STATISTICAL GRAPHICS: Mapping the Pathways of Science

Howard Wainer

Educational Testing Service, Princeton, New Jersey 08541; e-mail: hwainer@ets.org

Paul F. Velleman

Cornell University, Ithaca, NY 14853; e-mail: pfv2@cornell.edu

■ **Abstract** This chapter traces the evolution of statistical graphics starting with its departure from the common noun structure of Cartesian determinism, through William Playfair's revolutionary grammatical shift to graphs as proper nouns, and alights on the modern conception of graph as an active participant in the scientific process of discovery. The ubiquitous availability of data, software, and cheap, high-powered, computing when coupled with the broad acceptance of the ideas in Tukey's 1977 treatise on exploratory data analysis has yielded a fundamental change in the way that the role of statistical graphics is thought of within science—as a dynamic partner and guide to the future rather than as a static monument to the discoveries of the past. We commemorate and illustrate this development while pointing readers to the new tools available and providing some indications of their potential.

CONTENTS

INTRODUCTION: Graphs as Nouns, from Common to Proper

Graphic displays abounded in ancient times. For example, a primitive coordinate system of intersecting horizontal and vertical lines that enabled a precise placement of data points was used by Nilotic surveyors as early as 1400 BC. A more refined coordinate system was used by Hipparchus (ca. 140 BC), whose terms for the coordinate axes translates into Latin as *longitudo* and *latitudo*, to locate points in the heavens. Somewhat later, Roman surveyors used a coordinate grid to lay out their towns on a plane that was defined by two axes. The *decimanus* were lines running from east to west, and the *cardo* ran north to south (Smith 1925). There are many other examples of special-purpose coordinate systems in wide use before the end of the first millennium; musical notation placed on horizontal running lines was in use as early as the ninth century (Apel 1944); the chessboard was invented in seventh century India.

Costigan-Eaves & Macdonald-Ross (in preparation) found what appears to be one of the earliest examples of printed graph paper dating from about 1680. Large sheets of paper engraved with a grid were apparently printed to aid in designing and communicating the shapes of the hulls of ships.[1] Both Beniger & Robyn (1978) and Funkhouser (1937) describe Descartes' 1637 development of a coordinate system as an important intellectual milestone in the path toward statistical graphics. We join Biderman (1978) in interpreting this in exactly the opposite way—that it was an intellectual impediment that took a century and a half and William Playfair's (1759–1823) eclectic mind to overcome.
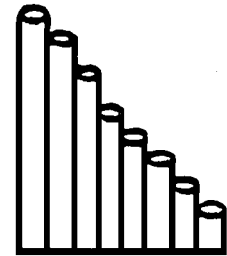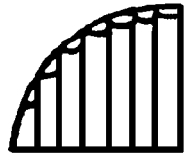
Because natural science originated within natural philosophy, it favored a rational rather than empirical approach to scientific inquiry. Such an outlook was antithetical to the more empirical modern approach to science that does not disdain the atheoretical plotting of data points with the goal of investigating suggestive patterns. Graphs in existence before Playfair (with some notable exceptions discussed below) grew out of the same rationalist tradition that yielded Descartes' coordinate geometry—that is, the plotting of curves on the basis of an a priori mathematical expression (e.g. Orseme's "pipes" on the first page of the Padua edition of his 1486 *Tractatus de latitudunibus formarum*[2] is often cited as an early example; see Figure 1).

This notion is supported by statements like that of Luke Howard, a prolific grapher of data in the late eighteenth and early nineteenth century who, as late as

---

[1]This material is classed in the "collection" category of the British Library with the entry, "A collection of engraved sheets of squared paper, whereon are traced in pencil or ink the curves or sweeps of the hulls of sundry men-of-war."

[2]Clagett (1968) argued convincingly that this work was not written by Oresme, but probably by Jacobus de Sancto Martino, one of his followers, in about 1390—yet another instance of how surprisingly often eponymous referencing is an indication only of who did not do it (Stigler 1980).

difformis vniformiter variatio reddit vnifor
miter diffoimiter difformes). ¶ Latum: vni
formin i oitioris e illa q̃ inf excellus graduuz
eq̃ oustatuz huat eade ꝓportoz e.ia ni a ꝓ
portoe eq̃litatis. Na ii inf excellus graduu z
inter ie eq̃ oustanuu huarent ꝓportoz eq̃nta
tis .uc cet .autm: vniformit oitheris ut p; ex
oiffinitionibus membroium secude ouuliois
Rurfus ii nulla proporcio feruat tuic nulla
poffet attendi vniformitas in latundine tali z
fic non effet vniformiter oihoim i oifformis
¶ Lat.m: oifformiter o.fformuter oifformis
e illa q̃ inter excellus graduu eque oiftantiuz
non ienuat eandem proportionem ficu.'in fe
cunda parte patebit. Notandum tame:n eft
ꝗ ficut in fupradictis oiffinitoib⁹ ubi loqtur
oe excello graduum inter fe eque oiftantium
oeb; accipi oiftancie fcom partes latitudinis
extefine z nó intefiue tita ut loquunt o.c:e oif
finitões o oiftatia z̃ duu fimali ii aut graduali

o.ffonc oiffona

of; of, o.ttofis

fdp.z fial ad iiz

Equit .feda ps in qua ut
fupradicta intelligautur ad
fenluz per figuras gecme
tricas oftenduntur. Et ut
omnem fpeciem latundis
in prefenti materia via oc
currat apparentior i.titudics ad figuras geo
metcaz e)plicant. Ifta ps oiuidit p tria ca
pitula q̃ ꝭ p̃ ꝛtmet o.onca. z̃ᵐ fuppofitio. s

Figure 1 Oresme's graphical illustration of functions taken from the first page of the Padua edition of his 1486 *Tractatus de latitudunibus formarum*. (British Library IA. 3Q024)

1847, apologized for his methodology and referred to it as an "autograph of the curve ... confessedly adapted rather to the use of the *dilettanti* in natural philosophy than that of regular students" (Howard 1847, p. 21).

It is not inaccurate to think of early graphic displays as nouns, indeed common nouns, that were used to depict some theoretical relationship. Thus, we can conceive of the first major revolution in the use of graphic display in science as a shift from its use as a common noun (e.g. the theoretical relationship between supply and demand) to that of a proper noun (e.g. England's imports and exports from 1700 to 1800). This revolution seems to have begun in 1665 with the invention of the barometer. This inspired Robert Plot to record the barometric pressure in Oxford every day of 1684 and summarize his findings in a remarkably contemporary graph (Figure 2) that he called a "History of the Weather." He sent a copy of this graph with a letter to Martin Lister[3] in 1685 with a prophetic insight on the eventual use:

> For when once we have procured fit persons enough to make the same Observations in many foreign and remote parts, how the winds stood in each, at the same time, we shall then be enabled with some grounds to examine, not only the coastings, breadth, and bounds of the winds themselves, but of the weather they bring with them; and probably in time thereby learn, to be forewarned certainly, of divers emergencies (such as heats, colds, dearths, plague, and other epidemical distempers) which are not unaccountable to us; and by their causes be instructed for prevention, or remedies: thence too in time we may hope to be informed how far the positions of the planets in relation to one another, and to the fixed stars, are concerned in the alterations of the weather, and in bringing and preventing diseases and other calamities...we shall certainly obtain more real and useful knowledge in matters in a few years, than we have yet arrived to, in many centuries. (Plot 1685)

Plot and Lister's use of graphic display was scooped by the seventeenth century polymath Christiaan Huygens (1629–1693). On October 30, 1669, Christiaan's brother Lodewijk sent him a letter containing some interpolations of life expectancy data taken from John Graunt's 1662 book *Natural and Political Observations on the London Bills of Mortality*. Christiaan's responded in letters dated November 21 and 28, 1669, with graphs of those interpolations (Huygens, 1895). Figure 3 shows one of those graphs, with age on the horizontal axis and number of survivors of the original birth cohort on the vertical axis. The curve was fitted to his brother's

---

[3]The origin of the graphic depiction of weather data sadly, for the obvious eponymous glory, rests not with Plot but rather with Lister, who presented various versions of graphical summaries of weather data before the Oxford Philosophical Society on March 10, 1683 and later in the same year a modified version to the Royal Society. Plot was not the only one enthusiastic about Lister's graphical methods. William Molyneux was so taken that he had an engraving made of the grid and he would faithfully send a "weather diary" monthly to William Musgrave. One of Molyneux's charts was reproduced in Gunther (1968).
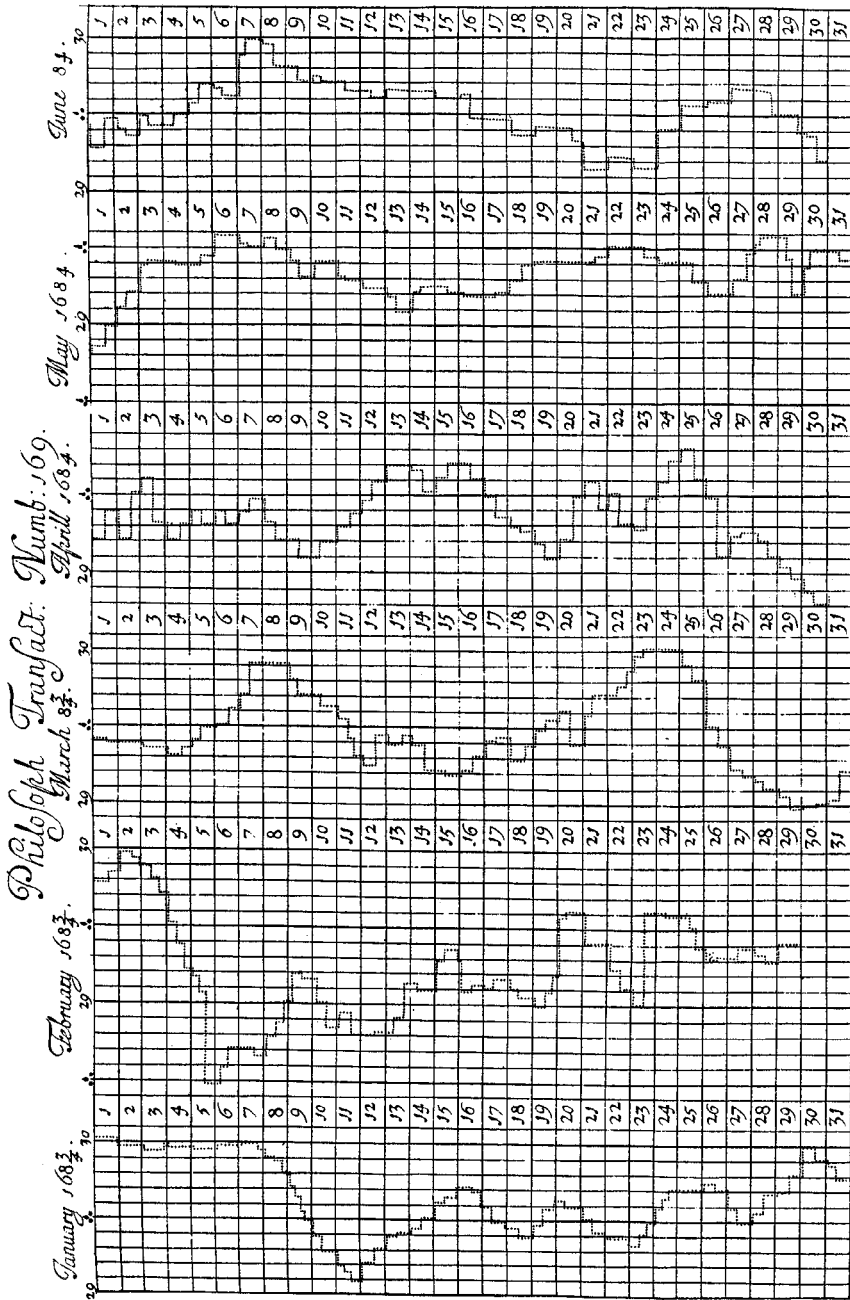
Figure 2 Robert Plot's (1685) "History of the Weather" recording of the daily barometric pressure in Oxford for the year 1684 (based on the original work of Martin Lister).
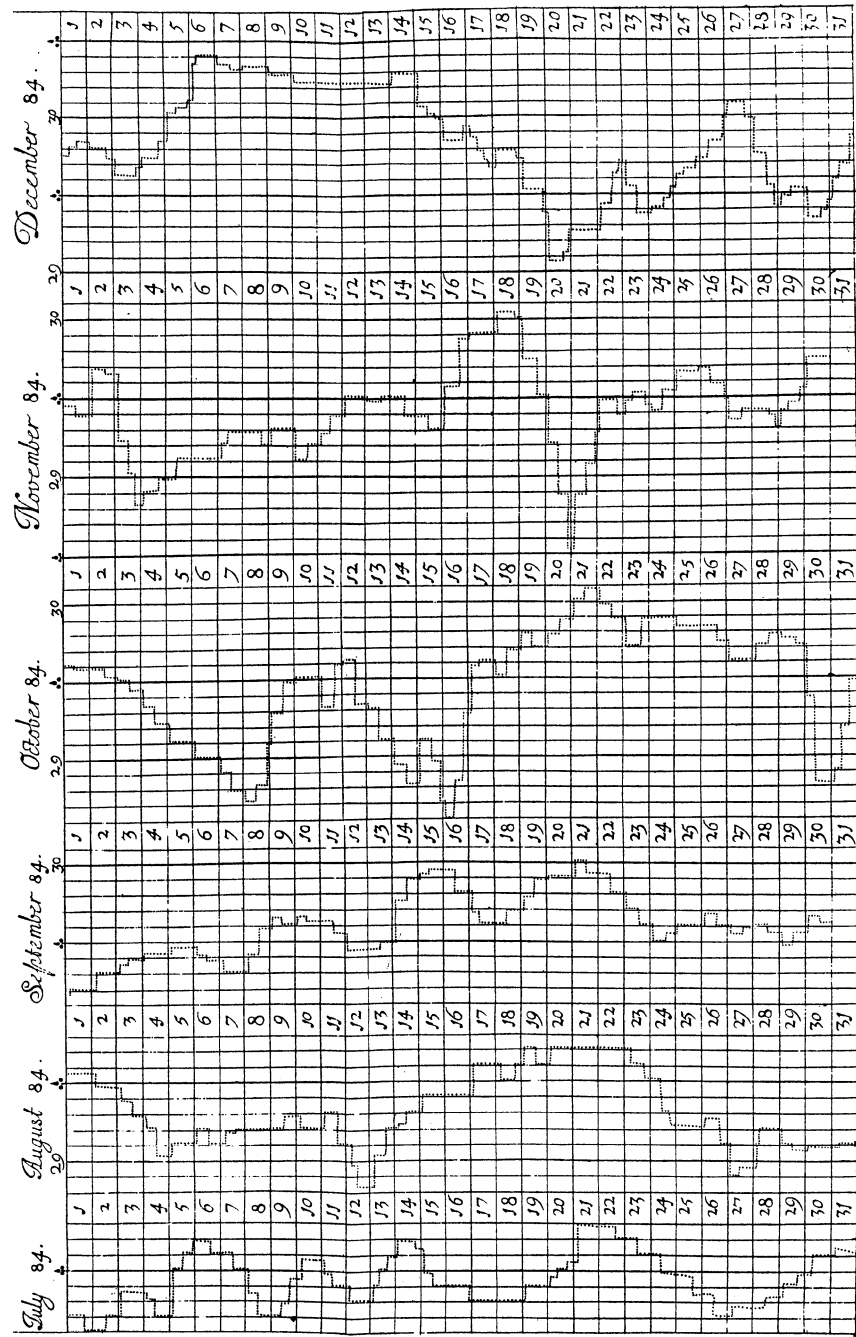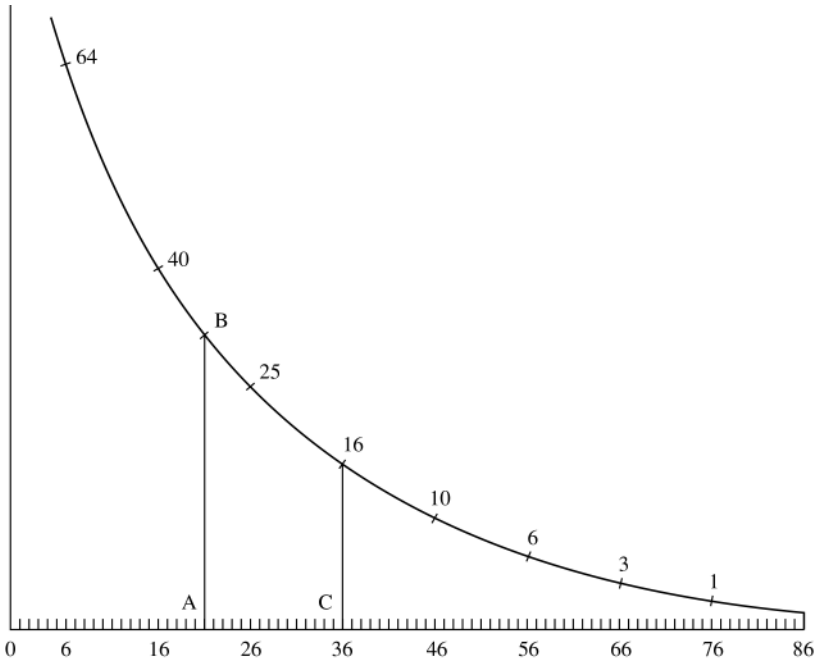
Figure 2 (*Continued*)

Figure 3   A redrafting of Christiaan Huygens' 1669 curve showing how many people out of a hundred survive between the ages of infancy to 86. [Data from John Graunt's (1662) *Natural and Political Observations on the Bills of Mortality*].

interpolations.[4] The letters on the chart are related to an associated discussion on how to construct a life expectancy chart from this one—that is, analyzing a set of data to yield deeper insights into the subject. Christiaan's constructed such a chart and indicated that it was more interesting from a scientific point of view; Figure 3, he felt, was more helpful in wagering.

There was a smattering of other examples of empirically based graphs that appeared in the century between Huygens' letter and the publication of Playfair's *Commercial and Political Atlas* (1786), for although some graphic forms were available before Playfair, they were rarely used to plot empirical information. Biderman (1978) argued that this was because there was an antipathy toward such a use as a scientific approach. This suggestion was supported by such statements as that made by Luke Howard. However, at least sometimes when data were available (e.g. Pliny's astronomical data, Graunt's survival data, Plot's weather

---

[4]There are many other graphical devices contained in the 22 volume *Oeuvres Complètes* (1888–1950) to be explored by anyone with fluency in ancient Dutch, Latin, and French. Incidentally, Huygen's graphical work on the pendulum proved to him that a pendulum's oscillations would be isochronic regardless of its amplitude. This discovery led him to actually build the first clock based on this principle.

data, and several other admirable uses) they were plotted. Perhaps part of the exponential increase in the use of graphics since the beginning of the nineteenth century is merely concomitant to the exponential growth in the availability of data. Of course, there might also be a symbiosis, in that the availability of graphic devices for analyzing data encouraged data gathering. For whatever the reasons, Playfair was at the cusp of an explosion in data gathering, and his graphic efforts appear causal. He played an important role in that explosion.

The consensus of scholars, well phrased by P Costigan-Eaves & M Macdonald-Ross (in preparation), is that until Playfair "many of the graphic devices used were the result of a formal and highly deductive science.... This world view was more comfortable with an arm-chair, rationalistic approach to problem-solving which usually culminated in elegant mathematical principles" often associated with elegant geometrical diagrams. The empirical approach to problem solving, a critical driving force for data collection, was slow to get started. However, the empirical approach began to demonstrate remarkable success in solving problems, and with improved communications,[5] the news of these successes, and hence the popularity of the associated graphic tools, began to spread quickly.

We are accustomed to intellectual diffusion taking place from the natural and physical sciences into the social sciences; certainly that is the direction taken for both calculus and the scientific method. However, statistical graphics in particular, and statistics in general, went the reverse route. Although, as we have seen, there were applications of data-based graphics in the natural sciences, it was only after Playfair applied them within the social sciences that their popularity began to accelerate. Playfair should be credited with producing the first chartbook of social statistics; indeed publishing an atlas that contained not a single map is one indication of his belief in the methodology (to say nothing of his chutzpah). Playfair's work was immediately admired, but emulation, at least in Britain, took a little longer (graphic use started up on the continent a bit sooner). Interestingly, one of Playfair's earliest emulators was the banker S Tertius Galton (the father of Francis Galton, and hence the biological grandfather of modern statistics) who, in 1813, published a multiline time series chart of the money in circulation, rates of foreign exchange, and prices of bullion and wheat.[6] The relatively slower diffusion of the graphical method back into the natural sciences provides additional support for the hypothesized bias against empiricism there. The newer social sciences, having no such tradition and faced with both problems to solve and relevant data, were quicker to see the potential of Playfair's methods.

Playfair's graphical inventions and adaptations look contemporary. He invented the statistical bar chart out of desperation because he lacked the time series data required to draw a line showing the trade with Scotland, and so used bars to symbolize

---

[5]The first encyclopedia in English appeared in 1704. The number of scientific periodicals began a rapid expansion at the end of the eighteenth century; between 1780 and 1789 20 new journals appeared, between 1790 and 1800 25 more (McKie 1972).

[6]Biderman (1978, 1990) pointed out that ironically, Galton's chart predicted the financial crisis of 1831 that created a ruinous run on his own bank.

the cross-sectional character of the data he did have. Playfair acknowledged Priestley's (1769) priority in this form, although Priestly used bars to symbolize the life spans of historical figures in a time line.[7]

Playfair's role was crucial for several reasons, but not for his development of the graphic recording of data; others preceded him in that. Indeed, in 1805 he pointed out that as a child his brother John had him keep a graphic record of temperature readings. However, Playfair was in a remarkable position. Because of his close relationship with his brother and his connections with James Watt he was on the periphery of science. He was close enough to know the value of the graphical methods, but sufficiently detached in his own interests to apply them in a very different arena—that of economics and finance. These areas, then as now, tend to attract a larger audience than matters of science, and Playfair was adept at self-promotion. [For more about the remarkable life and accomplishments of William Playfair (including the fascinating story of his attempted blackmail of Lord Archibald Douglas) the interested reader is referred to Spence & Wainer (1997, 2000), Wainer (1996) and Wainer & Spence (1997).]

In a review of Playfair's 1786 *Atlas*, which appeared in *The Political Herald*, Dr. Gilbert Stuart wrote, "The new method in which accounts are stated in this work, has attracted very general notice. The propriety and expediency of all men, who have any interest in the nation, being acquainted with the general outlines, and the great facts relating to our commerce are unquestionable; and this is the most commodious, as well as accurate mode of effecting this object, that has hitherto been thought of ... To each of his charts the author has added observations (which) ... in general are just and shrewd; and sometimes profound ... Very considerable applause is certainly due to this invention; as a new, distinct, and easy mode of conveying information to statesmen and merchants ..." Such wholehearted approval rarely greets any scientific development. Playfair's adaptation of graphic methods to matters of general interest provided an enormous boost to the popularity of statistical graphics.

THE NEXT GRAPHICAL REVOLUTION: Graphs as Dynamic Colleagues

> "*Eppur si muove!*"[8]
> Galileo (c. 1622)

For almost 200 years, from 1786 and the publication of Playfair's *Atlas* until 1977 and the publication of Tukey's *Exploratory Data Analysis*, the use

---

[7]Priestley's use of the bar as a metaphor is somewhat different then Playfair's in that the data were not really statistical. Moreover, Priestly was not the first to construct a graphical time-line; in 1753 the French physician Jacques Barbeu-Dubourg produced a graphic in the form of a 54 foot long scroll, configured in a way not unlike a torah, that contains thumbnail sketches of famous people from The Creation to 1750 (see Wainer 1998 for a fuller story).
[8]"And yet it moves!"

of graphics within science remained static. Statistical graphics became widely used to communicate information, to decorate and enliven scientific presentations, and to store information. Their use as the principal tool in the exploration of quantitative phenomena also grew in fits and starts, but sentiments, analogous to Luke Howard's were still voiced. Tukey's *Exploratory Data Analysis* changed things. Suddenly terms like data snooping, data dredging, and the currently trendy "data mining" were no longer pejorative.[9]

Coupled with the scientific acceptability, even desirability, of the clever plotting of data points in the search for suggestive patterns, was the ubiquitous appearance of cheap powerful computing. This manuscript is being prepared on a $2000 computer more powerful than any institutional mainframe available when Tukey's book was published. Although most of its MIPS are wastefully idle, they can be called upon whenever needed. However, the computer revolution does not stop with machinery (although it is surely powered by it). Enormous data sets, on varied topics, are readily available. A CD-ROM or two can provide you with the results of the decenniel census or the entire National Assessment of Educational Progress. Through the Cochrane Collaboration the results of 250,000 different random assignment medical experiments are immediately accessible for scrutiny and meta-analysis. Soon all three billion pieces of the human genome will be available to serve as biology's analog to the periodic table. And then there is "the web," overflowing with data (and nondata).

Software for data analysis and visualization when added to the assets of powerful computing and extensive data completes the scientific triumvirate. Studies that were either too expensive, too tedious, or too difficult can now be done with the click of a mouse. It is this ease of manipulation that characterizes the latest transformation of graphics in scientific inquiry. The graph is no longer a static object to be carefully constructed and enshrined for further study. It is a dynamic partner in the investigation.

The rest of this chapter focuses on some of the new dynamic tools that are available for examining data. We ignore the set of useful tools for data exposition that were described 20 years ago in an earlier incarnation of this chapter (Wainer & Thissen 1981) and instead refer interested readers to that review.

There are many more ways to display data badly (Wainer 1997, Chapt. 1), than there are to display data well—that is, to say what you mean about the data clearly and grammatically. Whereas the earlier chapter on graphical data analysis discussed clear, grammatical presentation of data, including methods that are resistant to influence by outliers, the balance of this chapter discusses how to hold a conversation about your data with a data display.

---

[9]Data mining, which usually implies fitting a very complex general model to an enormous data set, still seems to deserve critical scrutiny. Bert Green (personal communication) characterizes data mining as being akin to the Ganzwelt of the nineteenth century psychophysicists; sooner or later you begin to see things, whether or not anything is really there.

The key to conversational graphics is the recognition of a graph as dynamic and malleable. During the course of a good conversation, each party changes, learns, and grows. A good conversation about data is much the same: We may see something new in the data that leads us to want to view it in a new way. By viewing the data in many different but consistent ways, we have a greater chance of noting patterns, relationships, and exceptions. As the conversation leads us to a new point of view, we understand the data differently.

Conversational Graphics

Data graphics have evolved from depicting numbers, to depicting variables (e.g. distributions), to depicting relationships among variables. At each stage, however, the communication has been in only one direction: from the graph (or graph maker) to the viewer. But, as computers have taken over almost all graph drawing for data, we have come to realize the possibility of interacting with graphs, of holding a conversation with a graph in an attempt to mutually achieve greater insight. We have come to realize the extraordinary enhancement that such interaction brings to the understanding of data through graphs. There is good experimental evidence that we learn better through interaction. Such "active learning" is almost a fad among educators, but the principle that interacting with something new aids in understanding is sound.

Graphs that interact with the viewer first appeared in the early 1970s with projects such as PRIM-9 (Fisherkeller et al 1974), the first multidimensional rotating scatterplot and early experiments with plot brushing at AT&T Bell Labs (Becker & Cleveland 1984). It is only with the wide availability of powerful desktop computers, however, that they have become widely available. Various kinds of real-time interaction can be found in many statistics programs, although few offer all of the methods we discuss here. However, each method has usually been discussed on its own. We attempt here to bring together discussions of interactive graphics and provide unifying principles and insights.

The Absurdity of Graphing Data

The Nobel Laureate Eugene P Wigner (1960), in his address commemorating the opening of the Courant Institute, remarked on the unusual effectiveness of mathematics in science. He pointed out that "mathematics works so often in science that it is disquieting. It is like a man with a large key ring and a sequence of doors to open who finds that after choosing a key at random each door opens on the first or second try. Sooner or later you begin to doubt the relationship between the keys and the locks. So it is with mathematics and science." Why should the universe operate in such a way that human mathematics accurately describes it?

It is with the same sense of wonder that we ask the identical question about graphical display, for graphs of data are based on the somewhat absurd notion that we can usefully represent data values whose meaning relates to units of measurement in the real world by arbitrarily assigning them a position in space,

a color, a symbol, or a behavior. Moreover, although the data values themselves have no position, color, symbol, or behavior, an appropriate assignment will not only allow us to perceive patterns and relationships that might not otherwise be evident, but will meaningfully relate to the original measurement units.

Just as the unusual effectiveness of mathematics in science suggests something about the universal truth of mathematics, the unusual effectiveness of graphs for communication with humans suggests fundamental truths about human perception. In his Silliman Lectures, Jacob Bronowski (1978) notes that human perceptual abilities evolved along with our species and are thus optimized for certain survival-enhancing perceptions. We see edges well. We see straight lines and understand their relative slopes easily. We can compare areas and sizes visually unless distracted by an illusion of depth and volume. We are well-equipped to see smooth, physically-appropriate motion and we implicitly understand trajectories.

As a result, Bronowski points out, we see the world the way *we* look rather than the way *it* looks, which constrains what we perceive. Data graphics, however, must take account of how we look and what we will see. Properly designed graphics use human perception abilities wisely. Thus, well-planned layouts, straight lines, starkly different colors, areas of simple shapes, and smooth motion facilitate understanding and perception in graphs.

More generally, modern graphics take advantage of human perception by constraining the points and symbols representing the data to behave with a "cartoon reality" that obeys reasonable laws. These laws include the principle that elements in a graph move smoothly (not jumping from place to place), that they have a consistent color, shape, and selection state, and that the mapping of numeric value to physical plot attribute is consistent and shows an appropriate association (e.g. the well-established "area principle," which holds that the perceived size of a plot element should correspond to the magnitude of the value displayed).

In fact, the wise use of these principles makes it possible for modern statistical graphics to display greater complexity than humans can easily understand otherwise. Well-designed statistical displays enable analysts to understand relationships among four, five, or even more variables—certainly more than three-dimensional (3-D) creatures are usually comfortable manipulating in Cartesian coordinates.

Multiple Dimensions

Traditional graphics are limited by the two-dimensional page or screen on which they appear. It is difficult to display more than two variables, and nearly impossible to display more than three clearly. [The now famous Minard graph depicting Napoleon's disastrous invasion of Russia (see Wainer 1997, p. 64) is remarkable precisely because it surpasses these limits so gracefully.]

The world is not bivariate. The challenges of understanding multivariate relationships makes graphs that can help in this understanding particularly useful.

Time as a Dimension

Because the human eye tracks smooth motion well, motion can be an effective display dimension. Physicists have told us for a century that time should be regarded as a dimension along with the three spatial dimensions. Designers of data graphics have now taken this admonition to heart, although not in the sense that Einstein had in mind. Rather, it is possible to use motion to show how a relationship that has been graphed changes as some other term is modified. The ability of a graph to change in real time, in response to viewer action, can display relationships among variables in ways that are perceived by most viewers as naturally as the mapping of value to physical location on a bivariate plot.

One use of this capability that has become relatively common is the display of three variables in a three dimensional scatterplot, whose structure is displayed by rotating it smoothly on the computer screen. Even though the display in fact shows successive projections of the point cloud on the screen, the illusion of a three dimensional display seen in rotation is compelling.

Another use of such animation is to show a display changing as a parameter is altered. For example, the analyst might control the value of an exponent used in the reexpression of one variable by sliding an on-screen control with the mouse. Simultaneously, a display of the residuals from a regression analysis can be updated, smoothly changing as the reexpression changes. Some animations of this sort show residuals becoming more homoskedastic as an appropriate reexpression is found. Others might show a single data value drifting away from the others and becoming an outlier, vividly revealing the sensitivity of that particular value to the parameter change.

Yet another use of animation shows the relationship among two variables as a third variable is added smoothly to the model or otherwise modified. Such methods display an interaction effect—an aspect of statistical modeling that is notoriously difficult to understand and display, but that nevertheless is of great importance in discerning the truth about multivariate relationships.

To achieve this perceptually comfortable mapping, changes over time must follow their own rules of consistency. Displays must change smoothly and must keep up with mouse-based controls. (A delay of as little as 0.1 second can make the display appear to lag behind the mouse and destroy the illusion of physical reality.) Other rules usually lean toward simplification. For example, despite a number of attempts to simulate three dimensions accurately on a computer screen with perspective and shadow, most viewers are more comfortable with a flat projection of a three dimensional pseudoreality onto the screen, which then moves to show the third dimension. Such a view of the data is much like the view through a telescope at some distance, in which the depth of field is lost. It also corresponds to the mathematical operation of projecting from higher dimensions onto lower dimensions—an operation fundamental to most multivariate statistics.

Such a display sacrifices all cues about the direction of rotation. Some viewers can reverse the illusion, switching the perception that the frontmost points are

moving to the left (and the rearward points to the right) with the perception that the motions are reversed. Interestingly, the two displays are equivalent in data analysis content, so the ambiguity has no important consequences.

Kinds of Interaction

Modern data-display software provides several kinds of interaction with data graphics and some underlying principles that support them. All of these methods assume that what we are seeing shows the data from many points of view and in many different ways but continues to preserve the data's central reality and consistency. The displays observe the principle of "linking," in which multiple arrays of related data are consistent in how the data are displayed, in particular in the use of color, symbol, and highlighting of points. Changes in one view of the data alter all other views simultaneously, preserving the illusion that, for example, the color of a datapoint is the same regardless of how it is viewed.

Selecting is a fundamental operation because selected points stand out from the background of other points. Selected points and regions are usually highlighted by becoming brighter, by becoming slightly larger, by changing color, or by filling in open spaces. The unselected points are displayed as well, providing a context for the selected points. It is thus easy to see whether, for example, the selected points cluster together consistently or show a trend that differs from the background trend.

Linking shows each case consistently across several displays. When a case is selected in one plot, all views of that case are selected immediately and highlighted so that the selection can be seen. The selected case stands out from the other cases in each window, so its relationship to them becomes clearer, making it easy to see conditional relationships. Clusters of points in one display can be selected to see whether they appear as a group in other views of the data or whether the observed clustering is a local feature.

Linking makes it easy to answer questions such as

1. Is this extreme point also extreme in any other view of the data?
2. Do the points in this part of the histogram cluster on other variables?
3. Is the relationship between these two variables the same for each of the groups in this pie chart?
4. Does the pattern shown in this rotating plot correspond to any patterns shown in other views of the data?

These questions require sophisticated and complex statistical calculations to answer numerically but are easy to investigate with linked plots.

More fundamentally, linking treats each case as an object with a graphic reality. Just as real world objects have a shape, location, and color, graphic representations of data values benefit from having a consistent existence. Thus, graphing programs can also link plot symbols and colors. Each case is drawn in all of the plots with

the same symbol (where symbols are appropriate) and in the same color (where colors are possible).

Linking also makes possible the interactive actions *brush* and *slice*. These actions have emerged as fundamental parts of the conversation that data analysts can hold with graphic displays of their data.

Brushing and *slicing* can reveal joint patterns and relationships among many variables. Thus, they are actions appropriate for multivariate analysis.

Plot brushing was developed initially by statisticians at AT&T Bell Labs (Becker & Cleveland 1984) as a way to work with scatterplot matrices and is still offered in that specialized form by some statistics programs. Other programs generalize brushing beyond that isolated framework, making the plot brush a tool that works in any appropriate display.

Brushing focuses attention on a selected subset of points while showing them against the background of the rest of the points. Each kind of display can offer an appropriate way to define the selected subset. The simplest case is brushing a scatterplot in which a rectangle (whose size and shape can be controlled by the analyst) is dragged over a scatterplot controlled by mouse movements. Points covered by the rectangle are highlighted in the scatterplot and in all other displays simultaneously. One can usually define brushes of different sizes and shapes; a tall, thin brush, for example, selects small, local parts of an x-axis variable. The highlighted points in other plots show the patterns and distributions conditional on the selected slice of points.

By contrast, selecting points in a dotplot focuses attention on a subrange of the plotted brushed variable and shows where those points reside in other displays. Such a strip of values in effect, conditions on the selected subrange of the brushed variable, and shows the effects of changing the conditioning.

One can even brush bars in a histogram, watching the corresponding selection in other displays. More subtly, the effects of brushing can link into a histogram. Experience has shown that the best display for this is a highlighted "subset histogram" shown against the background of the full data histogram. By selecting points in a rotating plot, you can orient the rotation to identify a key dimension or to isolate a subgroup.

A *slicing tool* selects points in vertical or horizontal slices of a plot. The tool slices right to left, left to right, top to bottom, or bottom to top, according to its initial direction. In contrast to a plot brush, in which points leaving the brush lose their highlighting, points selected by a slicing tool are selected as the tool passes their position and remain selected unless you reverse direction and drag back over them.

Brushing and slicing help to answer questions such as

1. Do the same cases seem to be in roughly the same places in each plot?

2. Is there any trend in sales from east to west?

3. Which variables change systematically as I move along this principal dimension in a rotating plot?

4. How does the relationship between the gas mileage and weight of cars change as drive ratio increases?

Brushing and slicing are based on the principle that by emphasizing the common identity of cases in multiple displays, we can help analysts relate several displays to one another. They do not add information that is not already in the displays; rather, they provide easier access to that information.

80 Companies Slicing Example   As an example of how slicing can help, consider the scatter plot of Log(Assets) versus Log(Market Value) from 80 companies drawn randomly from the *Forbes* 500 (Figure 4, in which original data were in millions of dollars). We see three interesting features:

1. There is surely a trend of companies with greater market value to have greater assets (see the regression line in Figure 4);

2. There are about seven companies with a market value of about a billion dollars that have lower than expected assets; and

3. There seems to be a string of companies, of varying market values, that have unusually large assets.

It seems sensible to look at the residuals from the overall trend (item 1). These are shown, as a function of predicted value, in Figure 5. We next use our slicing
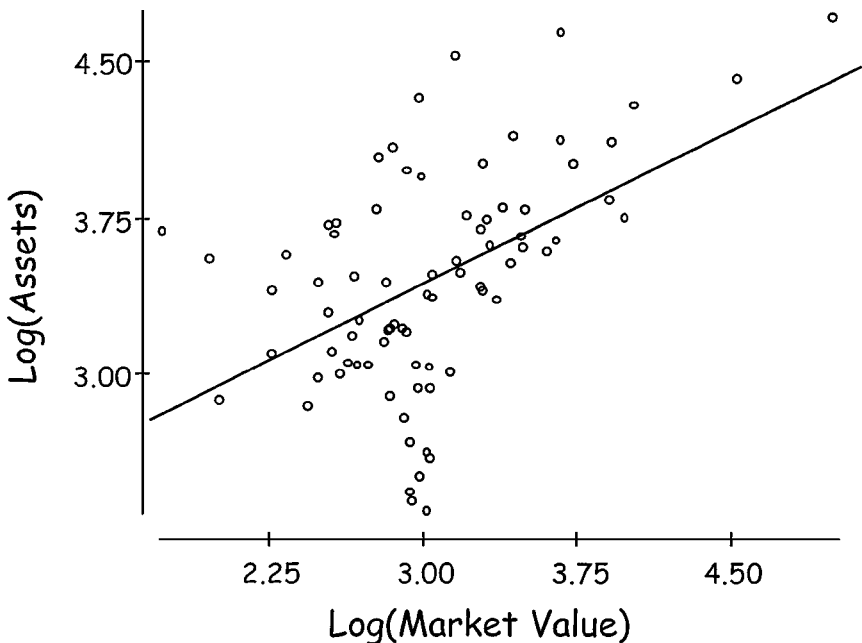

Figure 4   Scatter plot of Log(Assets) against Log(Market Value) for 80 companies drawn at random from the *Forbes* 500. A least squares regression line is drawn in.
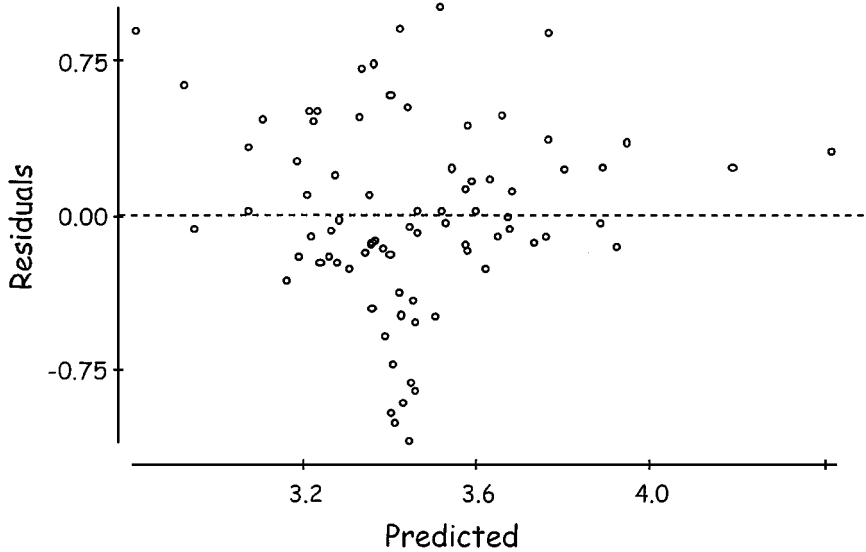
Figure 5   Residuals from the regression depicted in Figure 8 are plotted against their predicted values.

tool to select companies with large positive residuals (Figure 6*A*). The selection tool is indicated by the two horizontal lines and the selected companies are now shaded. As we select these companies a linked bar chart, which shows the number of companies within the sample that are drawn from each of nine industrial sectors, reacts. The reaction is in real time, but a snapshot of it is shown in Figure 6*B*. It shows us that most of the companies with large assets relative to their market value are finance companies (banks).

The linking of the scatter plot with the bar chart provided the environment within which the explanatory power of slicing can be effectively utilized. Slicing from the bottom up would show us that companies of less than expected assets seem to be distributed more or less uniformly across all of the industrial sectors.

Identification   Often simply identifying cases on a display proves to be a powerful way to add information to the display. It aids understanding by going beyond displaying general patterns and relationships in the data. Usually interpreting such patterns or trends requires that we know which cases make up each of the groups, which cases form the heart of the trend, and which cases fail to follow the pattern established by the others. For this, we need to be able to identify data points on a plot. The most common method of identifying points interactively on a display is to click on the points in question and have identifying text appear—usually near the mouse cursor, but occasionally in a related table.
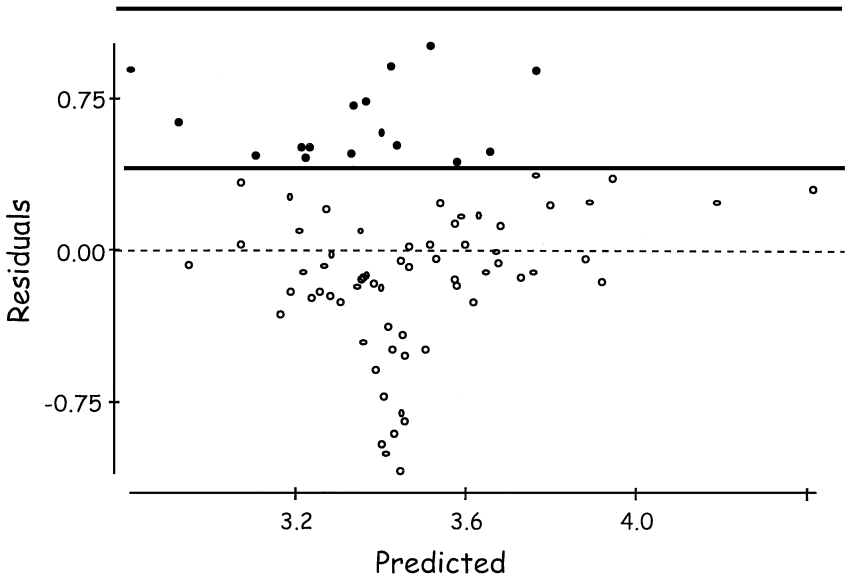
Panel A
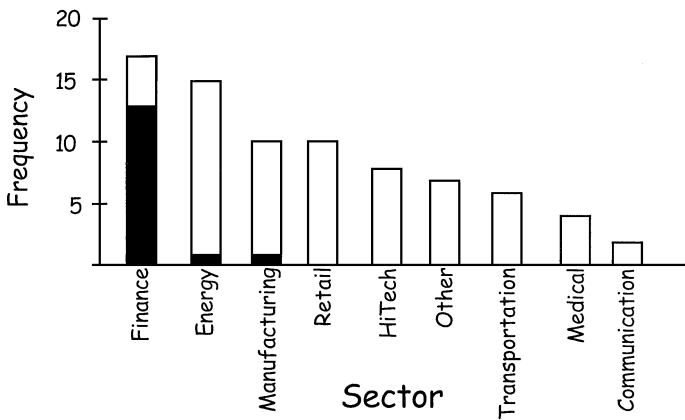
Panel B

Figure 6   (*Panel A*) The residual plot from Figure 9 is sliced downward from the top of the vertical axis. Those items selected by the slicer are shown darkened. (*Panel B*) A barchart showing the number of companies in each of nine industrial sectors. As a company is selected by the slicer in *Panel A*, the sector that it belongs to is shaded. This display shows that most of the companies with positive residuals are in the finance sector.

Subset Selection   Selection and linking can also work between graphics and quantitative statistical analyses, providing a powerful way to condition analyses. A quantitative analysis such as a regression, ANOVA, or contingency table can be constrained to be computed only for the selected cases. The data analyst can then select cases in an appropriate display and immediately see the quantitative analysis conditional on the selected cases. The lesson here is that graphics and quantitative analyses are part of the larger whole of data analysis and understanding, and are not two separate enterprises related only by their common database.

Subset selection is a first step from univariate and bivariate displays into analyses that depend on several variables. However, the most common subset conditioning selects levels in a categorical variable rather than ranges of a quantitative variable. Multivariate analyses of quantitative variables often turn to rotation for initial display.

Rotation   Rotating plots provide appropriate displays for many of the standard multivariate methods and can provide an intuitive way to learn about relationships among several variables without the need for advanced mathematics. The first program for rotating data was the PRIM-9 system developed by Fisherkeller, Friedman, and Tukey in 1972. It required several million dollars' worth of computer and display hardware so it remained a prototype system "proof-of-concept" implementation. PRIM stood for projection, rotation, isolation, and masking—the elementary operations that were found to be a basis for using plot rotations in data analysis. PRIM is a nice acronym, but the elements are more usefully discussed in "RIMP" order.

1. Rotation is an excellent, effective, general-purpose way to create the illusion of three dimensions. It provides both an immediate three-dimensional view of the point cloud and the ability to orient the point cloud in interesting ways. Early, special-purpose plot rotating programs restricted rotation to motion around one of the three standard axes, but modern software lets you rotate the points around any axis in the projection plane, often by "grabbing" the point cloud with a mouse and pushing it in any direction, much as you might rotate a globe mounted on gimbals by pushing lightly on its surface.

2. Isolation is the identification of subsets of points on the plot and the use of those subsets in further analyses, what we have called selection above. Often, rotated point clouds consist of several differently structured subgroups. Isolation makes it easy to focus attention on one at a time without viewing the others.

3. Masking is the ability to hide some part of the plot conditional on some other variable and concentrate on the remaining points. For example, one might want to know which part of the point cloud corresponds to points at one end of the range of some other variable or in some levels of a

categorical variable. Much of the masking principle is served in modern implementations by brushing and slicing.

4. Projection is the most subtle of the elementary operations. Rotating plots always show a projection of the point cloud on the screen. This projection establishes a relationship between the original variables (shown by the plotted data axis lines) and the plotting axes that point up-down, right-left, and in-out. Projection is especially powerful when the rotating plot accommodates more than three variables—a capability found in only a few of the current rotation implementations (see, for example, *Data Desk*, Velleman 1998). Often data originally recorded in several variables can be simplified to a few projected dimensions. A complete implementation of plot rotation should offer to record the linear combination of the original variables that results in the currently viewed projection, but this feature is often absent.

The Illusion of Three Dimensions

In the real world, we see three-dimensional objects in perspective. Objects farther away appear smaller; those closer appear larger. When we look at a real-world object we also have the benefit of stereo vision; each eye sees a slightly different view of the object, and our brain puts these views together to see the object at its true position in space. Rotating plots usually offer neither stereo views nor true perspective. Instead, the perception of depth comes from the animated rotation. In fact, the plotted points are just moving back and forth or up and down on the screen, but the viewer perceives this movement as a rotating three-dimensional cloud of points.

Because rotating plots show only a flat projection of the point cloud, true perspective plotting would be confusing. For data analysis, the initial view of the data (in which the y- and x-axes are in their ordinary orientation) is identical to a scatterplot. If a rotating plot showed true perspective, data points that were farther away (along the z-axis in-and-out of the screen) would shrink nearer to the center of the plot and data points that were closer to the viewer would spread away from the center of the plot, producing a distorted scatterplot.

To avoid this problem, rotating plots are usually drawn without any adjustment for perspective, much the way the world looks through a telescope or powerful telephoto lens. In this way, three-dimensional data analysis displays are different from representational three-dimensional drawings or computer-aided design displays. Such direct projection corresponds to the mathematical operation of projecting higher dimensional data into lower dimensional spaces that is fundamental to many multivariate analyses.

What Is an Interesting View?

Often we rotate a plot in search of interesting views of the data. Interesting views do not necessarily align with the data axes. If they did we could just plot simple

scatterplots and would have little need for plot rotation. Of course, the definition of "interesting" is deliberately vague. Sometimes an interesting view is a direction along which the data stretch out. Sometimes it is a view that shows distinct separated clusters. Often "interestingness" depends on the nature of your data or on your goals.

Fisherkeller et al (1974) discovered that many interesting orientations of the point cloud had the property that points seemed to clump together in separated clusters, which might then be isolated for further analysis. One might phrase this as groups with small within-group variance but large between-group variance, except that this phrasing suggests analysis of variance, which in turn suggests regularities such as homoskedasticity, which are definitely not restrictions on these patterns.

Instead, the pattern reflects the more general observation that one of the most important concerns of data exploration is with the homogeneity of the data. If our data do not describe a consistent, homogeneous population, it is difficult to imagine what it would mean to describe patterns with a statistical model or draw formal inferences from the data to the population. Thus, the discovery that a data set holds separate subgroups is often an important first step in understanding the data. We can then isolate each of the subgroups and analyze it separately, comparing the analyses along the way to understand how the subgroups resemble or differ from each other.

Projection is fundamental to many multivariate analyses. The combination of graphic techniques is often more effective than traditional multivariate computations at finding and clarifying multivariate structure in data. Principal components analysis and cluster analysis are among the methods that can be approximated by finding appropriate orientations of the point cloud and then using other statistical methods.

Seeing Patterns in Rotating Plots

Some statisticians have proposed that the best way to understand interesting patterns is to consider the least interesting pattern possible. For example, the normal distribution, useful though it may be in formal statistics, is fundamentally uninteresting in terms of real data. It is the deviations from normality that often prove interesting. A rotating plot of three random normal variables is basically uninteresting. It might then be argued that the more distant rotating plot data are from the multivariate normal, the more interesting they are in prospect. Unfortunately for this definition (and, of course, for all who would like to test their residuals to see whether they satisfy a multivariate normal assumption) there are many ways to deviate from normality. Fortunately, real data usually are interesting, although interesting patterns may be hidden from view at first. Several kinds of patterns are common and meaningful in data displays.

Orientations that show clusters of points separated from one another are often useful. Rather than showing most of the structure, a view of clusters often comes

about by looking at an interesting structure from the side. For example, if a point cloud consists of two separate stripes or "pencils" of points, rotating to look at the points of the pencils shows their separation but hides information about whether they are parallel or not. If you find an orientation with separated groups, consider assigning different plot symbols to the groups, or even hiding one temporarily and continuing the analysis with the other.

Uncovering* Randu's *Flaws: An Example of Discovery Through Rotation A well-known illustration of the value of a mobile display comes from IBM's ill-fated random number generator *Randu*. *Randu* is of a linear congruential type that yields numbers that depart from randomness in an interesting way. Suppose we generate 1200 numbers between 0 and 1 with this generator and consider each succeeding triple a point in three-dimensional space. We should end up with a uniform distribution of points on the unit cube. A two-dimensional projection of that cube is shown in Figure 7. Nothing in this display looks out of the ordinary. If we rotate the three-dimensional cube we find that most views support the conclusion that *Randu* has yielded a set of 400 points uniform in this space, yet suddenly we discover that (Figure 8) all the points line up on 15 planes in 3-space—a most decidedly nonrandom configuration. We note that this pattern of 15 stripes disappears quickly as we rotate away from this viewpoint by even a few degrees. This phenomenon is familiar to anyone who has ever driven past a cornfield and noticed how the corn rows sometimes line up and at other times look as if they are planted helter-skelter.

This flaw in *Randu* was described first by Marsaglia (1968), but is trivially uncovered with a rotation engine. The story might be more dramatically told if it was done dynamically, but the value of the outcome is fully appreciated with the static view of the end result.

In some displays, points cluster into isolated groups, but only in particular orientations of the display. It is often interesting to know whether the same cases cluster together in other displays of related variables. Assign a different plot symbol or color to each group, highlight clusters, or brush the plots to look for clustering across plots. Single variables with two clusters show up as two-humped, bimodal histograms. Slicing across one hump selects those cases so you can consider them in other plots.

Uncovering Differences Among Iris Species: An Example of the Power of Adding Identification As an example consider the 150 data points in Figure 9 [measurements made by the botanist Edgar Anderson, but first published by RA Fisher (1936)]. There were four measurements (in centimeters) made on each of three varieties of iris: sepal length, sepal width, petal length, and petal width. Originally there were only two varieties, *Iris Setosa* and *I. versicolor*, but Fisher added data [also gathered by Anderson (1928)] on *I. Virginica* to test Randolph's (1934) hypothesis that *I. versicolor* is a polyploid hybrid of the other two species.
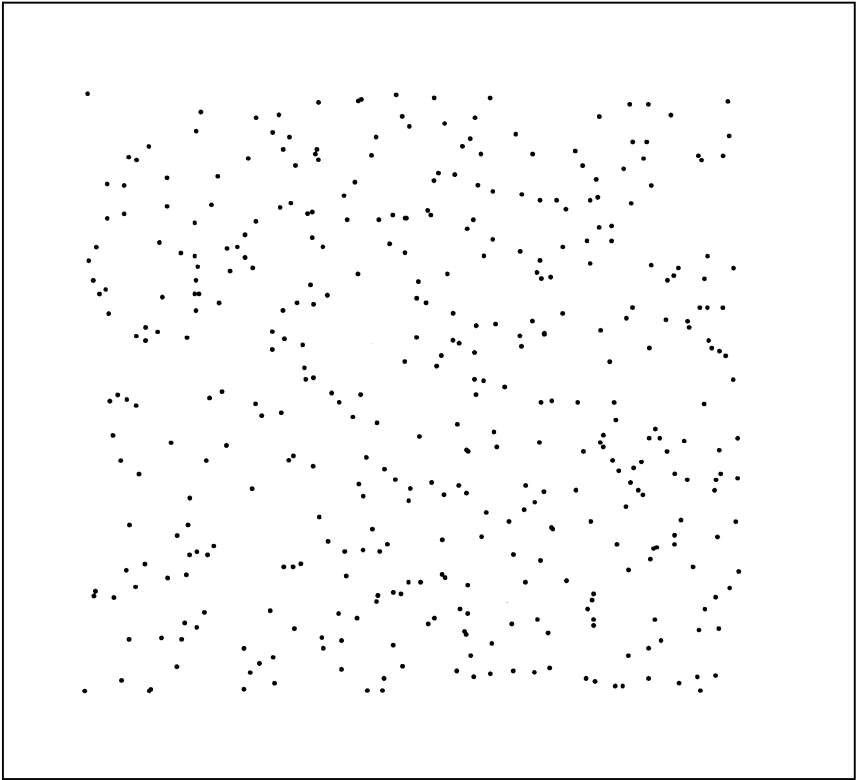
Unrotated Randu Data

Figure 7   A two-dimensional projection of 400 points plotted in the unit cube generated by the congruential random number generator, *Randu*.

We combine the four variables into two, sepal area (sepal width x sepal length) and petal area (petal width x petal length), and plot them (Figure 9). There seem to be two obvious groupings, but what are they? By assigning a different plotting symbol to each species we see that there are three, almost nonoverlapping distributions (Figure 10). This not only demonstrates the power of identification, but provides evidence about the relative power of graphical and analytic methods for scientific discovery. The graphic provides the primary evidence, and the analytic method (in this case discriminant analysis) is merely backup.

As we noted earlier, recognizing subgroups in data is an important exploratory step. When you find that your data can be split into subgroups, you may first want to find ways to characterize them. Often the best way to characterize clusters is to identify some of the cases in each cluster. For example, you may find that males and females form separate groups in your data (even though gender was not one of the
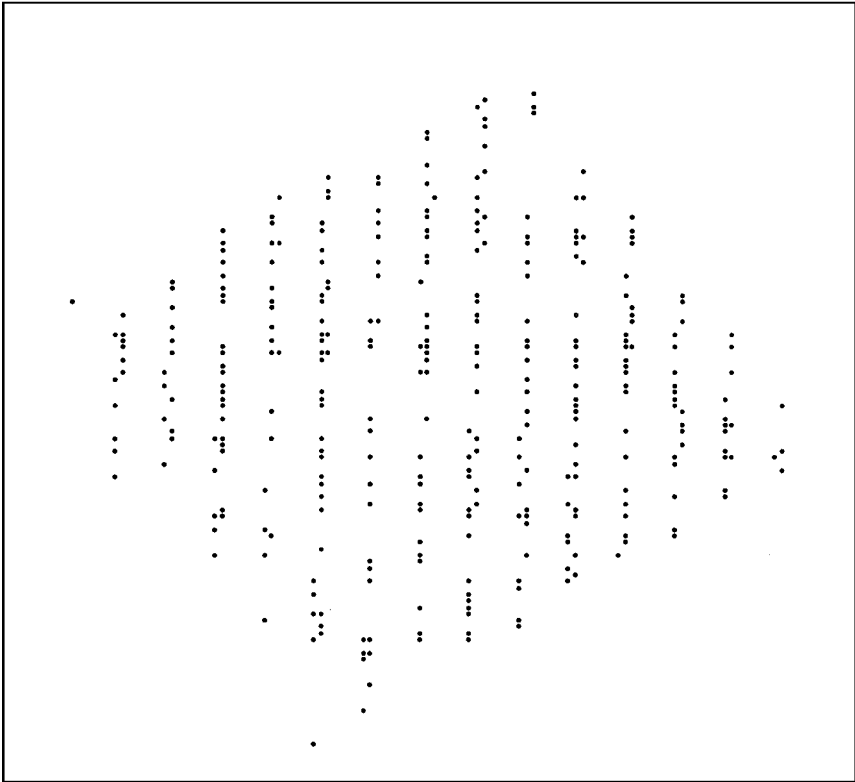
Rotated Randu data

Figure 8 A different two-dimensional projection of the cube shown in Figure 7 showing the striped pattern that is evidence for the conclusion that *Randu* does not yield entirely random numbers.

variables displayed), or that region, age, or season define subgroups. If the characterization is one that you did not anticipate, you have discovered a lurking variable.

Whether you can characterize the subgroups or not, it is often worthwhile to pursue analyses of the subgroups separately. Although it is rarely stated explicitly, a fundamental assumption of virtually every statistical analysis—even when no inference is planned—is that the data come from a single homogeneous population. If, in fact, the data come from two or more different subpopulations, it is usually more effective and more appropriate to analyze the groups separately.

Another interesting orientation is one in which the points are as spread out as possible along a particular axis (although in this case we must choose our scaling carefully; a common default scaling method, dividing each variable by its standard deviation, corresponds to a standard practice in multivariate statistics).
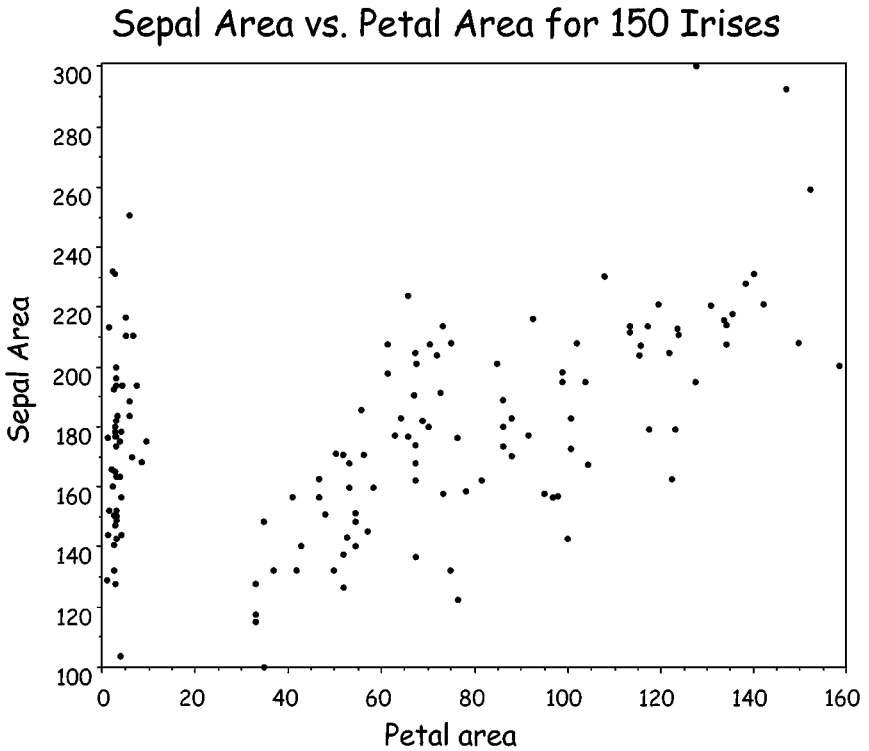
Sepal Area vs. Petal Area for 150 Irises

Figure 9   A scatter plot of sepal areas versus petal areas for 150 iris plants. These measurements were drawn from 50 of each of 3 varieties of iris; *Iris Setosa*, *I. versicolor*, and *I. Virginica*.

The direction of greatest variance, if we scale by standard deviation, is the first principal component of the data. One advantage of rotating plots is that it is relatively easy to ignore an outlier when positioning the display, even when the outlier might otherwise affect a multivariate calculation.

Many multivariate plots actually have only two or three directions of substantial variance, but these may not be aligned with the original variables. By identifying these principal axes we can simplify the analysis, reducing the number of dimensions to consider. An axis of great variance can also be a good axis to relate to other variables. For example, brushing along the axis while watching other plots can tell you much about its relationship to other variables.

Whether you can characterize the subgroups or not, it is often worthwhile to pursue analyses of the subgroups separately.

Some Rotating Plot Orientations Show a Clear Trend   Trends that are straight lines can be described with regression analyses or assessed with a correlation
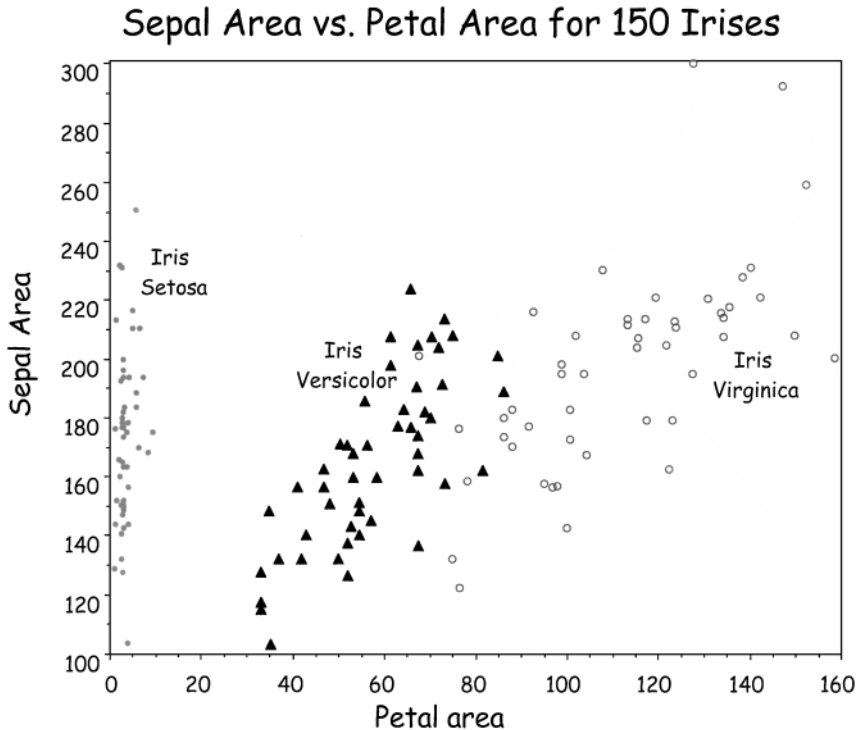
Figure 10   The scatter plot shown in Figure 9 with members of the three varieties of iris identified. The key aspect of this plot, which makes it different from similar plots done in earlier times, is how easily the identification was accomplished.

coefficient (although, of course, the relationship is probably more concisely described in terms of the projected variables). Trends that are not straight can be assessed with a nonlinear regression analysis or a nonparametric correlation coefficient such as the Spearman or Kendall correlations. Alternatively, they might become both clearer and more useful by transforming one or more of the variables.

Some Rotating Plots Show a Flat Surface   Flat surfaces can be described statistically with a multiple regression analysis. They tell us there is a combination of the variables that varies little, suggesting that we do not really need three dimensions to describe the data.

One of the Most Common and Useful Patterns Is the Simple Extraordinary Point or Outlier   Points can be extraordinary by being very far from the rest of the data or by failing to conform to a pattern, even though they are not particularly far from the data. An extraordinary point may be a sign of errors in the data such as a

misplaced decimal point or swapped digits. It may be a point that should not be a part of the data (for example, a motorcycle or truck included with cars).

An extraordinary point may be a perfectly correct and valid point that simply does not fit. These are often the most interesting points because we can learn a great deal by discovering why the point does not fit with the rest. Sometimes it may be better to remove or suppress an extraordinary point during part of the analysis so it does not dominate the calculations.

Occasionally a Rotating Plot Reveals a Complex Pattern   Examples of such patterns are planes that twist into a helical shape, parallel or intersecting lines or planes, and patterns with multiple extraordinary points. These represent patterns beyond the reach of any static statistics computation. The only really good way to describe such patterns is with several pictures or with a rotating plot. Unfortunately, rotating plots pasted into text documents and printed no longer rotate, so you may find that you must spend the traditional 1000 words to describe the picture.

Rotation and Color as an Additional Dimension

Used wisely, color can be a valuable addition to a rotating plot. You can use color to identify different groups or to represent values on another variable. When color represents a continuous variable it provides another dimension of information in addition to the three dimensions seen in the rotation. This can be an effective way to see four variables together, especially if the colored dimension is well ordered relative to the spatial dimensions. Sadly, combining the words color and well-ordered in the same sentence is typically an oxymoron, at least as it concerns human perception (Bertin 1973). The only aspect of color that is well ordered is saturation, and hence if we wish to represent an ordered variable with color we ought to do it by varying the saturation. Of course, using color for purposes of identification (e.g. "note the red points") can work very well indeed.

Four Variables and More

Some rotating plot implementations can handle more than 3 spatial dimensions (PRIM-9 could work with nine, *Data Desk* can work with twelve or more). Although most people find it hard to visualize four or more dimensions, virtually every multivariate statistical method searches for patterns or structure in a multidimensional array of points. Multidimensional rotating plots let you see the patterns and relationships that multivariate methods describe with numbers. Along with linking, symbols, color, and brushing, multidimensional plots make concrete what could only be imagined before. When a rotating plot has more than three dimensions, the viewer must select three dimensions to rotate. All other dimensions are held perpendicular to the chosen dimensions and do not rotate.

Practical Multivariate Graphics

Multivariate analyses are a constant struggle to reduce high-dimensional patterns to fewer dimensions to facilitate our understanding. Interactive displays can play a valuable role in this quest. For that to happen, displays must be integrated with analyses so that the data analyst can move smoothly from looking at aspects of the data to quantitative descriptions and tests and then back again to examine residuals or look for additional patterns. For multivariate analyses the investment in learning to use interactive graphics pays great dividends.

CONCLUSIONS AND LIMITATIONS

In the early part of the last century the poet Edna St. Vincent Millay wrote,

> Upon this gifted age in its dark hour
> Falls from the sky a meteoric shower
> Of facts. They lie, unquestioned, uncombined.
> Wisdom enough to leach us of our ills is daily spun,
> But there exists no loom to weave it into fabric.

This chapter is our attempt to chronicle the progress that has been made toward the construction of a glorious loom.

Space limitations have precluded more detailed discussions, and the obvious practical limitations of a print medium have forced us too often to tell rather than show. We hope we have been able to convey a sense of the exciting developments that widely available, powerful computers have made possible. Simultaneously, we would like to emphasize that the same perceptual system that led to the design of efficacious static displays remains with us for dynamic displays. Multicolored pseudo-three-dimensional pie charts that communicated data structures poorly when they were static, are not likely to improve if they spin through space in real time. The popularity of flashy (and often expensive) data-mining software demonstrates how easy it is to be seduced by the sizzle. In the assessment of new display technology we must ask first what can we learn using it that we would have missed without it. Or, more weakly, how much easier is it to have learned it this way?

Psychology, because of its long history and expertise in the measurement of perceptual phenomena, ought to take a lead role in such an assessment. We would like to encourage psychologists' involvement.

Penultimately, although there is an enormous amount of graphical software commercially available, very little of it thoughtfully melds the analytic side of the data analysis with the visual. *Data Desk* (Velleman 1998) is one such realization.[10]

---

[10]The reader must forgive the apparent self-serving nature of this recommendation. It is made defensible by two facts–the recommendation was written by the first author and it is true.

Details on programming graphics, and more importantly, how to think about programming graphics are now available in Wilkinson (2000).

Readers interested in pursuing the last two decades of developments toward a graphical loom should consult the marvelous work of Edward Tufte (1983, 1990, 1997), Bill Cleveland (1994a,b; Cleveland & McGill 1984, 1988), and of course, John Tukey (Tukey 1990, Basford & Tukey 1999).

ACKNOWLEDGMENTS

Visit the Annual Reviews home page at www.AnnualReviews.org

LITERATURE CITED

Anderson E. 1928. The problem of species in the northern blue flags, *Iris versicolor L.* and *Iris virginica L. Ann. Bot. Gard.* 15:241–332

Apel W. 1944. *The Notation of Polyphonic Music.* Cambridge, MA: Mediaeval Acad. Am.

Basford KE, Tukey JW. 1999. *Graphical Analysis of Multiresponse Data.* New York: Chapman & Hall

Becker RA, Cleveland WS. 1984. *Brushing a Scatterplot Matrix: High-Interaction Graphical Methods for Analyzing Multidimensional Data.* AT&T Bell Lab. Tech. Memo.

Beniger JR, Robyn DL. 1978. Quantitative graphics in statistics: a brief history. *Am. Stat.* 32:1–10

Bertin J. 1973. *Semiologie Graphique.* The Hague: Mouton-Gautier. (In French) 2nd ed. (W Berg, H Wainer. 1983. *Semiology of Graphics.* Madison: Univ. Wisc. Press)

Biderman AD. 1978. *Intellectual Impediments to the Development and Diffusion of Statistical Graphics, 1637–1980.* Presented at 1st Gen. Conf. Soc. Graph., Leesburg, VA

Biderman AD. 1990. The Playfair enigma: toward understanding the development of schematic representation of statistics from origins to the present day. *Inf. Des. J.* 6(1):3–25

Bronowski J, 1978. *The Origins of Knowledge and Imagination.* Binghamton, NY: Vail-Ballou

Clagett M. 1968. *Nicole Oresme and the Medieval Geometry of Qualities and Motions.* Madison: Univ. Wis. Press

Cleveland WS. 1994a. *The Elements of Graphing Data.* Summit, NJ: Hobart

Cleveland WS. 1994b. *Visualizing Data.* Summit, NJ: Hobart

Cleveland WS, McGill ME, eds. 1988. *Dynamic Graphics for Statistics.* Belmont, CA: Wadsworth

Cleveland WS, McGill R. 1984. Graphical perception: theory, experimentation, and application to the development of graphical methods. *J. Am. Stat. Assoc.* 79:531–54

Fisher RA. 1936. The use of multiple measurements in taxonomic problems. *Ann. Eugen.* 7:179–88

Fisherkeller MA, Friedman JH, Tukey JW. 1974. PRIM-9: An interactive multidimensional data display and analysis system. A.E.C. Sci. Comp. Inf. Exchange Meet. (Movie available from the Am. Stat. Assoc. An allied technical report same title, available from the Stanford Linear Accelerator Cent., SLAC-PUBL-1408)

Funkhouser HG. 1937. Historical development of the graphic representation of statistical data. *Osiris* 3:269–404

Galilei Galileo 1622. The assayer. In *The Controversy on the Comets of 1618*, Galileo Galilei, Horatio Grassi, Mario Guiducci, and Johannes Kepler. Transl. S Drake, CD O'Malley, 1960. Philadelphia: Univ. Penn. Press

Graunt J. 1662. *Natural and Political Observations on the London Bills of Mortality*. London: Martyn

Gunther RT. 1968. *Early Science in Oxford*, Vol. XIII. *Dr. Plot and the Correspondence of the Philosophical Society of Oxford*. London: Dawsons of Pall Mall

Howard L. 1847. *Barometrigraphia: Twenty Years' Variation of the Barometer in the Climate of Britain, Exhibited in Autographic Curves, With the Attendant Winds and Weather, and Copious Notes Illustrative of the Subject*. London: Richard & John E. Taylor

Huygens C. 1895. *Oeuvres Completes, Tome Sixieme Correspondance*, pp. 515–18, 526–39. The Hague: Nijhoff

Marsaglia G. 1968. Random numbers fall mainly in the planes. *Proc. Natl. Acad. Sci. USA* 61:25–28

McKie D. 1972. Scientific societies to the end of the eighteenth century. In *Natural Philosophy Through the 18th Century and Allied Topics*, ed. A Ferguson, pp. 133–43. London: Taylor & Francis

Playfair W. 1786. *The Commercial and Political Atlas*. London: Corry

Plot R. 1685. A letter from Dr. Robert Plot of Oxford to Dr. Martin Lister of the Royal Society concerning the use which may be made of the following history of the weather made by him at Oxford through out the year 1684. *Philos. Trans.* 169:930–31

Priestley J. 1769. *A New Chart of History*. London. (Reprinted 1792. New Haven: Amos Doolittle)

Randolph LF. 1934. Chromosome numbers in native American and introduced species and cultivated varieties of Iris. *Bull. Am. Iris Soc.* 52:61–66

Smith DE. 1925. *History of Mathematics*, Vol. 2. Boston: Ginn & Co.

Spence I, Wainer H. 1997. William Playfair: A Daring Worthless Fellow. *Chance* 10(1):31–34

Spence I, Wainer H. 2000. William Playfair (1759–1823): an inventor and ardent advocate of statistical graphics. In *Statisticians of the Centuries*, ed. CC Heyde. Voorburg, The Netherlands: Int. Stat. Inst. In press

Stigler SM. 1980. Stigler's Law of Eponymy. *Trans. NY Acad. Sci.* 239:147–57

Tufte ER. 1983. *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press

Tufte ER. 1990. *Envisioning Information*. Cheshire, CT: Graphics Press

Tufte ER. 1997. *Visual Explanations*. Cheshire, CT: Graphics Press

Tukey JW. 1977. *Exploratory Data Analysis*. Reading, MA: Addison-Wesley

Tukey JW. 1990. Data based graphics: visual display in the decades to come. *Stat. Sci.* 5:327–29

Velleman PF. 1998. *Data Desk*. Ithaca, NY: Data Description, Inc.

Wainer H, Spence I. 1997. Who was Playfair? *Chance* 10(1):35–37

Wainer H, Thissen D. 1981. Graphical data analysis. *Annu. Rev. Psychol.* 32:191–241

Wainer H. 1996. Why Playfair? *Chance* 9(2):43–52

Wainer H. 1997. *Visual Revelations: Graphical Tales of Fate and Deception from Napoleon Bonaparte to Ross Perot*. New York: Copernicus/Hillsdale, NJ: Erlbaum. 2nd print. 2000

Wainer H. 1998. The graphical inventions of Dubourg and Ferguson: two precursors to William Playfair. *Chance* 11(4):39–41

Wigner EP. 1960. The unreasonable effectiveness of mathematics in the natural science. *Commun. Pure Appl. Math.* 13:1–14

Wilkinson L. 2000. *The Grammar of Graphics*. New York: Springer-Verlag