VISIONS OF MAPS AND GRAPHS

William Kruskal
University of Chicago

Cartography and statistics have much in common and many bonds beyond the presence of a statistician as speaker at your excellent banquet. In general terms, cartography and statistics both deal with the compression or distillation of complex quantitative data sets into packages comprehensible to human eyes and minds. Another general similarity is that both statistics and cartography are much concerned with the efficient design of investigations.

The two disciplines have long historical links. Let me read you a short quotation from G. N. Clark's 1948 book, Science and Social Welfare in the Age of Newton. Under discussion are the advances in cartography in England during the late sixteenth century; Professor Clark says that

> "This new cartography was used in the service of the state, for instance by the great Cecil in Elizabethan England, who employed more than one map-maker. In all its aspects it was closely allied to statistics. The estate maps often had in the corners tables of the amounts of land held by different owners, or the numbers of beasts they had a right to pasture. The county maps of Saxton, the most notable of Cecil's cartographers, were accompanied by descriptions which gave figures. This was no accident but arose from the nature of the two methods. A map is an abstract statement based on measurement; statistics are abstract statements based on measurement, counting, and calculation. If this appears to be a farfetched identification, let it be remembered that two of the pioneers of statistical science, Petty and Gregory King, were surveyors before they were statisticians. Neither of them ... got very far away from the geographical point of view."

Editor's Note: Dr. William H. Kruskal, Ernest DeWitt Burton Distinguished Service Professor of Statistics, and Dean of the Division of the Social Sciences at the University of Chicago, delivered the address at the closing banquet of Auto-Carto II. A professor of statistics at the University of Chicago since 1950, Dr. Kruskal has received numerous awards and fellowships in this field. He was President of the Institute of Mathematical Statistics (1970-71) and Vice President of the American Statistical Association (1972-74). He is a member of the American Academy of Arts and Sciences, the International Statistical Institute, the Royal Statistical Society, the American Mathematical Society, the Mathematical Association of America and the Biometric Society, just to name a few of his many affiliations. In his address, Dr. Kruskal explored the relationship between cartography and statistics. *Preparation of his paper was partly facilitated by the National Science Foundation, Grant Number SOC 72-05228 A03.*

Modern connections between the two fields are superbly discussed by Brian J. L. Berry in his 1968 article on "Statistical Geography" in the International Encyclopedia of the Social Sciences. I mentioned three specific connections that have had special interest for me.

First, there is in geography a literature on the many possible ways of describing centers—centers of land mass, of population distribution, or whatever—with lengthy discussion of the merits of various proposals. This literature is much like the statistical literature on what used to be called "measures of central tendency," indexes to give some sensible idea of where the middle of a distribution might be. In the case of a two-dimensional distribution, perhaps on the surface of a sphere, we are essentially in the geographer's position.

The utility of such summary measures presumably comes from the facilitation of comparisons that flows from them; we have all, I expect, seen maps of how the center of gravity of the U. S. population has moved westward with the years.

Any one measure, like the center of gravity, may have disadvantages in particular contexts. A problem that has interested me—and also some geographers—is whether there is a sensible analogue to the median for bivariate distributions; "sensible" includes the requirement that rotating the coordinate system leaves the median-like point unchanged. So, in particular, taking the ordinary univariate medians along two conventional axes does not satisfy that condition.

Second, the development of least squares methods, including nonlinear regression, owes a great deal to the motivation and energies of nineteenth century cartographers, geodesists, and surveyors. From the very start, least squares theory was intertwined with difficult calculations in astronomy, terrestrial magnetism, and map-making. The issues discussed were both practical and theoretical.

Charles Sanders Peirce, perhaps best known for his philosophical thought, was, I believe, exposed to statistical-cartographic problems at the Coast and Geodetic Survey; he was an early, thoughtful, and inadequately appreciated American statistician.

I am told that continuing needs for handling redundant, not fully consistent, data hold today even with the most up-to-date technologies, especially if one includes whatever cartography is called as it travels out to space along with rockets and satellites.

Third, there are contexts in which cartography and statistics are so intertwined that separation is impossible. A recent article by Robert Hoover and others (1975), for example, uses U.S.A. maps by counties, with specific cancer mortality rates shown by shading or color. From these, clusters of counties are chosen for comparison with patterns of industrial concentration.

After this brief description of some links between statistics and cartography, I turn to another that leads to a main theme of these comments. That theme is the importance of empirical experiments dealing with what gets communicated by maps on the one hand, and by statistical graphics on the other.

I cannot speak for cartography with special authority, but for statistical graphics I may safely say that we are at a primitive state: in choosing, constructing, comparing, and criticizing graphical methods we have little to go on but intution, rule of thumb, and a kind of master-to-apprentice passing along of information. You need only look at a good text on statistical graphics. Much of its advice will be excellent, no doubt, but it will also be dogmatic or arbitrary, in the sense that there is neither general theory nor systematic body of experiment as a guide. What we have instead are accumulated experiences, social conventions, and prescriptions. Actual practice in statistical graphics often does not rise to the level expounded by good texts: witness, for example, graphs in otherwise excellent scientific journals that show curves faired through observed data points, but not the data points themselves; or again, graphs of economic statistics in the daily press--even in the Reports of the Council of Economic Advisors--in which choices of origin and scale appear to have been made so as to magnify or diminish, in a self-serving way, some behavior of the data. I must immediately qualify that statement, for it is always dangerous, and perhaps ultimately evil, to ascribe motivation when one sees error. Even, with all good will, honesty, and candor humanly possible, however, can one hope to present graphical material without distortion? What's more, how do we know when distortion occurs? Even if we knew, it would doubtless vary from viewer to viewer: a graphical display honestly presenting material to those of us at this conference might be quite misleading to a conference of lawyers, of automobile salesmen, or of musicians.

My initial theme then is to deplore the paucity of empirical experiments on statistical graphics. Yet there have, of course, been such experiments ... there is a literature. Together with Ian Eggleton, a doctoral candidate at the Graduate School of Business of the University of Chicago, I have been exploring the literature of empirical work in statistical graphics, and also in cartography, for the boundary between the two is fuzzy and work in one may well inform the other.

Our review to date suggests a number of serious shortcomings in the literature of empirical investigations on statistical graphics and cartography. I must apologize in advance of the shortcomings listed really reflect the inadequacy of our search.*

First, the literature appears sketchy in a number of senses. There is relatively little systematic, cumulative research in specific topics. For example, the early work by F. E. Croxton and others in the 20's and 30's on the comparative perception of different shapes for symbols in statistical graphics has apparently received little further attention. Yet these early results were inconclusive and the problem is of continuing importance.

Exceptions to my critical remark are the 1953 paper by J. R. MacKay on cubic symbols and the 1959 paper by John I. Clarke on statistical map reading. On the whole, however, we sense a tendency to introduce new sorts of symbols, perhaps highly imaginative and interesting, but without systematic empirical investigation. Examples are the superimposed symbols of H. R. Wilkinson (1967), the six pointed stars of D. R. MacGregor (1967), and Roberto Bachi's graphic rational patterns (1968). Professor Bachi's book calls for careful experimentation, but I do not think that it reports any; I hope that he will tell me of great work done since or planned soon.

Sustained attention has, it appears, been given to graded series of shadings for maps, and we took special interest in the 1961 paper by G. F. Jenks and D. S. Knos.

---

*Further search has brought us in touch with a few more careful empirical studies; in particular, I cite with admiration the 1969 monograph by Henry W. Castner and Arthur H. Robinson.

Second, we sense generally inadequate attention to the psychological aspects of empirical trials and their reporting. Many papers are remarkably silent about the population of subjects, the experimental instructions and conditions, the exact nature of the stimuli and instructions, etc., all those kinds of information necessary for peer criticism, replication, and extension.

To illustrate the importance of full reporting, I cite a recent paper by Judy Olson (1975) on the subject of map complexity. To its credit, this paper is relatively detailed; in particular, it is reported there that two similar instructions to subjects gave rise to different results, a puzzling and possibly valuable finding.

Third, the papers we have found show relatively little evidence of close cooperation with experimental psychologists concerned with perception and cognition. A few papers show awareness of the psychological literature, but we have found little towards the construction of a psychologically informed theory or of specific hypotheses for future testing.

Fourth, serious questions come to mind about the external validity, so-called, of much of the experimental work to date. I.e., how broadly do the results apply? Subjects seem to be narrowly drawn, often traditionally from student populations, and the effect of training is often neglected. (When training is taken into account it may be important, as was shown by MacKay in 1953.) Stimuli may be stylized or schematic, as in the recent work by Olson with checker-board-like patterns, and she raises herself the question of generalizability.

I must say that statisticians doing this sort of experiment on, for example, visual fitting of lines to data, or numbers of intervals for histograms, are subject to similar criticisms, so I come as a fellow sinner. Surely part of the problem is that of arousing the interest of psychologists, to whom our concerns may appear special and technological. On the other hand, one might hope to elicit their concern, since they themselves use, or should use, statistical graphics in their own research and teaching.

Now surely putting procedures and practices to empirical check is pragmatic and scientific. Let me mention three examples to illustrate this next theme. The first is from an unrelated area in which I have a personal interest: swimming. When I first learned to do the crawl, as a small boy, I was told how important it was to keep the fingers together. Dutifully I accepted that dictum and by now the habit is ingrained; it would be distinctly uncomfortable for me to separate my fingers while swimming.

James Counsilman (1968; pp. 9-12), swimming coach at the University of Indiana, decided to put this established wisdom to empirical test--along with other swimming traditions--by means of special apparatus. It turns out that there is hardly any difference between fingers together and fingers apart! What is the world coming to? Next thing you know, someone will announce that peppermint ice cream is healthier than spinach.

My second example is medical; it is one of many similar examples described and analyzed by Dr. Thomas Chalmers, head of Mt. Sinai Medical School in New York. There is a surgical operation called the portocaval shunt that is traditionally carried out on cirrhotic alcoholics on the basis of plausible arguments about prolongation of life. Yet no proper experiment--with control group and randomization--had been carried out until a few years ago. When the experiment was done, lo and behold, it did not at all prolong life for alcoholics--quite the contrary, it had a shortening effect on average. See Chalmers (1970) and Grace et al. (1966).

(Of course it might be that some categories of patients would benefit from the operation. It is, as we all know, in the discovery of such relevant strata that much scientific advance resides.)

My third example is closer to the statistical tone of this talk; I draw it from the valuable writings of Amos Tversky and Daniel Kahneman (1974). Tversky and Kahneman have been concerned with many aspects of how people actually deal with probabilistic concepts, as opposed to how they might ideally deal with them. In particular, they were curious about the extent to which increasing sample size is perceived as tending to bring a sample average closer to the population average. Among a number of imaginative experimental procedures, they presented subjects with a carefully worded description of two hospitals, one of which has about 15 births a day and the other about 45. Then they asked, for example, about the fraction of days one would expect the percentage of boy babies to be greater than 60%. It turned out that there was hardly any difference in reported fractions greater than 60% despite the considerable difference in sample size, and hence in tightness of distribution of the newborn sex ratio in the two hospitals.

In fact, over half the respondents said that the two hospitals would have the same fraction of days with more than 60% boy babies. Yet it seems clear to us that the smaller hospital, with its higher variability, is bound to have more days with boy babies in excess of 60%.

Reading the Tversky-Kahneman article is a chastening experience for me; what is the point in research and teaching of advanced statistical theory and practice when the general public exhibits such complete ignorance? A discussion of that question would inevitably lead to issues like those that arise in discussing statistical graphics and maps, especially for distribution to wide audiences. I suppose that issues like this are a recurrent topic within government statistical agencies.

To summarize the theme to date then: empirical testing of graphic and cartographic methods is important. That testing should bring in the competences of good psychologists, and it should be based on proper experiment designs. I might add the hope that such activity would lead to usable perceptual theory for the future. After all, it is impossible to test everything empirically; and it is impossible to draw subjects from all groups. Thus ultimately we rely on a combination of experiment and intuition cum experience cum theory. As someone said--I wish I know who it was--, "There is nothing so practical as a good theory."

Next I'd like to discuss briefly the role of statistical graphics within statistics generally. That role has had tremendous ups and downs: at one time, graphical methods were near the core of statistics--Karl Pearson devoted considerable attention to graphics and he was following the emphasis of his hero, Francis Galton. Later on, statistical graphics became neglected and even scorned in comparison with the blossoming of the mathematical side of statistics. In recent years, however, there has been a renaissance of concern with graphics and some of our best statistical minds have suggested new graphical approaches of great interest.

For example, John Tukey (1976) has made highly imaginative suggestions to improve some of the simplest graphical devices: I think in particular of his so-called stem and leaf way of tallying observations and simultaneously producing a rough histogram; I think too of his hanging histogram suggestion, in which the histogram bars hang from an approximating curve rather than poke up towards it.

Several recent suggestions deal with the fundamental problem of exhibiting graphically more than two or three numerical variables at once. Here in the Bureau of the Census, for example, there is current work with color to gain new ground in this direction. A completely different approach has been taken by Edgar Anderson (1957) with his so-called metroglyphs: circles to represent points on a chart or map, each with four or five lines sticking out to represent other variates.

George Barnard, an eminent British statistician, suggests (1969) starting out with a two-dimensional perspective drawing of a surface of y as a function of x_1 and x_2; then, by showing that drawing on a moving picture or television screen and letting it move, one can introduce a third independent variable x_3. Indeed he suggests adding another variable x_4 by shifting x_3 in a relatively slow cycle and x_4 in a relatively fast one, thus presenting the viewer with a pictured surface that heaves (for x_3) and quivers (for x_4). Whether this has been tried, and how practicable it is, I do not know.

Perhaps less limited in dimensionality is an idea discussed by David Andrews (1972): one lets each numerical coordinate determine a coefficient of a finite Fourier series, and then looks at the resulting graphs with their bumps and wavinesses.

One of the most dramatic suggestions is that of Herman Chernoff (1973), who starts from the fact that human beings are remarkably good at recognizing, remembering and discriminating among the faces of other human beings. So Chernoff suggests using schematic faces: the first coordinate might provide degree of ovalness of the face, the second, interocular distance, etc. I have seen computer drawn faces of this kind for some 10 or 12 variates, and they are fascinating to work with. The cartoon-like faces make one laugh at first, and then one takes them seriously.

Note that most of these suggestions depend, for their practical effectiveness, on high speed computers or on television and motion picture technology. So be it, although I fear the decrease in objectivity that may result from the growth of technological graphics.

Ordinary graphics and traditional cartography carry along their own opportunities for distortion, whether conscious or not. We all know the standard cases, for example, misleading scales that make statistical graphs dishonest; in cartography, purposive choice of a projection or of a color can make the red threat look bigger or smaller. Vision is notoriously subject to emotion and predisposition. Just consider the famous, or infamous, canals of Mars.

Yet the addition of <u>motion</u> to graphic or cartographic displays may permit far more extensive distortion and departures from objectivity. Let me read you a passage written by Pauline Kael, motion picture critic for <u>The New Yorker</u> magazine,

> ". . . it is perhaps the most spectacular example of agitprop moviemaking so far, and it demonstrates in a classic way the problems that seem to be inherent in propaganda movies. It is painfully affecting, since it shows the diseases and miseries of the poor, but it is also upsetting and maddening, since it throws facts and figures at us that we cannot evaluate while we're watching it, and calls for revolution as if the case for it had been made on plain, objective grounds." <u>The New Yorker</u>, 6 March 1971.

I note also that — to my knowledge — there has been little serious psychological experimentation on the characteristics of any of the above suggestions.

How stable, for example, are the results of intuitive clustering with Chernoff-like faces under permutation of the coordinates? How much do any of these methods depend on practice and experience?

Chernoff (1975) himself, together with M. Haseeb Rizvi, have carried out just such an experiment. I should also again cite Castner and Robinson (1969).

I would like to push these themes, with your indulgence, in a direction orthogonal to the main thrust of this conference, and yet not one that is absolutely irrelevant.

Earlier I suggested that statistical graphics was in a primitive state, that it was far more an art or a craft than a science. That is even more true for the making of statistical tables. Yet tabulation is the traditional communication mode for the Census and other government statistical agencies. Some agencies, like the Census, are — it seems to me — very good at statistical tabulations and with long experience. Some other agencies, with much less experience, have published tables that are dreadful . . . unreadable, too many figures, poor or missing legends, lots of broken numerals, poor use of white space, and so on.

Yet when I say that Census tables are very good — and I say that with the contented thought that absolute honesty and common courtesy here go hand in hand — when I say that Census tables are very good, I neither know exactly what I mean nor can I point to a body of extensive, careful empirical work to document the assertion. There are publications that give conventional rules for good table making, but these tend to be dogmatic and conventional. There are other books, for example, Hans Zeisel's Say it with Figures, that go much deeper, yet not to first principles as regards tables themselves. There is, to my knowledge, at most a handful of pilot experiments. For example, Andrew Ehrenberg in London has carried out some empirical trials, including comparisons between tables and graphs. (He comes out in favor of tables, to my surprise, but I am not sure how widely his results can be generalized.)

All the problems one has with graphical materials are there for tables: What are the criteria of honesty, clarity, cost, insightfulness? How can one think about variability of readership or of use? The table that a Census demographer will take in almost at a glance would take my youngest son two hours to understand — if then. To what extent can one generalize from results on one or two kinds of table, with one or two kinds of reader or viewer?

Now let us move a little further along this road to the text, the ordinary prose text, of a statistical — or a cartographic — report. Some authors seem to be much better at writing clear, communicative, interesting prose than others. Surely those are desirable characteristics, at least so long as clarity and interest are not used in the service of meretricious propaganda.

I recently ran across a relevant quotation from Tom Margerison (1965); he says "Report writing, like motor-car driving and love-making, is one of those activities which almost every Englishman thinks he can do well without instruction. The results are of course usually abominable." That's a funny, but perhaps an arrogant remark.

How do we really know what gets communicated, and what prose devices are superior to others? Who does careful empirical work on scientific prose — much less careful theory?

Here there are, I believe, starts towards the building of an empirical base: psychologists, linguists, and others have, I understand, been hard at work to develop empirical knowledge and some relevant theory about language. Presumably we have a long way to go, and I doubt that we will ever be able to put some tables in a computer, push a button, and have a finished standardized report come tumbling out.

Research into all these modes of display, exposition, and communication is enormously difficult. As Vincent Barabba says in a recent Census publication, no one knows how decision makers actually use or fail to use a graphic display . . . and I might add a numerical table or a prose exposition. Nor does anyone seem to know how scientists use graphs, tables, etc. for insight. One can safely venture the generalization that people use materials differently. Harold Lasswell said that

> ". . . a trained imagination is necessary before one can perceive with full vividness the significant events referred to in a table of figures, a map, or a chart. Our perceptions of current and past events are facilitated by the context provided by the concreteness of news stories, anecdotes, and personal observations. By contrast the charts, graphs, and tables that refer to the future lack support. This is a problem especially for nonspecialists, since, if laymen are to grasp the meaning of a technical communication, they must relay upon equivalencies with common experience." Lasswell (1959), p. 105.

I began with the title "Visions of maps and graphs" and I intended that in a double sense. First, I look ahead to great improvements in statistical graphics, cartography, tabulations, and other modes of quantitative communication. So that is one sense of vision. And second I think that such improvements must be founded in better knowledge of human perception and cognition, that is to say, if you will forgive the metonymy, a better knowledge of vision.

REFERENCES

Edgar Anderson, 1957, "A Semigraphical Method for the Analysis of Complex Problems," Proceedings of the National Academy of Science, 13, 923-7. Reprinted, with an appended note, in Technometrics 2 (1960) 387-391.

D. F. Andrews, 1972, "Plots of High Dimensional Data," Biometrics, 28, 125-36.

Roberto Bachi, 1968, Graphical Rational Patterns, Israel Universities Press, Jerusalem.

George Barnard, 1969, "Summary Remarks," pp. 696-711 in Norman L. Johnson and Harry Smith, Jr. (Editors), New Developments in Survey Sampling, Wiley-Interscience, New York.

Brian J. L. Berry, 1968, "Statistical Geography," International Encyclopedia of the Social Sciences, Macmillan and Free Press, New York.

Albert D. Biderman, 1971, "Kinostatistics for Social Indicators," Educational Broadcasting Review, 5, 13-19.

Henry W. Castner and Arthur H. Robinson, 1969, Dot Area Symbols in Cartography: The Influence of Pattern On Their Perception, American Congress on Surveying and Mapping, Cartography Division, Technical Monograph No. CA-4.

Thomas C. Chalmers, 1970, "A Clinician Looks at the Role of Statistics in Clinical Research," Symposium on Statistical Aspects of Protocol Design, 9-10 December (National Cancer Institute), San Juan, 5-12.

Herman Chernoff, 1973, "Using Faces to Represent Points in K-dimensional Space Graphically," Journal of the American Statistical Association, 68, 361-8.

Herman Chernoff and M. Haseeb Rizvi, 1975, "Effect on Classification Error of Random Permutations of Features in Representing Multivariate Data by Faces," Journal of the American Statistical Association, 70, 548-554.

G. N. Clark, 1948, Science and Social Welfare in the Age of Newton, Oxford University Press, London. The quoted passage is from page 19 of a reprinting of the chapter, "Social Science in the Age of Newton," in pp. 15-30 of Anthony Oberschall (Editor), The Establishment of Empirical Sociology, Harper and Row, New York, 1972.

John I. Clarke, 1959, "Statistical Map-Reading," Geography, 44, 96-104.

James E. Counsilman, 1968, The Science of Swimming, Prentice Hall, Englewood Cliffs, N.J.

N. D. Grace, H. Muench, and T. C. Chalmers, 1966, "The Present Status of Shunts for Portal Hypertension in Cirrhosis," Gastroenterology, 50, 684-691.

Robert Hoover, Thomas J. Mason, Frank W. McKay, and Joseph F. Fraumeni, Jr., 1975, "Cancer By County: New Resource for Etiologic Clues," Science, 189, 1005-7.

George F. Jenks and Duane S. Knos, 1961, "The Use of Shading Patterns in Graded Series," Annals of the Association of American Geographers, 51, 316-334.

Harold D. Lasswell, 1959, "Strategies of Inquiry: The Rational Use of Observations," pp. 89-113 in The Human Meaning of the Social Sciences, Daniel Lerner (Editor), World Pub., Cleveland.

D. Ronald MacGregor, 1967, "The Mapping of Industry," International Yearbook of Cartography, 7, 168-185.

J. Ross MacKay, 1953, "A New Projection for Cubic Symbols on Economic Maps," Economic Geography, 29, 60-62.

Tom Margerison, 1965, Review of Writing Technical Reports, by Bruce M. Cooper, Sunday Times, 3 January. Quoted on p. 49 of A Random Walk in Science, R. L. Weber (compiler) and E. Mendoza (Editor), Institute of Physics, London, and Crane, Russak, New York, 1974.

Judy M. Olson, 1975, "Autocorrelation and Visual Map Complexity," Annals of the Association of American Geographers, 65, 189-204.

John W. Tukey, 1976, Exploratory Data Analysis, Addison-Wesley, Reading, Massachusetts.

Amos Tversky and Daniel Kahneman, 1974, "Judgment Under Uncertainty: Heuristics and Biases," Science, 185, 1124-31.

H. R. Wilkinson, 1967, "Regional Features in Yugoslavia," in R. W. Steel and R. Lawton (Editors), Liverpool Essays in Geography, A Jubilee Collection, Longmans, London.