had then become clear is that statisticians cannot just be reapers, they must also be threshers.

Now much as I like agricultural metaphors, threshing is not quite right as a metaphor for extracting the important aspects of a set of data, separating the wheat from the chaff. A more modern metaphor, which speaks of separating the signal from the noise, is much more enlightening I think. It encourages us to think of our data as an instance of a process where a systematic signal is overlaid by chaotic noise, and to define our basic activity as extracting the best representation of the signal given its inevitable distortion by the noise. Success depends upon three components:

(i) a good description of the signal,
(ii) a good description of the noise,

which together comprise the *model*, and

(iii) matching the model to the data.

As indicated above, I propose to give a central position in statistics to the third component, which requires both data and model. As we look back over the history of our subject we can see swings towards views of statistics that are either data-heavy or model-heavy. The original data-heavy views of our founders eventually gave way to a period which can be described as model-heavy. At this time much more attention was given to the mathematical properties of models, particularly those describing the noise, than to data or the matching process itself. Furthermore the emphasis on the noise rather than the signal in models tended to create a gulf between the statistician and the scientist or technologist; for whereas the latter were mainly interested in the signal, they often found that the former was mainly interested in the noise. The result was, and is, misunderstanding and non-communication on a considerable scale.

2.1. *The Arrival of Data Analysis*

"*Data! data! data!*" *he cried impatiently.* "*I can't make bricks without clay*"
Conan Doyle — *The Copper Beeches*

It was inevitable that there would be a swing away from theory towards data again, and this movement is very much associated with the name of J. W. Tukey (1977). So much had statistics come to mean statistical mathematics to many (particularly perhaps in the USA) that Tukey introduced the term 'data analysis' to describe techniques for looking at data. The extent to which statistics had indeed come to mean statistical mathematics to many people was shown at a recent meeting in N. America where a speaker explicitly distinguished between statisticians and those who looked at data. I reject this distinction, believing it to be a contradiction in terms; however, the fact that it was made shows how far the balance had previously shifted in the model-heavy direction.

While we should welcome this attempt to shift the balance back towards data, there are nonetheless some worrying features in the current scene. First, there is the matter of nomenclature: personally I have always disliked the term 'data analysis', believing that process to be a part (valuable and under-described though it may be) of 'statistical analysis'; the danger of giving it a separate name is that people may come to believe that it is an entirely separate activity and not part of statistics, rather than being an important component of what a statistician does. The recent paper given to the Society by Chatfield (1985) illustrates very well its true place (as I see it).

A second worry is that data analysis may degenerate into the mere playing of games with numbers, as an activity in its own right, whereas it should constitute a component of the crucial matching process. A particular danger here is that of ignoring prior information about possible patterns of both signal and noise when looking at the data. At an elementary level this can lead to unhelpful activities such as looking at the areas of the States of the USA as a set of 50 numbers, ignoring all that is known about their settlement and history of acquisition. In my view these 50 numbers cannot represent a homogeneous set of anything and no fledgling statistician should be encouraged to think that they can; to attempt to do so is to enter the complex process of statistical analysis at the wrong place.

3. THE STRUCTURE OF DATA

It is a capital mistake to theorize before one has data.
Conan Doyle — *Scandal in Bohemia*

A noticeable result of many scientists', technologists', and statisticians' education is a lack of understanding of the possible structure and types of data. There are several possible reasons for this. For one thing, data are untidy; they do not satisfy exact mathematical relationships, and so do not lend themselves to deductive reasoning. Secondly data arrive embedded in a more or less complex context, which describes what they measure, the units of measurement, the conditions under which they were collected and the reasons why they were collected. It includes prior knowledge of the system, the design of the present experiment or study, known sources of heterogeneity (e.g. as expressed by blocking factors), the properties of the experimental units, and so on. This context, though vital to understanding the data, may be seen by some, wrongly, as irrelevant to the making of statistical points. Finally it may be thought that the description and classification of the various structures of data is a rather low-level activity, not to be compared with the mathematics of statistical theory.

Whatever may be the reason for the neglect of data in the education of statisticians in particular, and of scientists and technologists in general, the effects are both noticeable and deplorable. People whose work will be intimately connected with acquiring and making sense of measurements often finish their training (i) unable to look intelligently at a three-way table, (ii) unable to distinguish between crossed and nested classifications, (iii) unaware of the existence of gross errors in data or of the reasons for their occurrence, and (iv) with no feel for dimensionality. A common reaction to complaints about this state of affairs is that 'this is the sort of thing you pick up on the job'; however, it could be argued that, if anything, the reverse is the case, and that it may be easier to pick up some unfamiliar theory later when it is clearly seen to be needed. Certainly lack of understanding of differences in the structure of data has widespread consequences.

Firstly, it can lead to inappropriate models for analysis being selected; thus it is common to find multivariate techniques originally developed for isolated and unstructured data sets, e.g. principal component analysis, being applied unheedingly to structured data. The analysis is forced into an unsuitable framework through ignoring vital properties of the data. If models were introduced from the beginning by matching them with compatible data structures, much of this misuse would be avoided.

Secondly, there can be a failure to distinguish properly between distinct forms of analysis. For example, correspondence analysis (Greenacre, 1981, 1984) is a name applied to operations on two quite distinct data structures. One is a simple data matrix of counts in which the columns represent distinct measurements in the form of frequencies and the rows represent experimental or observational units; the other is a two-way contingency table of counts, obtained by tabulating occurrences with respect to two classifying factors. The mathematical apparatus used, which could be described as the singular-value decomposition of the matrix of Pearson residuals (McCullagh and Nelder, 1983) from a main-effects log-linear model, is the same, but the structures on which it operates, and hence the inferential contexts, are quite different. It seems unfortunate that the distinction of data structure is often ignored, while we can be sure that a corresponding difference in the theoretical component would not have been.

4. THE CULT OF THE ISOLATED STUDY

In the full tide of successful experiment

Jefferson — *First Inaugural Address*

The ignoring of prior information, which is often apparent in expositions of data-analytic techniques, raises a general question about the value of statistics in science and technology. Most statistical books and papers place enormous emphasis on the analysis of the unique experiment or study. Much statistical expertise is deployed to make inferences from a single isolated data set, treated as if it were essentially unique. The fact that (almost always) it is one of a whole sequence of comparable data sets, and that inferences need to be embedded in what we have learnt from previous experiments, is largely ignored. This emphasis on the isolated study, with the corresponding lack of emphasis of problems of combining information from many experiments, is, I believe, an unsatisfactory feature of much statistical writing.

Historically emphasis on the single (often fairly small) data set can be traced to Student (1908). He saw, rightly, that the procedures of the biometric school did not offer much to the statistician dealing with small experiments or sampling schemes. Fisher greatly extended Student's work, while maintaining the idea of extracting maximum information from limited data. These men were fully aware that the single experiment was only a small part and not the whole of a study, but others have been less so. As a result there has grown up a literature based on the reporting of a single experiment in which great emphasis is placed on the notion of a significant difference, with its associated significance tests. The effects claimed may never be checked by painstaking reproduction of the study elsewhere, and when this absence of checking is combined with the possibility that the original results would not have been reported unless the effects could be presented as significant, the result is a procedure which hardly deserves the attribute 'scientific'. Very little of the statistics in such papers is done by statisticians, and the view of the experimenters is frequently that they are going through these procedures in order to satisfy the requirements of supervisors, editors or referees rather than to help their own thinking. Not surprisingly such people see statistics as a chore to be done rather than a tool to be used. When they do use the tools it is often to reinforce an existing binary mode of thinking in which questions are posed with yes–no alternatives such as 'is this new treatment better?' Such questions are often the wrong ones for the investigator to ask, and the fact that (s)he does so is often the result of lack of ability to think quantitatively. There are welcome signs that this kind of investigation with its associated statistics is coming under attack from various directions (Carmer and Walker, 1982; Carver, 1978; Guttman, 1985; Preece, 1982).

If we look now at established procedures in the physical sciences we find that the scientist begins to believe that (s)he is winning when (s)he gets reproducible results from several experiments done under various conditions, perhaps with different instruments at different sites etc. Looking for reproducible results is a search for significant *sameness*, in contrast to the emphasis on the significant *difference* from a single experiment. The problem here is that the scientist or technologist may be quite unaware that statistics has a useful role in this kind of activity. Certainly most textbooks give little sign that it has.

We should be paying much more attention to the modelling of the variation *between* data-sets and to trying to formalize the ways in which the scope of associated models may be defined and measured. This will be a challenging task, if only because the well-controlled set-up which exists *within* an experiment largely disappears when we look at differences *between* experiments (Ehrenberg, 1975, p. 346); thus all the problems attending the interpretation of observational data come crowding back.

One really useful insight of the statistician is that (s)he sees why throwing every possible effect into a model doesn't lead to enlightenment. (S)he is, therefore, in a strong position to press for the establishment of *protocols*, i.e. a standardization of conditions under which trials or experiments are done, whereby the scope of inferences is deliberately constrained in the hope that more nearly reproducible results will emerge. Progress in this field is sure to meet an immediate response from many scientists and technologists, who will see the statistician concentrating on matters of obvious relevance to their own work. Furthermore work in this area will meet basic problems in probability and inference, and require a theory of great intrinsic interest as well as one having many applications.

A fundamental notion seems likely to be that of *homogeneity*: an anecdote may help to illuminate the point. A physicist colleague who had no time at all for statistics in his work (he appeared to believe that physics could, in principle, explain *all* variation) said to me one day "I've come to the conclusion that taking a mean is a very difficult business". He had been wrestling with the problem of combining information from many data sets and had recognised that to take a mean of some quantity requires some kind of homogeneity of the values averaged if the resulting value is to be useful. He could not, in the manner of a probabilist, *postulate* a population from which his set of values was a random sample; rather he had to *identify* a population which the data themselves suggested was homogeneous. The sad thing, to me, was that it had never entered his head that statisticians might have encountered this problem, or that statistics might have anything useful to say about it.

We ought to look carefully at this state of affairs. Recently the term 'meta-analysis' has been introduced (Glass *et al.*, 1981; Hedges and Olkin, 1985) to describe the combination of information from many studies. The use of this, to me, rather pretentious term for a basic activity of science is a clear indication of how far some statisticians' view of statistics has diverged from the basic procedures of science. I think we can do without it and, indeed, should try to make it unnecessary; then scientists and technologists would have much less cause to see statistics as at best marginal to their interests.

5. MAKING OURSELVES MORE USEFUL

If statisticians are to make the contributions which they could to science and technology several other features of our work need attention. These include:

(i) the relationship with mathematics
(ii) notation
(iii) communication
(iv) the relationship with the computer

There are many cross-links between the topics.

5.1. *The Relationship of Mathematics to Statistics*

Mathematics is a subject in which we do not know what we are talking about, nor care whether what we say is true.

Bertrand Russell

The development of the theoretical part of statistics would be impossible without the body of mathematical notation and theory, and yet the relation of mathematics to statistics is complex. Some who think of themselves as statisticians adopt the aims and objectives of mathematics in their writings, while others (for example, Pearce (1983)), distrust certain effects of mathematics on the development of our subject.

A mathematical theory, such as group theory, constructs an edifice of theorems built on a well-defined set of axioms. The method of exposition (though not usually the method of discovery) is deductive, and some of the results are of enormous power and generality. But the theorems are totally abstract, as Russell's characteristic aphorism so aptly declares. That is, the theory stands on its own, without reference to possible interpretation in terms of objects in the world outside, their properties and behaviour. In statistics, by contrast, we ought to know what we are talking about, in the sense of relating our theory to external objects. We should also care about whether what we say is true, in the sense of our inferences and predictions being well supported by the data.

When mathematicians construct theories they do not seem in general to think of themselves as constructing *tools* for others to use. That they frequently, and apparently inadvertently, do just that has often been remarked upon; the development by Ricci and Levi-Civita (1901) of the tensor calculus, which proved to be just what Einstein needed for relativity theory, is a famous example. If the applicability of mathematical theories as tools in statistics is indeed unplanned, then we should not be surprised if their application can be both liberating and constricting. We need both to take what is useful from a theory and to refuse to be constrained by it where it proves unsuitable for our purposes. Experimental design provides an interesting example. The use of combinatorial theory was and is of enormous use in the construction of incomplete block designs, for which the design matrix $\mathbf{X}$ has elements either zero or one. However statistical interest centres on the properties of the information matrix $\mathbf{X}'\mathbf{X}$; the idea of uniform information can be deformed to that of nearly uniform information, and remain useful to the experimenter. Insistence on working inside the combinatorial framework delayed, I believe, the exploitation of useful designs which did not fit into that framework.

The main danger, I believe, in allowing the ethos of mathematics to gain too much influence in statistics is that statisticians will be tempted into types of abstraction that they believe will be thought respectable by mathematicians rather than pursuing ideas of value to statistics. One origin of this temptation is undoubtedly the siting of statisticians working in Universities in Departments of Mathematics; the pressure on the statisticians to develop their researches in directions thought to be acceptable to mathematicians may then become too strong to be easily resisted. However, there is little doubt that it ought to be resisted, for the two disciplines have very different objectives.

5.2. *Notation*

The vocabulary of 'Bradshaw' is nervous and terse, but limited.
Conan Doyle — *Valley of Fear*

Something that we might hope to take from mathematics is a simple and clear notation for expressing our ideas. It is a truism that enormous strides have been made in improving notation in the last 500 years — anyone who doubts the value of standard modern mathematical notation should try reading the original paper by Tartaglia (Struik, 1969) on the solution of the cubic, or try to imagine the problems that Kepler faced in finding his third law with its two-thirds power without Cartesian co-ordinates, polar co-ordinates or logarithms to help him.

Leibnitz (Hollingdale, 1985) hoped to develop a universal notation for the expression of all mathematical problems and was perhaps the first man to see clearly the importance of notation as such. Though the grand plan failed, his monument will always be the modern notation for the calculus. Many problems of notation remain unsolved and a look at the statistical literature will quickly confirm that there is still much to be done. It is quite easy to write a definition of a probability density function, using standard mathematical notation, in such a way as to disguise almost totally the main point of the idea. The mathematical niceties, hardly ever needed, swamp the basic notion. An interesting discussion could be generated with our continental colleagues on the extent to which the Bourbaki notation, apparently mandatory in many of their papers, actually helps to make clear the statistical points at issue. Admittedly I am biased because I was not brought up on this notation; however, in several cases I concluded that, even allowing for this, the notation had served to obscure rather more than it revealed about the points at issue.

I hope you will allow me an example in which I have a personal interest; this concerns the definition of factorial models and their fitting by least-squares. In the discussion of a paper by Pearce (1963) given to this society I proposed a way of describing such models, which I later developed further (Nelder, 1965). These ideas were considerably extended by Wilkinson and Rogers (1973) to produce the syntax for the linear predictor in *GLIM*, and also for the block and treatment structure formula used in the part of Genstat that deals with the analysis of designed experiments (Payne and Wilkinson, 1977). This syntax often needs less than half the number of symbols of some alternatives, without being less informative. The underlying least-square equations can be written reasonably compactly in terms of Kronecker products of patterned matrices, though this may well not be the last word. Certainly least-squares theory, a basic part of statistics, is often made to look intimidating by unsatisfactory notation.

5.3. *Communication*

Thou shalt not sit with statisticians/nor commit a social science.
W. H. Auden — *Under Which Lyre*

The consequences of poor notation are serious at several levels: we may fail to communicate our ideas to our colleagues; we may make life difficult for the next generation trying to learn statistics; we may fail to get out ideas across to other scientists; and on a much wider scale we may fail to interest the public at large in what we are doing.

It might help communication in the first circle, other statisticians, if the algorithmic way of describing procedures were used more often in place of the purely mathematical. In algorithms we can use identifiers with more than one symbol, '*yield*' instead of '*y*', so helping understanding; iterative processes are naturally described in terms of loops, and if-then-else constructs deal simply with branching. An algorithmic description can provide a useful bridge between plain language and mathematics, and may be clearer than either. The whole subject of notation in relation to communication deserves more study; for while prolixity is common and often self-defeating, extreme terseness can be equally so, witness the not-entirely-unfair description of APL as a write-only language. I remember being very impressed as a student by the reduction of classical dynamics to the single principle that the passage from one state to another state followed the path of least action (A): mathematically this was expressed by

$$\delta A = 0.$$

Terseness can hardly go further, and it would be interesting to know how enlightening this is generally found to be.

The much wider question of communication with the general public was ably discussed by my immediate predecessor (Bodmer, 1985b), who knows much more about it than I do. It is often argued that at this level of communication we should dispense entirely with mathematical notation; one reason given is that many have been rendered fearful of the whole subject by their education and so will be put off. Another, rather different, argument claims that mathematics (or jargon generally) is too often used as a deliberate form of mystification, whose aim is to show that the expert knows best. I am not entirely convinced by either of these arguments; for while bad, unnecessarily prolix, notation may convince some readers that just because it looks incomprehensible therefore the writer must be very clever indeed, good notation, I believe, should have the opposite effect. It makes clear what is being discussed and should actually be better than its verbal equivalent.

As a Society we face a major task in making clear at all levels what our special expertise and insights consist of. We rightly believe that these are important in very many areas of science, technology, and public affairs. But whereas a topologist, say, may justifiably aim his papers at other topologists, and not try to explain himself to number theorists, statisticians need to take a different approach. Even to aim one's papers at statisticians means envisaging a most heterogeneous target population; to look further, at the users of statistics, is to enlarge this population hugely.

The French have a word for it—'vulgarisation', which carries with it none of the implicit lack of intellectual respectability that its English nearest equivalent 'popularisation' has. I do not know how to make this activity respectable, but that it should become respectable I have no doubt at all.

No discussion of communication can today ignore the huge potential of the computer, a unique tool which needs separate discussion.

5.4. *The Computer*

In the last 35 years the computer has gone through four generations, and the industry is bracing itself for the appearance of the fifth. I have tried to summarize elsewhere the history and prospects of statistical computing (Nelder, 1984), distinguishing a first generation of isolated programs, a second of statistical packages, and a third (beginning to appear) of the expert system. Here I want to consider briefly the impact of the computer on our relations with science and technology. Second-generation packages have been widely used by scientists and technologists - some would say widely misused. The fact that it is as easy and simple for the user to do the wrong thing (such as calculating t-statistics for wildly non-Normal data) as the right thing continues to worry many of us. The packages themselves reflect a view of statistics taken by the developers (not all of whom have been statisticians), some aspects of which are unhelpful to scientists for reasons I have argued earlier. They have often been used as magic boxes, to give respectability to an analysis because it has been done on a computer; we are, I hope, now moving out of that era. Packages are often constricting, forcing users into particular channels because those are what the package has algorithms for, while at the same time making it difficult to move to another package that may have the facility required. They usually contain little or nothing in the way of checks that the procedure being used is relevant for the analysis of the current data set.

Whatever may be the shortcomings of packages (and as the instigator of two, I am only too well aware that they exist), we should not forget that, for better or worse, they have been a powerful force in statistical education. To a considerable extent many scientists and technologists will have had their statistical ideas shaped by the packages they have used; it is vital, therefore, that statisticians play their full part in the development of computer software, and do not leave it to computer scientists, who may imagine that all they have to do is to implement the arithmetic of the procedures in a standard text. This is especially important now that third-generation software in the form of expert systems is beginning to emerge.

5.5. *Expert Systems*

We are now faced with the challenge of formalizing what we do in designing studies and analysing data, in a way that can be described to a computer and so displayed to guide the user. We must move from knowing *how* to do something (and so being able to write an algorithm for it) to advising on *what* to do.

To help us we have all the tools developed by the Artificial-Intelligence (AI) community. These include rule-based declarative languages, like Prolog, knowledge-elicitation procedures for debriefing the expert, expertise on the man-machine interface (how people react with computers), and so on. Gale (1985) gives a good picture of the current state of the art.

I am currently working on a project to provide a front-end for *GLIM* (Nelder and Wolstenholme, 1986); though this is still in its early stages, a number of general points have emerged. One is that the rules in a statistical expert system are at a level of abstraction above those of a diagnostic program in, say, medicine or geological prospecting; the skill of the statistician embodies experience in many fields of application, and there is no question of incorporating in the expert system the special knowledge of those fields. What must be included are rules abstracted and potentially applicable to *any* future data set, and this makes knowledge elicitation by use of a knowledge engineer difficult, perhaps impossible.

We must distinguish between systems with a fixed rule base, and those which create new rules, or modify existing ones, as the result of their interaction with users. Some would deny the term 'expert system' to the first kind, though such a system can certainly incorporate expertise. For an account of an ambitious plan to construct a system of the second kind see Gale (1985), Chapter 10.

Another important dichotomy is between authoritarian and libertarian systems. While an authoritarian system will constrain the user at any stage to a limited set of options, a libertarian system will not; instead it will concentrate on giving good advice conditional on the path already taken. If we are to allow adequately for the user's special knowledge about his data we must aim to construct a libertarian system. This will not please, however, the user who does not want to think or the kind of implementer who thinks he always knows best.

A final point concerns the ways in which the user can query the system. (S)he may want to ask

(i) What options are available at this stage?
(ii) What does a word (e.g. *aliasing*) mean in a question?
(iii) Why is this question being asked?
(iv) What does the system advise?

While the expertise is primarily contained in the answers to user's questions of type (iv), provision of answers to the other types of question are vital if the system is to appear friendly to a wide range of users.

We have much to learn about this new kind of software, which should not be thought of as a panacea. Procedures unhelpful to the scientist and technologist can be codified as rules just as easily as helpful ones, and we can confidently predict that the misuse of packages will be followed by the misconstruction of expert systems.

6. BRIDGING THE GAP

The statistician and the scientist/ technologist will find more in common, to the advantage, I believe, of both, if their education can be more closely linked.

Firstly, the statistician: during training (s)he should take part in a designed study or experiment, becoming involved in the design, the planning of data collection, *the actual collection of the data themselves*, and the subsequent analysis. This in turn will involve becoming *au fait* with some relevant existing theory in a particular field of application together with the previous data. The aim will be to give the trainee statistician insight into the problems of the effective collection of quantitative information, including some understanding of its dynamic nature, with the one study embedded in a larger sequence. More emphasis needs to be placed in his courses on measurement, its accuracy and how to assess it, and less on significance tests. Models need to be introduced with much more emphasis on the systematic part, and the finding of suitable models compatible with the structure of the data needs to be developed from the beginning. This in turn implies more emphasis on basic data structures. Early familiarity with well-structured computer packages is needed, and in time with good expert systems.

Secondly the scientist/technologist: (s)he needs also to understand the basic kinds and structures of data, the distinction between a parameter and its estimate (an important statistical insight currently little understood in wide areas of science and technology), how to collect a set of data free of gross errors and major sources of bias, how variation in the mass can be described via the random part of models, why the design stage of a study is important, and the idea of the amount of information about a parameter value in a data set.

Finally both need to understand that models are necessarily simplifications of the system being modelled; that they are, in an absolute sense, wrong; that they are certainly provisional, but nonetheless are useful and necessary for successful quantitative thinking. Statistics provides powerful arguments for simplicity or parsimony in models, and these are little understood. I remember several cases where the development of a simple model, replacing an initially complex one, was met by the scientist not with interest or pleasure, but with indifference or hostility. We cannot afford to have people so ignorant of the tools of their trade that they behave like this.

CONCLUSION

I am convinced of the great value of statistical ideas and methods to the scientist and technologist, but we have much to do before that value can be realised. Yet surely the goal is worth striving for. Let the poet have the penultimate word:

Upon this gifted age, in
its dark hour,
Falls from the sky a
meteoric shower
Of facts . . . they lie
unquestioned, uncombined.
Wisdom enough to leech
us of our ill
Is daily spun; but there
exists no loom
To weave it into fabric.

Edna St. Vincent Millay
Huntsman, What Quarry?

We should be loom-makers.

ACKNOWLEDGEMENTS

It is a pleasure to record helpful discussions with Professors A.S.C. Ehrenberg and M. J. R. Healy, who helped to clarify several points. Remaining deficiencies are entirely mine.

REFERENCES

Bodmer, W. F. (1985a) Understanding statistics (Address of the President). *J. R. Statist. Soc.* A, **148**, 69–79.
Bodmer, W. F. (1985b) *The Public Understanding of Science.* London: The Royal Society.
Carmer, S. G. and Walker, W. M. (1982) Baby bear's dilemma: a statistical tale. *Agronomy J.*, **74**, 122–4.
Carver, R. P. (1978) The case against statistical significance testing. *Harvard Education Rev.*, **48**, 378–399.
Chatfield, C. (1985) The initial examination of data (with Discussion). *J. R. Statist. Soc.* A, **148**, 214–253.
Ehrenberg, A. S. C. (1975) *Data Reduction.* Chichester: Wiley.
Gale, W. A. ed. (1985) *Artificial Intelligence and Statistics.* Reading, Mass.: Addison-Wesley.
Glass, G. V., McGaw, B. and Smith, M. L. (1981) *Meta-Analysis in Social Research.* Beverly Hills: Sage Publications.
Greenacre, M. J. (1981) Practical correspondence analysis. In *Interpretating Multivariate Data* (V. Barnett, ed.). New York: Wiley.
Greenacre, M. J. (1984) *Theory and Applications of Correspondence Analysis.* London: Academic Press.
Guttman, L. (1985) The illogic of statistical inference for cumulative science. *Appl. Stochastic Models and Data Analysis*, **1**, 3–9.
Healy, M. J. R. (1978) Is statistics a science? *J. R. Statist. Soc.* A, **141**, 385–393.
Hedges, L. V. and Olkin, I. (1985) *Statistical Methods for Meta-Analysis.* Orlando: Academic Press.
Hollingdale, S. H. (1985) Leibniz and the first publication of the calculus in 1684. *Bull. Inst. Math. Appl.*, **21**, 88–94.
Kuhn, T. S. (1970) *The Structure of Scientific Revolutions*, 2nd Ed. University of Chicago Press.
McCullagh, P. and Nelder, J. A. (1983) *Generalized Linear Models.* London: Chapman and Hall.
Nelder, J. A. (1984) Present position and potential developments: some personal views—statistical computing. *J. R. Statist. Soc* A, **147**, 151–160.
Nelder, J. A. (1965) The analysis of randomised experiments with orthogonal block structure. I: block structure and the null analysis of variance. *Proc. Roy. Soc.* A, **283**, 147–162.
Nelder, J. A. and Wolstenholme, D. (1986) A front-end for GLIM. *Computer Science and Statistics: Proceedings of the 18th Symposium on the Interface.* Washington D.C.: American Statistical Association.
Payne, R. W. and Wilkinson, G. W. (1977) A general algorithm for analysis of variance. *Appl. Statist.*, **26**, 251–260.
Pearce, S. C. (1963) The use and classification of non-orthogonal designs. *J. R. Statist. Soc.* a, **126**, 353–377.
Pearce, S. C. (1983) The monstrous regiment of mathematicians. *The Statistician*, **32**, 375–378.
Preece, D. A. (1982) The design and analysis of experiments: what has gone wrong? *Utilitas Math.*, **21A**, 201–244.
Ricci, M. G. and Levi-Civita, T. (1901) Méthodes de calcul differentiel absolu et leurs applications. *Math. Ann.*, **54**, 125–201.
Struik, D. J. (1969) *A Source Book in Mathematics, 1200–1800.* Harvard University Press.
'Student' (Gosset, W. S.) (1908) On the probable error of a mean. *Biometrika*, **6**, 1–25.
Tukey, J. W. (1977) *Exploratory Data Analysis.* Reading Mass.: Addison-Wesley.
Wilkinson, G. N. and Rogers, C. E. (1973) Symbolic description of factorial models for analysis of variance. *Appl. Statist.*, **22**, 392–399.

Two recent Presidents of the Society responded as follows

Professor P. Armitage (University of Oxford): Professor Nelder is, during the current session, guiding the fortunes of the Society with both acumen and tact, and we are all grateful to him for taking on this task of leadership. That he should feel concern at the additional requirement to produce a Presidential Address on time is entirely understandable, and he has my sympathy. Let me therefore assure him that his labours have not been in vain. He has given us a penetrating survey of the place of statistics in science and technology, written in a delightful style that makes me sense that he enjoyed it after all. I regret only that I am unable to be present in person this evening.

Professor Nelder is almost uniquely qualified to tackle the subject he has chosen. He has made important theoretical contributions to our subject and he has for several decades been a leading figure in the application of statistics to agricultural science and technology—and I must be careful to use both terms here. Moreover, he is a principal architect of two of the most widely used statistics packages, and has, in this respect alone, had an enormous influence on current practice.

The President has for many years had a keen interest in the philosophy of science, and we can appreciate this evening the broad and balanced view to which he has been led. I am surprised in this connection to see no reference to Popper, whom he used to admire, and wonder whether he has transferred his affections to rival schools of thought. He is to be congratulated on referring to Kuhn and yet making only one use of the word 'paradigm'. Moreover he avoids the usual adjective for this word, which, because I am not allowed to be controversial I will quote as 'B—ian'. As regards the distinction between science and technology I am inclined to think that although in principle it is clear and important, in practice it may often be difficult to draw. The experiments by Cushny and Peebles used by 'Student' in his 1908 paper could be regarded as exploring the hypnotic effects of different drugs (surely technology), or as part of a long-term programme of fundamental research into the nature of optical isomers (surely science). Enquiries often have many purposes and many consequences, and we should perhaps not try to categorize them too rigidly.

I am glad the President referred to meta-analysis, where I agree the methodology has been unduly mystified, partly because of the curious tendency to combine standardized effects rather than raw effects. However, the general idea of combining results from similar studies is important, and has been put to very good use by Richard Peto and his colleagues in combining the results of trials in various branches of medicine.

Those of us who attended the less intellectual parts of the 150th Anniversary Proceedings will recall that Professor Nelder is a formidable pianist as well as a connoisseur of some of the recondite corners of the 19th century pianoforte literature. (A piece called "The Fireman's Galop" comes to mind). He has this evening displayed catholic tastes in poetry. I applaud this move to link statistics with the arts, and recall one or two meetings, of this and other societies, which were enlivened by musical performances. Perhaps we should look wider afield and seek the help of visual artists to depict stochastic phenomena in space and time. If this seems too fanciful, my excuse is that the President's address should encourage us to look outwards from our discipline towards other modes of thought. I have much pleasure in proposing the vote of thanks.

Professor Sir David Cox (Department of Mathematics, Imperial College, London): The unavoidable absence of Sir Walter Bodmer gives me the unexpected duty and pleasure of seconding a vote of thanks to Dr Nelder. Until fairly recently tradition on these occasions forbade comment on, let alone criticism of, a Presidential Address. Nowadays ruthless criticism is expected. In fact, however, in this wide-ranging address on very important issues, I find myself in virtually total agreement with Dr Nelder and therefore all I can do is pick out one or two points for further comment and reinforcement.

It is good to see the stress on the importance of notation. A key word here should be flexibility, I think. A good notation for general theory can be impossibly ponderous for special cases. The general notation for functions $f: R \rightarrow R$ has merit, particularly when special sets are involved. Yet it and the notation $f(.)$ are very poor for special functions. Various notations are useful for matrices and it is interesting to see emerge in recent months renewed emphasis on Cartesian tensors. Those who in fact do most matrix calculations secretly in suffix form take heart! Deliverance and even respectability may be at hand. I am however far from enthusiastic about the suggestion of more use of algorithmic definitions; it is purpose that counts initially not technique of computation.

Associated with questions of notation are those of nomenclature. GLIM is an evocative name! Abbreviations in general, however, seem most suitable for private notetaking; ca and pca met manova in eda ...

The distinction between the role of statistics in science and in technology and the distinction between normal and revolutionary science puzzle me. In some sense applications of statistics in applied science are more visible than in basic science, but is this more than a comment on the signal to noise ratios involved and possibly on educational background? I see a distinction between investigations in which one objective is to seek understanding of some phenomenon in terms of underlying concepts, and to aid in the formation of those concepts, and those investigations which are concerned with the direct description of the phenomenon under study; but it is perfectly possible to do both types in textile technology, say. Such a distinction is not one of subject matter but of purpose.

Also it is clear that a very few investigations are of very profound and lasting consequence, in particular in concept formation, rather more are very important, more still quite important and so on, rather a lot being concerned with very minor detail or elaboration. Yet is it really sound to pick out the extreme tail as qualitatively different from the rest? I don't know, but rather doubt it. More importantly, it seems likely that the whole edifice hangs together, and that it is easy to underrate the importance of the building up of detail. Quantum theory and special relativity, while in one sense revolutionary, were, as I understand it, natural developments from classical physics, quite slowly arrived at with much detailed work, indispensible at the time, even if largely forgotten now.

Finally I would like to comment on what I take to be a central theme of Dr Nelder's address, some alarm at the present position of statistics. We are, I agree, in a paradoxical and in many ways disturbing situation. Dr Nelder has described negative features. On the other hand, many scientific journals contain statistical material, sometimes innovative and subtle and, in wandering down University corridors at random, it is quite frequent that one overhears some well loved word, not perhaps "ancillary", but certainly "least squares", "sampling", or, dare I say it, "statistically significant". Rapid developments in the subject are taking place, but except perhaps in medical statistics, professional statisticians and societies are often not very prominent; this is an international phenomenon. It needs great attention, although there are positive, as well as very worrying aspects to it.

We have heard a fine Presidential Address with important comments on important topics. It gives me great pleasure to second the vote of thanks.

The vote of thanks was passed by acclamation.

As a result of the ballot held during the meeting, the following were elected Fellows of the Society.

Cassin-Scott, Anthony R.
Chudy, Bernard
Davies, Richard B.
De Morgan, Richard M.
Else, Elisabeth R.
Eltinge, John L.
Gilmour, Tom
Gough, Jacqueline J.
Harding, Simon A.
Harrison, Michael R.
Kelly, Krystyna A.
Laing, John B. G.
Lee, Patrick J.
Legge, Kathryn
Mead, Andrew
Mohamad, Hadi A.
Moore, Leslie M.
Osborne, Christine
Pickles, Andrew R.
Saidian, Masouda
Solly, Karen J.
Struthers, Lesley P. L.
Van Der Ploeg, Carol E.
Walton, Karen L.
White, Fredrick J.
Williams, Linda A.
Wilson, Michael J.