Graphical Methods for Data Presentation: Full Scale Breaks, Dot Charts, and Multibased Logging

WILLIAM S. CLEVELAND*

Experimentation with graphical methods for data presentation is important for improving graphical communication in science. Several methods—full scale breaks, dot charts, and multibased logging—are discussed. Full scale breaks are suggested as replacements for partial scale breaks, since partial breaks can fail to provide a forceful visual indication of a change in the scale. Dot charts show data that have labels and are replacements for bar charts; the new charts can be used in a wider variety of circumstances and allow more effective visual decoding of the quantitative information. Logarithms are powerful tools for data presentation; base 2 or base *e* is often more effective than the commonly used base 10.

KEY WORDS: Statistical graphics; Graphical perception.

1. INTRODUCTION

Graphs are important tools for presenting data in science. In the preceding article (a companion piece; Cleveland 1984), I have argued that graphical communication in science needs improvement. One research area that can contribute to such improvement is the assessment of current graphical methods for presenting data and the development of new methods.

Within the field of statistics, there has developed during the past 20 years a wide base of research on graphical methods for *data analysis* (Tukey 1977; Fienberg 1979; Chambers et al. 1983). But in statistics the experimentation with new graphical methods for *data presentation* has been less widespread. One example is Gonzalez et al. (1975), in which methods for portraying the sampling variation of survey statistics are given; this work is reflected in the final chapter of Schmid (1983). Another example is Tufte (1983), in which some new ideas about graph design are presented. Clearly there is much overlap of the area of graphical data analysis and the area of graphical data presentation, but there are many special considerations that arise when a graph is made to present data to others.

In this paper I illustrate the assessment and development of graphical methods for data presentation by describing three methods: full scale breaks, dot charts, and multibased logging. Some of the ideas are new; others are old but do not appear to be widely known despite their usefulness.

*William S. Cleveland is a statistician at AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, NJ 07974. Jon Kettenring, Colin Mallows, Marylyn McGill, Robert McGill, Ronald Snee, Walter Tapp, Edward Tufte, Paul Tukey, and anonymous referees made many helpful comments that very much affected this work.

2. FULL SCALE BREAKS

In science today the convention for indicating a change or gap in the scale of a graph along one axis is to break the axis either with two short wavy parallel curves or with two short parallel line segments. This is illustrated in Figure 1 with some made-up data. But the partial scale break is a weak indicator that the reader can fail to appreciate fully; visually the graph is still a single panel that invites the viewer to see, inappropriately, patterns between the two scales. Figure 1 gives the false impression that the rightmost two points continue the nearly linear behavior of the points on the left. The partial scale break also invites authors to connect points across the break, a poor practice indeed; this invitation was accepted in many of the graphs of the surveys reported in Cleveland (1984).

A better procedure is the full scale break illustrated in Figures 2 and 3 (and in Fig. 2 of the companion article). The full break results in a graph with two juxtaposed panels. This use of juxtaposition to provide a full scale break, *with each panel having a full frame and its own scales,* shows the scale break about as forcefully as possible and discourages mental visual connections by viewers and actual connections by authors. Full scale breaks are rare in scientific graphics but are not unprecedented; it has been argued that two juxtaposed panels should be used when a time variable has a break in its scale (Dept. of the Army 1966).

3. DOT CHARTS

Figures 3–8 are *dot charts;* so are Figures 1–4 of the companion article (Cleveland 1984). As will be discussed shortly, several aspects of dot charts vary according to the nature of the data and the purpose of the presentation.

The data in Figures 3 and 4 are the populations of European cities around 1800. The numbers appeared in a graph of William Playfair (see Tufte 1983). The data in Figures 5–8 are from one survey reported in the preceding article. For 50 articles in each of 57 scientific journals, the fraction of space devoted to graphs and the fraction devoted to illustrations were measured. (The graph areas are discussed in Cleveland 1984, but not the illustration areas.) An illustration is any figure that is not a graph; examples are schematic diagrams, photographs, and maps with no statistical information. Fractional figure area is the sum of fractional graph area and fractional illustration area. Figure 5 shows the fractional figure areas for the nine journals whose areas were above .25. All three fractional areas are shown in Figures 6 and 7 for these nine journals. The data in Figure 8 are the fractional graph areas for social science journals (these data are also shown in the bottom panel of Fig. 3 of Cleveland 1984).

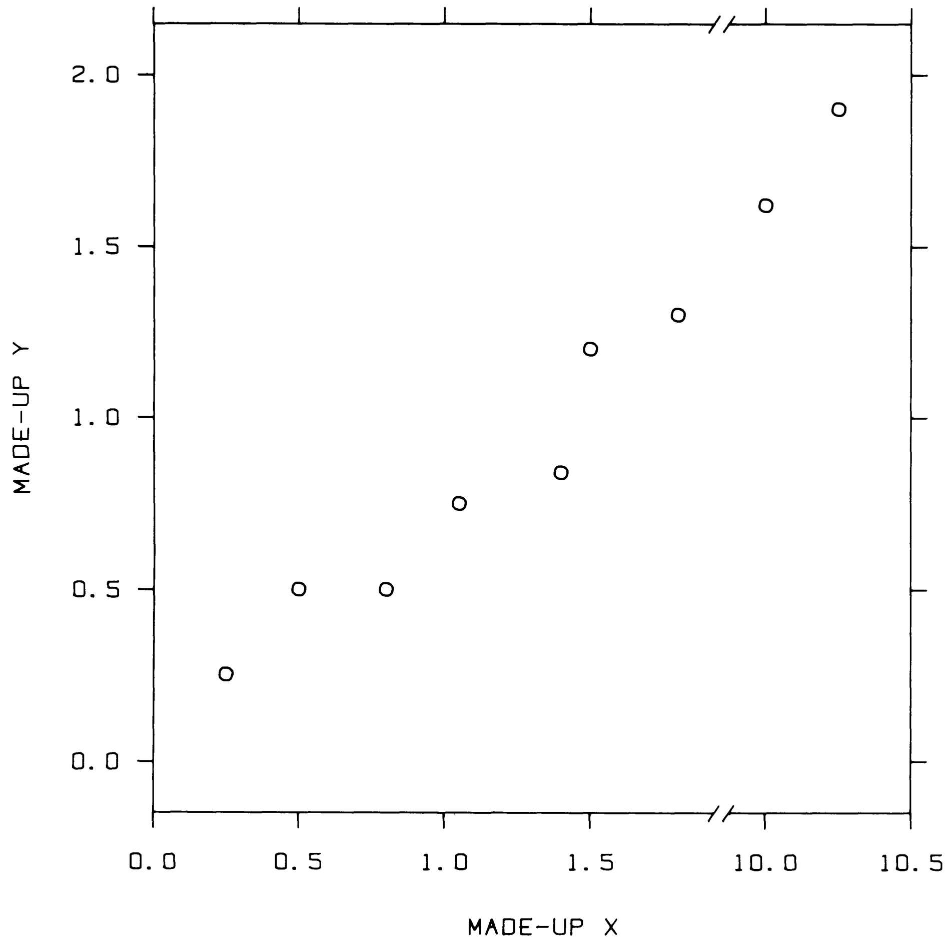

Figure 1. Partial Scale Break. The partial scale break (two parallel slanted lines breaking each axis) is a weak indicator that the viewer can fail to fully appreciate. The graph gives the false impression that the rightmost two points continue the nearly linear behavior of the points on the left.

Dot charts can be used when the values in a data set have names that one wants to convey. When there are many values, the light dotted lines on the graph enable the visual connection of a plotted point with its name. When the number of values is small, as in Figure 5, the dotted lines can be omitted, since the visual connection can be performed without them.

The dotted lines on the dot chart have been made light to keep them from being visually imposing and obscuring the large dots that portray the data. When we visually summarize the *distribution* of the data, the data dots stand out and the graph is, in effect, a sample cumulative distribution function, provided the data are ordered from smallest to largest as in Figures 3–5 and 8. When we want to emphasize this distribution as well as the labels, a scale for the sample distribution can be put on the right vertical axis of the graph, as in Figure 8. Following a standard convention (Chambers et al. 1983), the sample cumulative distribution function at the ith order statistic is taken to be $(i - .5)/n$, where n is the number of observations.

When there is a zero on the scale—or some other meaningful baseline value—and no scale break, the dotted lines can emanate from the baseline value and end at the data dots, as in Figure 6. When there is no meaningful baseline value on the scale or when there is a scale break, the dotted lines should go across the graph, as in Figures 3 and 4; the reason for this will be explained shortly.

Sometimes data values that have names fall into two or more groups; in many of these cases, we want to show a statistic, such as a total, for each group. Such data can be conveyed by a *dot chart with groups*, as illustrated in Figures 6 and 7 (and in Fig. 2 of the companion article). There are two ways to make the dot chart with groups. In one (illustrated in Fig. 6) each value in a group is on a separate line and the group structure is conveyed by indenting labels. In the second (illustrated in Fig. 7) the values for each group are on the same line and the items within each group are identified by a key. This second variation of the dot chart with groups is successful only when there are not too many items within each group and when the plotting symbols do not overlap.

On a dot chart with grouping, the choice of the symbols for the different items within each group is a vital issue. In Figures 6 and 7 fractional figure area is encoded by partially filled circles, fractional graph area is encoded by unfilled circles, and fractional illustration area is encoded by filled circles. The objective is to use an encoding scheme that provides high visual contrast so that we can focus on all of the values for one type of item, mentally filtering out the rest of the values. For example, in Figures 6 and 7 it is possible to focus on the fractional graph data as a whole and ignore other values. It turns out that changing the method

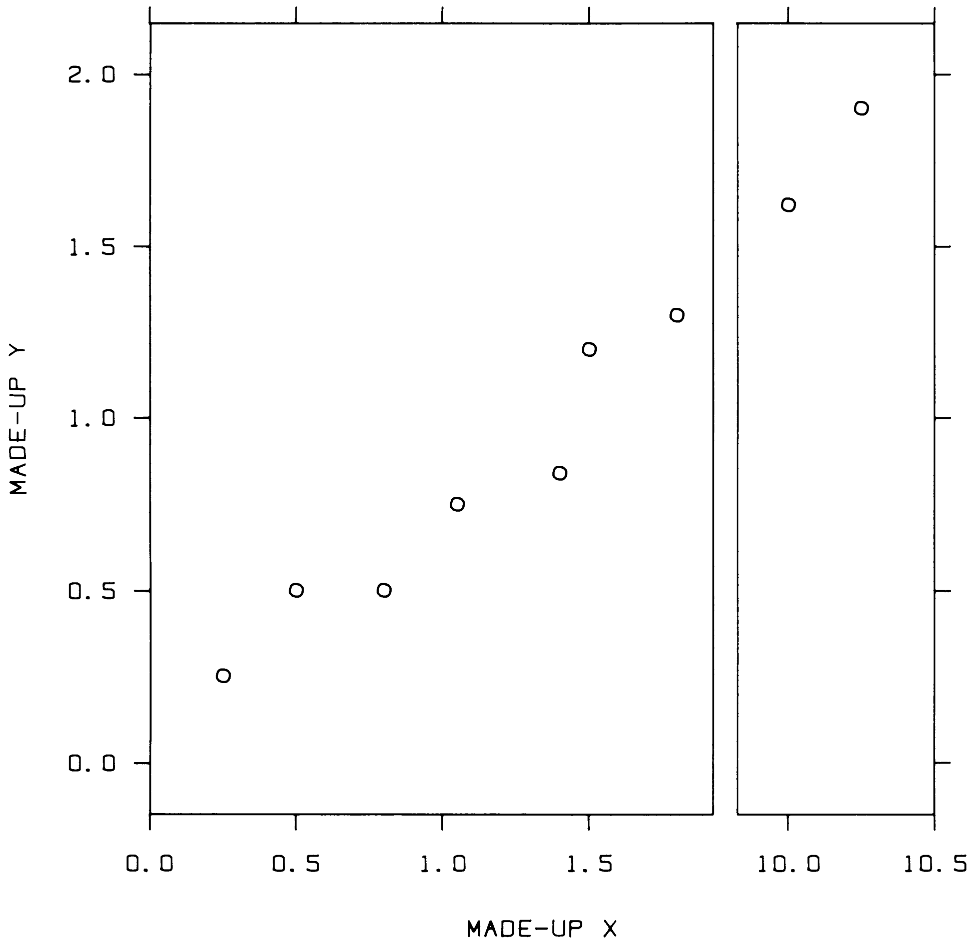

Figure 2. Full Scale Break. The full scale break shows a scale change much more forcefully than does the partial scale break and discourages mental visual connections by viewers and actual connections by authors.

of filling circles provides high visual discrimination; other schemes, such as different letters or different geometric shapes (e.g., circle, square, and triangle) do not do as well.

To understand why dot charts are a useful graphical method, it is important to appreciate that a bar chart should not be used in the way the dot chart is used in Figures 3–5. That is, the basic graphical element on these graphs—the filled circle together with the dotted line—should not be replaced by a bar emanating from the left and ending at the data value. The bar of a bar chart has two aspects that can be used to visually decode quantitative information—size (length and area) and the relative position of the end of the bar along the common scale. The changing sizes of the bars is an important and imposing visual factor; thus it is important that size encode something meaningful. The sizes of bars encode *the magnitudes of deviations from the baseline*. If the deviations have no important interpretation, the changing sizes are wasted energy and even have the potential to mislead (Schmid 1983).

In Figure 3 there is a scale break, and in Figures 4 and 5 the baselines have no significant meaning; thus bars should not be used in these figures. Suppose that in Figure 5 we used bars to encode the data and kept the scale the same. The baseline then would be slightly below .25, which has no significance in this application; thus bar size would have no significant meaning. We could, of course, change the horizontal scale so that the baseline is 0, and use a bar chart; then bar length and area would encode fractional figure area. But this would result in wasted space and a decrease in the resolution of the graph. In Figure 4 I can think of no baseline that seems compelling for this population data shown on a log scale. Thus a bar chart seems inappropriate for these data. Bars should not be used in Figure 3 because the areas of bars broken by a scale break have no valid meaning in any situation and run the risk of leading a viewer to attach visual meaning, either consciously or not, and be misled.

For these same reasons the version of the dot chart in which the dotted lines go only as far as the data should not be used in Figures 3–5. The varying lengths of such lines would not be encoding anything important. By bringing the dotted lines all of the way across the graph, the portion between the left vertical axis and the data dot is visually deemphasized.

The data of Figures 6 and 7 are displayed in Figure 9 by a divided bar chart. A grouped dot chart—or even a grouped bar chart designed like the grouped dot chart in Figure 6 but with the data dot and dotted line replaced by a bar—will usually display such grouped data more effectively. On a divided bar chart, some values can be compared by making either area–length judgments or judgments of position along a common scale, but others can be compared only by area–length judgments. In Figure 9 the values of the fractional figure areas can be compared with one another by using length–area or position judgments, since there is a common

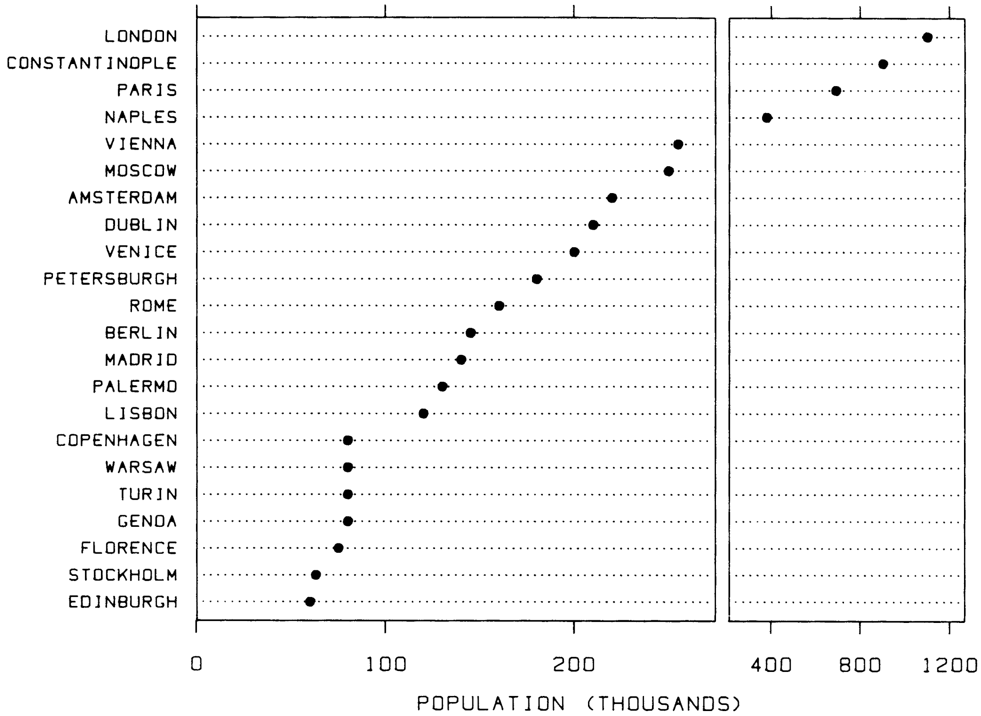

Figure 3. Dot Chart With Full Scale Break. The data are the populations of European cities around 1800 and are from a graph by William Playfair. The scale break is used, since without it the values below 250,000 are forced into a small region of the scale and lose much resolution. The data are displayed using a dot chart. A bar chart should not be used in the same way, since the scale break would make the bar sizes meaningless.

baseline for these values; the same is true of the illustration areas. But the fractional graph areas can be compared only by length–area judgments, since the lack of a common baseline does not allow judgments of position along a common scale; the same is true for the comparison of the fractional illustration area with the fractional graph area for a

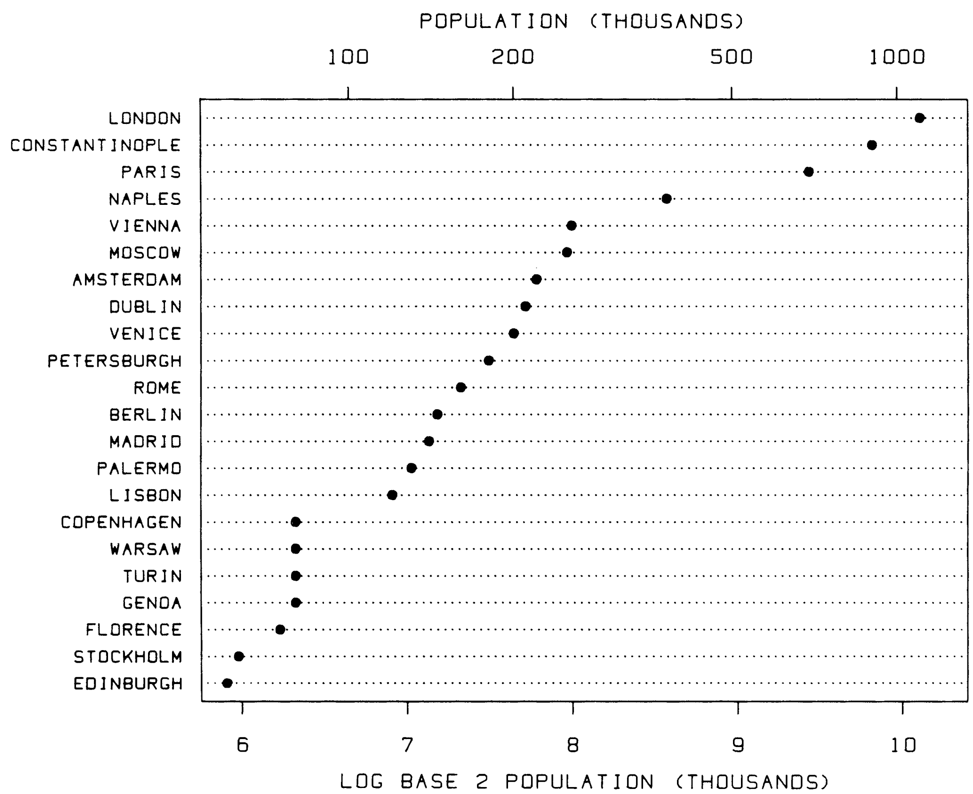

Figure 4. Dot Chart With a Log Base 2 Scale. The European city population data are shown using a dot chart. A bar chart should not be used in the same way, since the bar sizes would encode meaningless numbers. The log scale is convenient for these data, since it is natural to think of ratios of city populations and since the skewness is reduced by the transformation. Log base 2 is used rather than log base 10 to provide more easily interpretable scale markings.

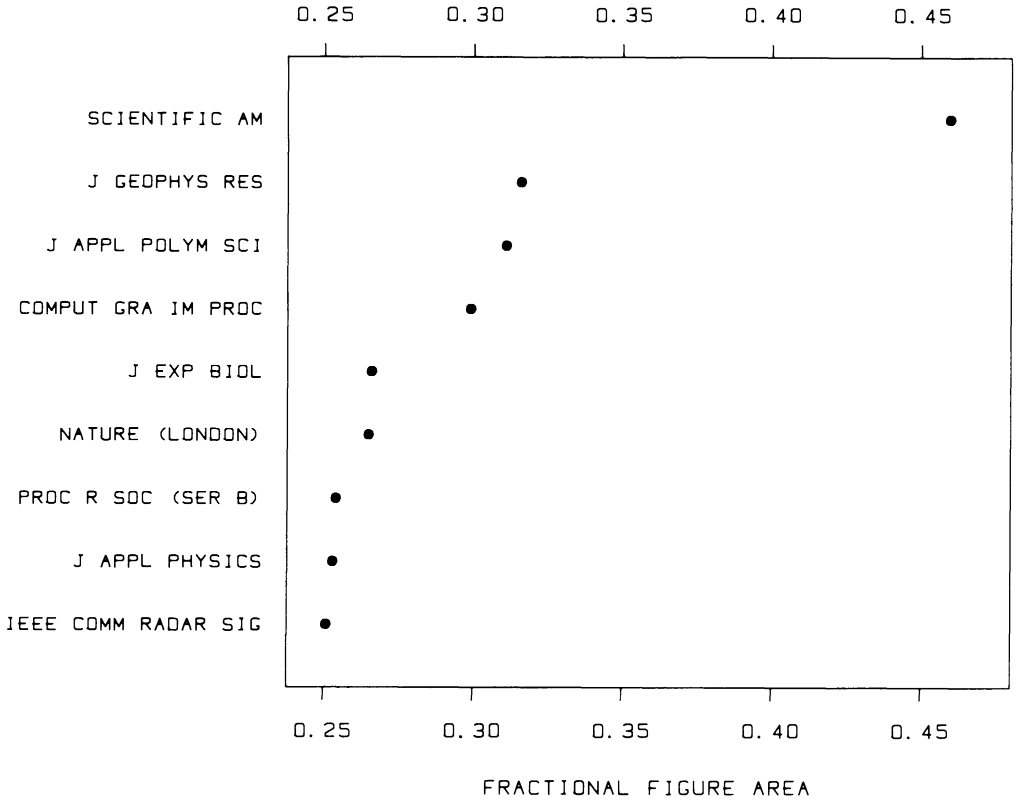

Figure 5. Dot Chart. The data are the fractions of space devoted to figures for nine scientific journals. A bar chart should not be used in the same way, since the bar sizes would encode meaningless numbers.

particular journal.

One of the implications of a theory of human graphical perception in Cleveland and McGill (1984) is that we perform judgments of position along a common scale much better than we do judgments of length and area. For example, in one experiment reported by Cleveland and McGill, errors of length–area judgments were 40–250% larger than those that involved judgments of position along a common scale.

If we use a dot chart with grouping in place of a divided bar chart, all values can be compared by making judgments of position along a common scale. Figures 6–8 illustrate the benefits. It is harder to compare the fractional graph area data on the divided bar chart than on the grouped dot charts. The reader is invited to determine which are the three smallest fractional graph area values and which are the three largest. For my eye–brain system the task is much easier on the grouped dot charts. Similarly, comparisons of the fractional illustration area data with the fractional graph area data are harder on the divided bar chart; for example, on the divided bar chart, it is somewhat harder to determine the journal that has the most nearly equal values of fractional illustration and graph areas; on the grouped dot charts, it is immediately obvious that it is the *Proceedings of the Royal Society*.

The final issue is whether dot charts—that is, ordinary bar charts and not divided bar charts—are more effective than bar charts even when there is a meaningful baseline, as in Figure 8. I will argue that they are. A reasonable principle for the design of graphs is to make the graphical elements representing the data as nearly equal in area as possible; this gives equal visual emphasis to all data values. On bar charts the areas of the elements representing the data—the bars—can be very unequal. Consider the bar chart of Figure 10; the data are the fractional graph areas for the social science journals displayed by the dot chart of Figure 8. The bar areas change radically, and the value for the *Journal of Social Psychology* disappears from sight. This does not happen on the dot chart of these data in Figure 8. Indeed, light dotted lines were chosen for giving the visual connection between the labels and the data dots to make the areas and visual impacts of the graphical elements as nearly equal as possible. We could, of course, go one step further and insist that the dotted lines always go across the graph, as in Figures 3 and 4, so that the areas of the graphical elements would be identical; but this seems unnecessarily restrictive, since the visual impacts of the different graphical elements in the Figure 8 version of the dot chart are so very nearly equal.

In addition, error bars are more effectively portrayed on dot charts than on bar charts. This is illustrated in Figures 11 and 12, which show fractional graph areas and approximate 95% confidence intervals for 25 journals; the data are the 25 largest fractional graph areas from the survey reported in Cleveland (1984). On the bar chart the upper values of the intervals stand out well, but the lower values are visually deemphasized and are not as well perceived as a result of being embedded in the bars. This deemphasis does not occur on the dot chart.

Dot charts—with all of the variations and the construction details—are essentially a new graphical method. In the version in which the dotted line goes from a meaningful base-

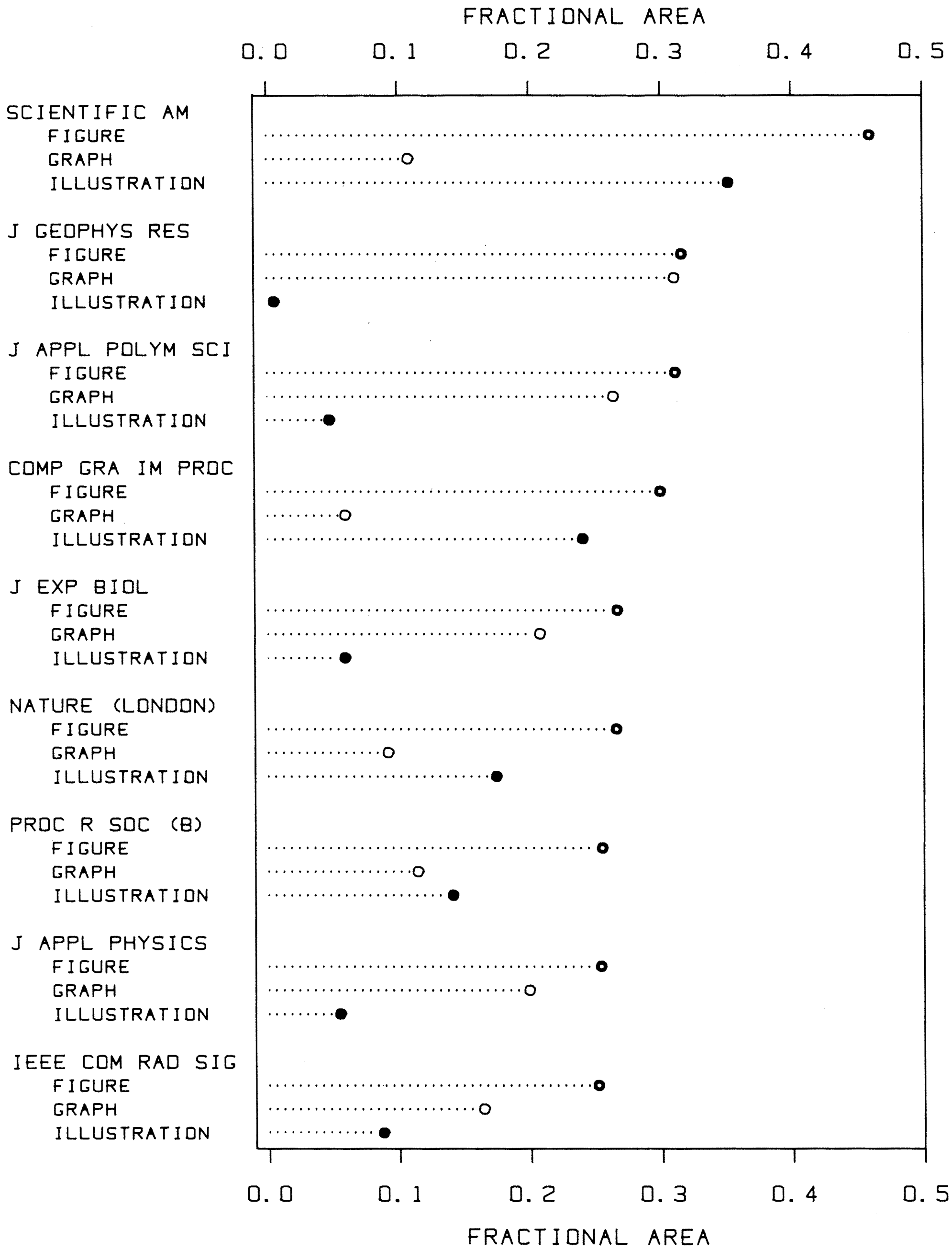

Figure 6. Dot Chart With Multiline Groups. The data are the fractions of space devoted to illustrations, to graphs, and to all figures (illustrations plus graphs) for nine scientific journals. The dot chart with grouping allows values to be compared by making graphical perceptual judgments of position along a common scale. For this reason it portrays data more effectively than does a divided bar chart. This is one variation of the dot chart with grouping. The second is shown in Figure 7.

line to the data dot, however, the graphical element is similar to what some call *thermometers* (Lydon 1961), a solid line with a data dot at the end.

4. MULTIBASED LOGGING

The logarithm is an extremely powerful and useful tool for graphical data presentation. One reason is that logarithms turn ratios into differences, and for many sets of data, it is natural to think in terms of ratios. Suppose that t, u, v, and w are four positive numbers with $t/u = v/w$; suppose that t is much or moderately bigger than v. Then on a graph involving these four numbers on the original scale, it is difficult to visually detect the equality of the ratios without reading numerical values from the scale and doing mental arithmetic. If we plot logarithms we have a much better chance of detecting equality, since it is now a matter of perceiving that the distance along the log scale of log (t) to log (u) is equal to the distance from log (v) to log (w); this is a reasonable graphical perceptual chore.

Consider the city populations in Figures 3 and 4. For such data it is natural to think about how many times bigger one city is than another. On the log plot in Figure 4, it is easy to see that the percentage by which Naples is bigger than Vienna is about the same as the percentage by which Lisbon is bigger than Copenhagen and that London is 16 times as big as Edinburgh and 4 times as big as Moscow.

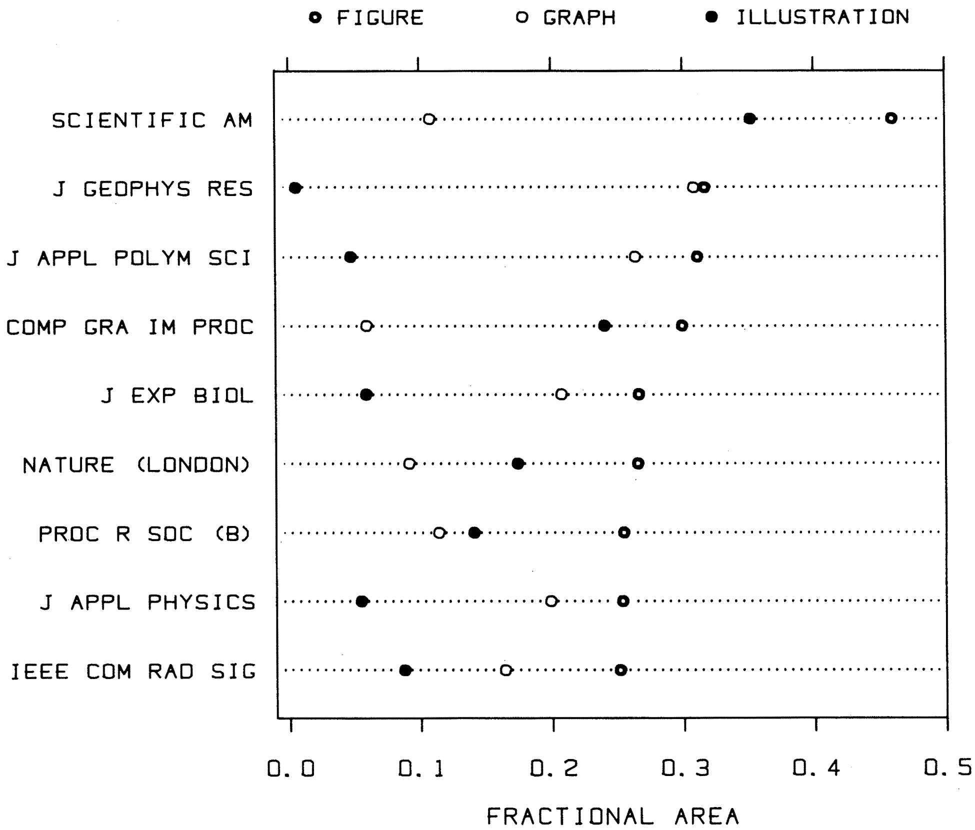

Figure 7. Dot Chart With Single-Line Groups. The data are the same as in Figure 6. The graph is the second variation of the dot chart with groups. This variation works only if the number of items in each group is not large and if the plotting symbols for each group do not overlap.

Another reason for the power of logarithms is resolution. Data that are amounts or counts are often very skewed to the right; on graphs of such data, there are a few large values that take up most of the scale and the majority of the points are squashed into a small region of the scale with no resolution. This is the case for the population data; plot-

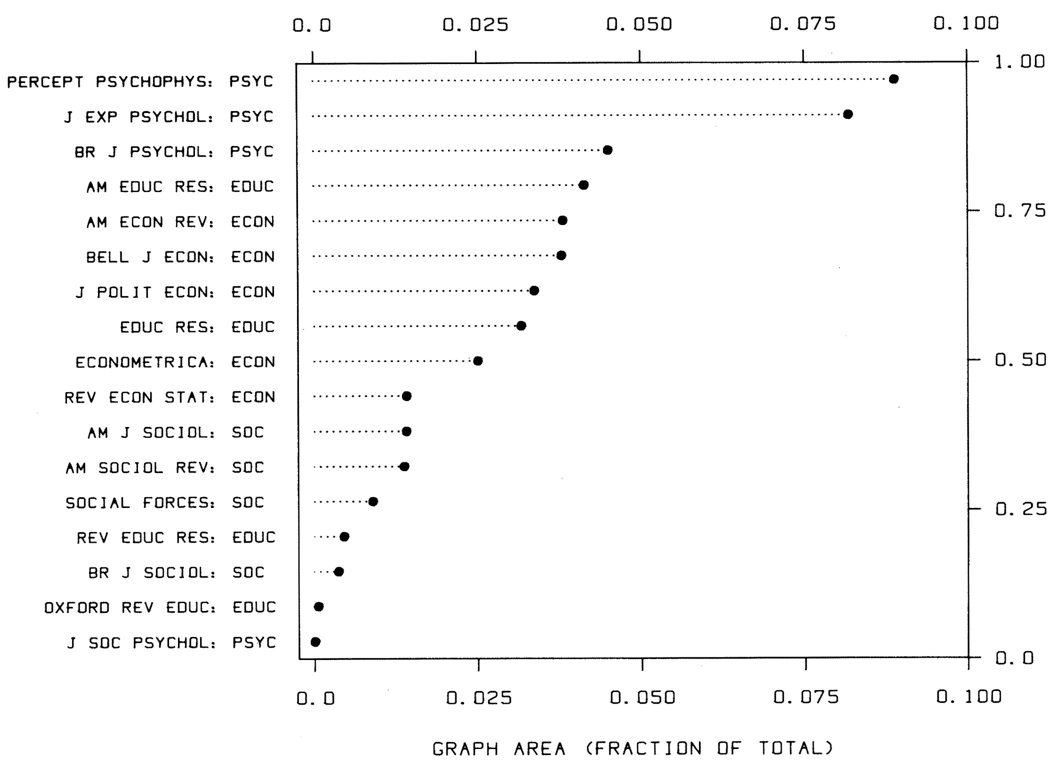

Figure 8. Dot Chart. The data are the fraction of space devoted to graphs for 17 social science journals. In this case the dotted lines can stop at the data dots because line length encodes something meaningful. The scale on the right vertical axis portrays the cumulative distribution of the data.

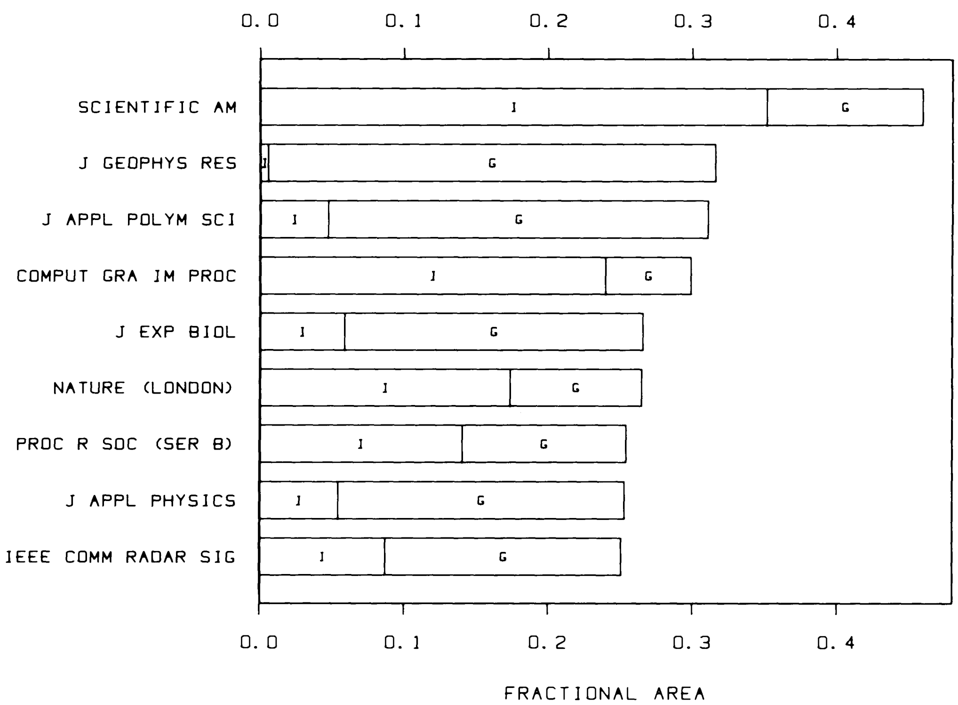

Figure 9. Divided Bar Chart. The journal data of Figure 6 are displayed using a divided bar chart. This display is not as effective as the dot chart with grouping, since a viewer must make difficult length–area judgments to compare some of the values.

ting the data on the original scale in Figure 3 forces us to use a scale break to get reasonable resolution for the values less than 250,000, but no scale break is needed on the log plot in Figure 4.

In some cases in which log plots appear in scientific publications, the tick marks and their labels show values on the original scale and are not equally spaced on the log scale; an example is the top scale markings of Figure 4. It is reasonable to do this *together with* equally spaced tick marks on the log scale, as in Figure 4; but omitting equally spaced log scale markings subverts the power of the logarithm to convey equal ratios.

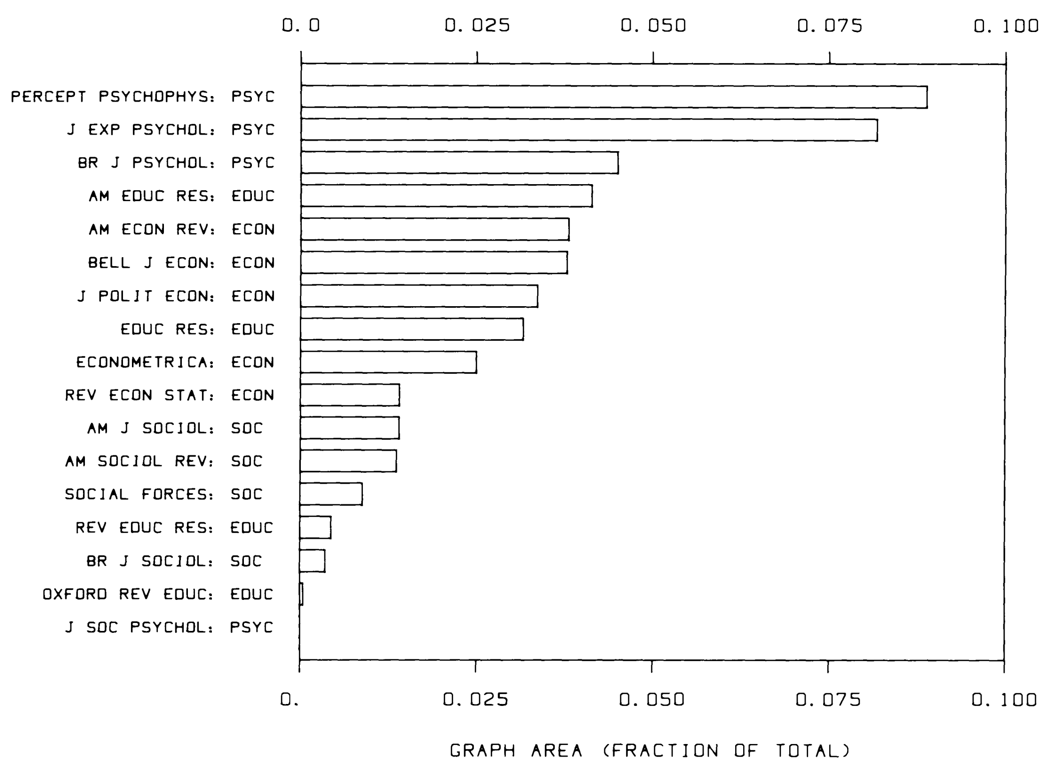

Figure 10. Bar Chart. The journal data of Figure 8 are displayed. Bar charts have the undesirable property that the graphical elements representing the data—the bars—change substantially in area. This is undesirable, since we do not make area judgments as well as other graphical perceptual tasks and since some data values are given more visual emphasis than are other data values. On the bar chart in this figure, one bar disappears altogether because the datum it represents is zero.

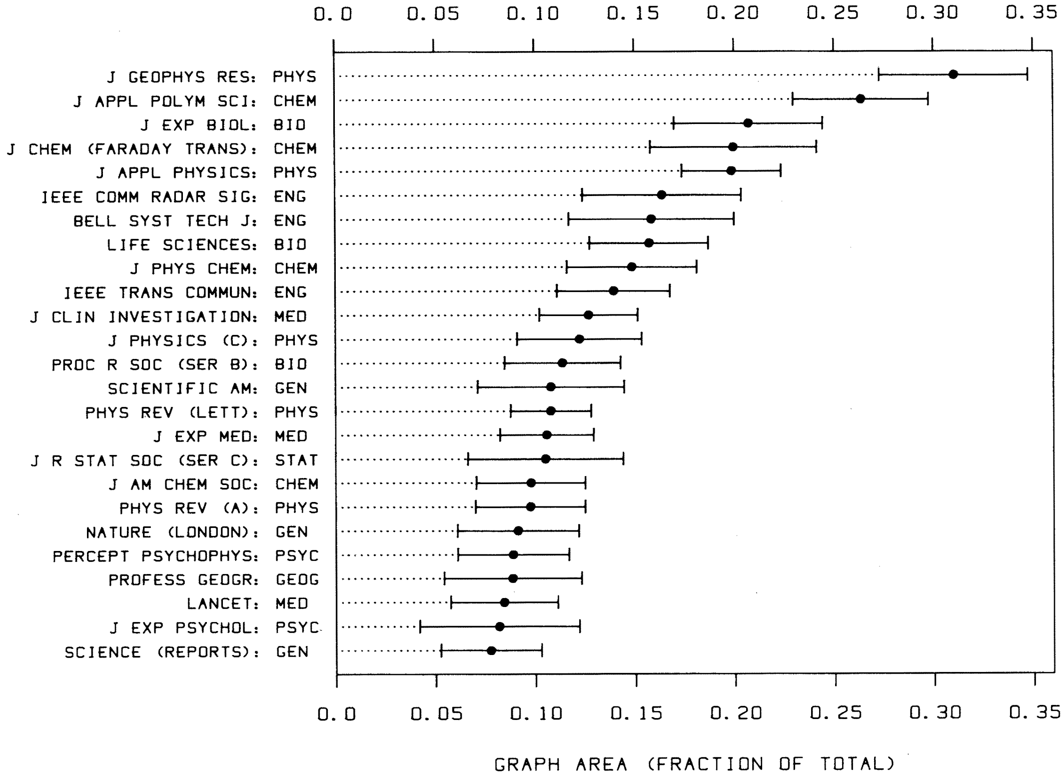

Figure 11. Dot Chart With Error Bars. The fraction of space devoted to graphs is shown for 25 journals; the data are the 25 largest fractional graph areas from the survey reported in the companion article (Cleveland 1984). The error bars show approximate 95% confidence intervals.

Log base 10 is almost always used in scientific presentation graphs for a log scale. This is much too limiting. Log base 2 and log base *e* should always be considered. The choice of the base depends on the range of the data values that need to be visually compared. Suppose the data go through many powers of 10. For example, brain weights of

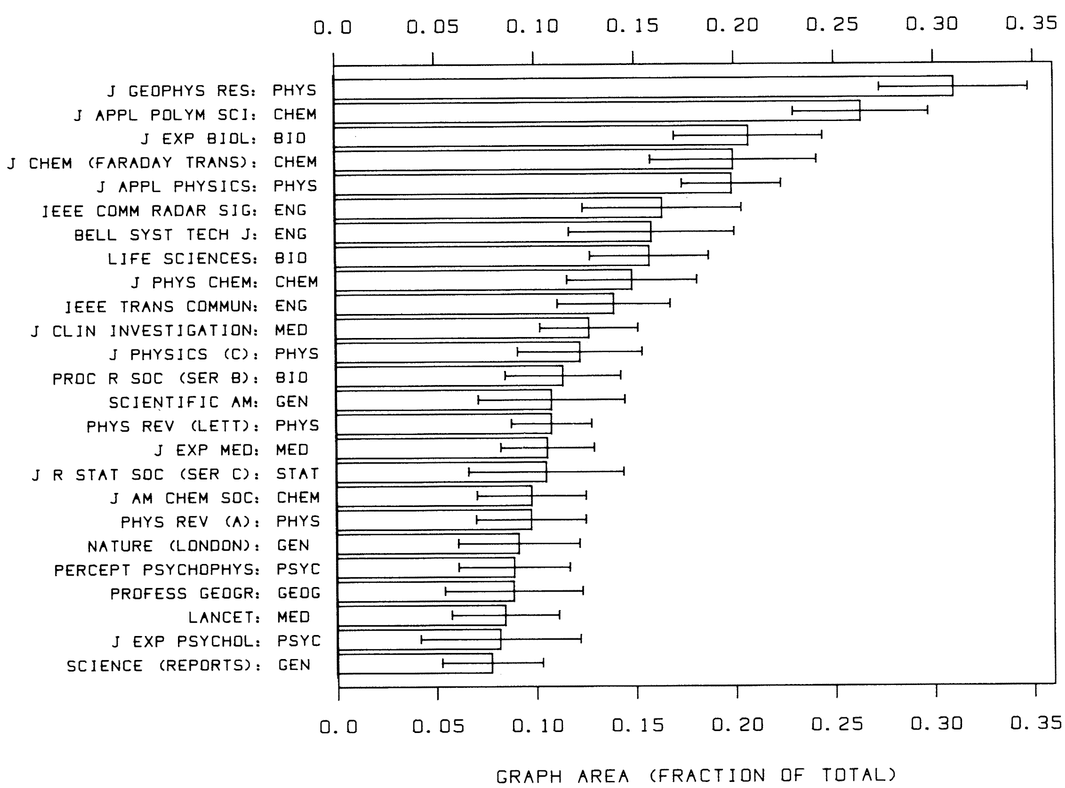

Figure 12. Bar Chart With Error Bars. The values portrayed are the same as those in Figure 10. The confidence intervals are not as clearly perceived on the bar chart as on the dot chart, since the lower sections of the intervals are visually camouflaged by the bars.

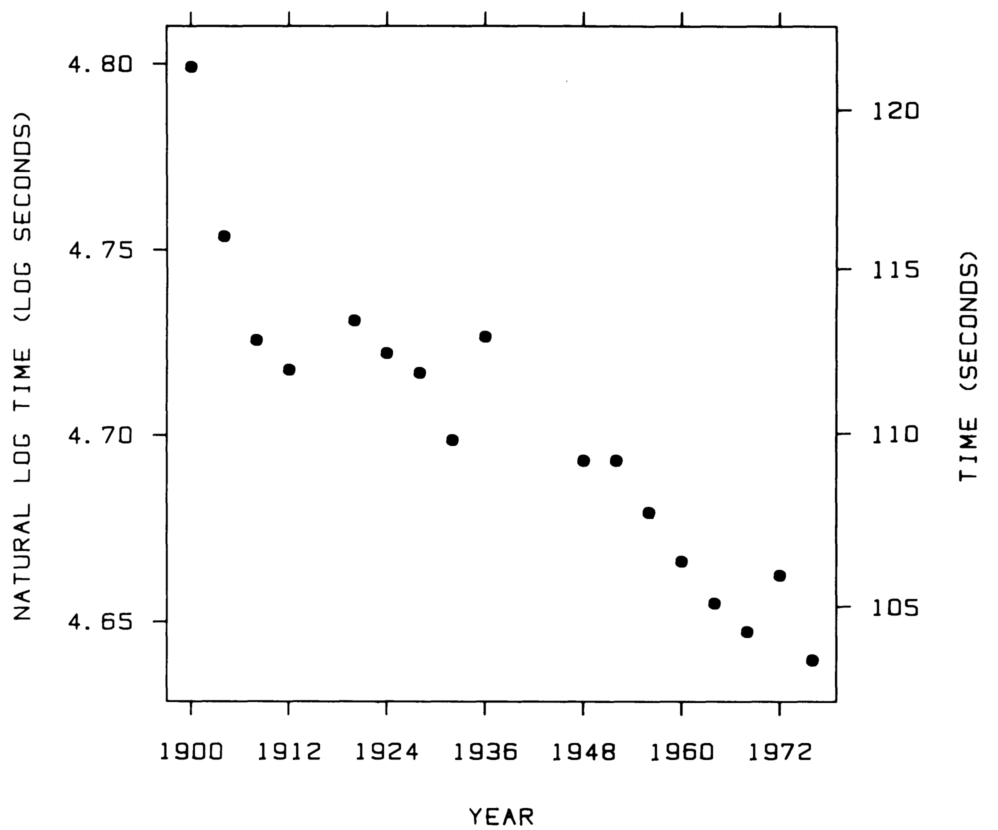

Figure 13. Olympic Data With a Natural Log Scale. The data are the winning times for the 800-meter dash at the Olympics from 1900 to the present. Since the data are plotted on a natural log scale and change by no more than .15 on that scale, 100 times the changes in the log values can be interpreted as percent changes of the winning times on the original scale. For example, since the values have decreased by about .1 on the log scale from 1904 to the present, the decrease in the winning times has been about 10% over this time period.

adult mammals range from less than .15 grams to more than 5,000 grams. In such a case it is reasonable to use log base 10.

But suppose that the data range over two powers of 10 or less, which is the case for the city population data in Figure 4. Then it is inevitable that equally spaced scale markings for log base 10 would involve fractional powers of 10. How many of us understand such fractional powers? Most know $10^{.5}$ is a little bigger than 3, but our understanding falls off rapidly after that. In such a situation it makes sense to convert to log base 2, as in Figure 4 (and in Fig. 4 of Cleveland 1984). Working with multiples of 2 in our heads when the data range over two or even three powers of 10 is easy. Recalling exact powers of 2 up to 2^{10} is not hard. (Those who used the fast Fourier transform in its early days have exact powers of 2 branded in their brains up to about 2^{12}.) We can even go higher by using the computer scientists' trick: let $k = 2^{10} = 1{,}024$, and now think of k as 1,000 so that $2^{14} = 2^4 \times 2^{10} = 16k \approx 16{,}000$.

Let us go further. Suppose that the ratio of the maximum to the minimum data values is too small even for log base 2 to be useful. We are unlikely to have much understanding of fractional powers of 2, so considering logarithm base e is reasonable because natural logarithms have a wonderful property. Suppose t and u are positive numbers and $t/v = 1 + r$, where r is between $-.25$ and .25. Then to a good approximation, log (t/v) is r. For r positive the largest error in this range is for $r = .25$: log $(1 + .25) = .223$. For r negative the largest error in this range is for $r = -.25$: log $(1 - .25) = -.287$. Thus on a natural log plot, if two values—log (t) and log (u)—differ by an amount log $(t) -$ log $(u) = r$ that is between $-.25$ and .25, then we know t is approximately $100r\%$ larger than u if r is positive and $100r\%$ smaller if r is negative.

Figure 13 illustrates the use of natural logarithms. The data are the winning times in the 800-meter dash for the Olympics and are analyzed in Chatterjee and Chatterjee (1982). We can clearly see that there was a large decrease of about 5% from 1900 to 1904 and that from 1904 to the present, there has been a steady overall decline of about 10%.

It is not necessary for the entire data range to increase by less than 25% for log base e to be useful. It is sufficient that the values being visually compared differ by no more than $\pm 25\%$. For example, on a time series chart, it may well be that we are primarily interested in the changing values only of points close in time; so it may be useful to use log base e, even if the overall increase of the data is much more than 25%.

5. SUMMARY

Experimentation with graphical methods for data presentation is important for improving graphical communication in science. Several graphical methods are discussed. *Full scale breaks* are suggested as replacements for partial scale breaks, since the full breaks provide a more forceful indication of a scale change that discourages mental visual connections by viewers and actual connections by authors. *Dot charts* are suggested as replacements for bar charts. The replacements allow more effective visual decoding of the quantitative information and can be used for a wider variety of data sets. *Multibased logging* is important for presenting data on log scales; using base 2 or base e in place of base 10 can frequently enhance understanding of the variation in the data.

[*Received May 1983. Revised June 1984.*]

REFERENCES

CHAMBERS, J.M., CLEVELAND, W.S., KLEINER, B., and TUKEY, P.A. (1983), *Graphical Methods for Data Analysis*, Belmont, Calif.: Wadsworth (hard cover), and Boston, Mass.: Duxbury Press (paperback).

CHATTERJEE, S., and CHATTERJEE, S. (1982), "New Lamps for Old: An Exploratory Analysis of Running Times in Olympic Games," *Applied Statistics*, 31, 14–22.

CLEVELAND, W.S. (1984), "Graphs in Scientific Publications," *The American Statistician*, 38, 261–269.

CLEVELAND, W.S., and McGILL, R. (1984), "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods," *Journal of the American Statistical Association*, 79, 531–554.

DEPT. OF THE ARMY (1966), "Standards of Statistical Presentation," Pamphlet No. 325-10, Washington, D.C.: U.S. Government Printing Office.

FIENBERG, S.E. (1979), "Graphical Methods in Statistics," *The American Statistician*, 33, 165–178.

GONZALEZ, M., OGUS, J.L., SHAPIRO, G., and TEPPING, B.J. (1975), "Standards for Discussion and Presentation of Errors in Survey and Census Data," *Journal of the American Statistical Association*, 70 (Part 2), 5–23.

LYDON, P.L. (1961), *The Graphic Primer*, Brooklyn Heights, N.Y.: Pol Lydon, Inc.

SCHMID, C.F. (1983), *Statistical Graphics*, New York: John Wiley.

TUFTE, E.R. (1983), *The Visual Display of Quantitative Information*, Cheshire, Conn.: Graphics Press.

TUKEY, J.W. (1977), *Exploratory Data Analysis*, Reading, Mass.: Addison-Wesley.