THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 62, No. 6, July–August 1983
Printed in U.S.A.

Human Factors and Behavioral Science:

Experiments on Quantitative Judgments of Graphs and Maps

By W. S. CLEVELAND,* C. S. HARRIS,* and R. McGILL*

(Manuscript received August 18, 1982)

Behavioral studies are essential for devising guidelines for effective communication of quantitative information from graphs. Three experiments in which subjects made quantitative judgments from three different kinds of graphs lead to several recommendations: use pastel rather than highly saturated colors on statistical maps; standardize the point cloud size relative to the frame on a scatterplot; scale circles by making the circle area proportional to the variable represented, but expect widely varying judgments of the areas.

I. INTRODUCTION

With the proliferation of computer graphics, there is an increasing reliance on visual displays to convey quantitative information. Maps, graphs, and diagrams have been in use for a long time, but in recent years the variety, complexity, and ease of preparing such visual displays have increased greatly. It is often assumed that visual displays allow people to quickly and accurately appreciate quantitative information and relationships that might be much harder to grasp from other representations, such as tables of numbers, equations, or verbal

* Bell Laboratories.

descriptions. Graphs, for example, are regarded as powerful tools for analyzing data[1] and for presenting data to others.[2]

In preparing a graph, the usual assumption is that if the mathematical relationships are represented accurately in visual form, they will convey the correct quantitative impressions. But this may not be right. Numerous experiments with very simple displays have shown that people's perceptual judgments often differ from physical measurements of such attributes as length, area, orientation, separation, brightness, and color (for reference lists, see Refs. 3 through 5). A much smaller number of studies have used displays that are similar to graphs or maps (see Refs. 6 through 9). Considering the enormous usage and variety of graphs, the number of directly applicable experiments is quite small. More information about human factors in graphical judgment is essential for the design of better visual displays.[10,11]

We summarize here three sets of experiments on quantitative judgments of three kinds of graphs. (More detailed accounts are given in Refs. 6, 7, and 12). The experiments differ in the information being conveyed, the coding of the information, the experimental procedures, the subject populations, and the methods of statistical analysis of the results. In passing, we will mention some useful statistical techniques that may be new to many readers.

II. A COLOR-CAUSED ILLUSION OF AREA

A clear example of erroneous perception of a relatively simple display is shown in a study of the perception of areas of colored regions within a map.[12] Figure 1 is a map of the counties in Nevada. Maps like this, with subsets of the counties colored red or green, were shown to 24 subjects (12 scientists and 12 secretaries, clerks, and craftspeople, all from Bell Laboratories). In each map the total red area differed from the total green area by no more than 0.01 percent. Each subject saw ten variations of the map, with different subsets of counties colored red or green. The maps for one group of 12 subjects had the same partitioning of counties as for the other group, but the counties that were red for one group were green for the other, and vice versa. The maps were paper prints produced by a Bell Laboratories Printing System-Multicolor (PRISM) printer (a computer-driven modified Xerox 6500 copier), using standard options, which include highly saturated colors. Matching the colors with Munsell color chips,[13] under illumination similar to that in which subjects viewed the maps, gave Munsell values of hue = 7.5 red, value (brightness) = 4, chroma (saturation) = 14 and hue = 2.5 green, value = 5, chroma = 12. Thus, the red and green were quite similar in brightness and saturation.

An instruction sheet told the subjects that the colors indicated various geological features in each county. The subjects' task was to

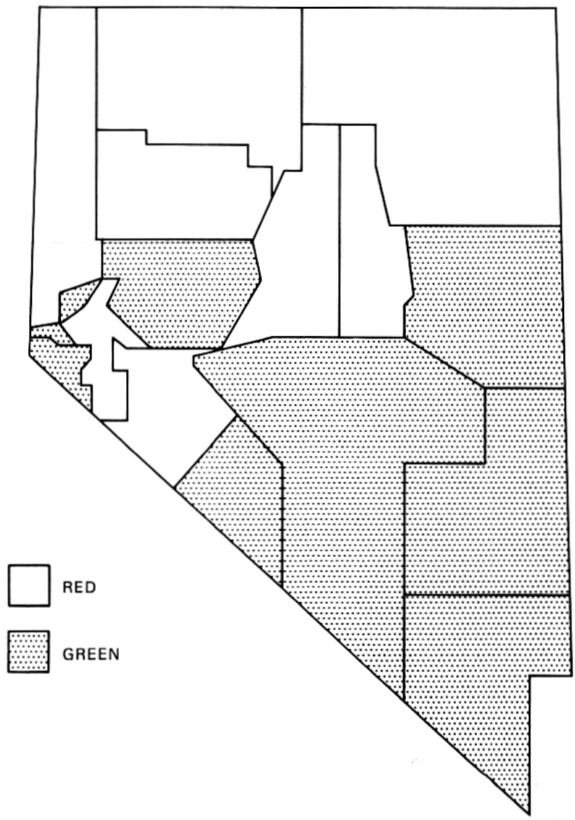

Fig. 1—Map of Nevada, where light and dark areas represent red and green, respectively.

decide which colored area within each map was larger and to mark "more red," "more green," or "same" on a checklist. The outcome was that most of the 24 subjects marked "more red" more frequently than "more green." That is, although the red and green areas were actually equal, they didn't look equal.

Figure 2 shows the data on a trilinear plot,[2] which accommodates within a single graph three variables that sum to a constant (100 percent in this case). Each open circle represents a single subject's data. The vertical axis indicates the percentage of maps for which a subject said the areas of the two colors looked the same. The two diagonal axes measure the percentage of maps in which the red or green area was called larger. Just as in the more familiar Cartesian xy plot, the three percentages for each data point are given by its perpendicular projection onto each of the three axes.

Because each subject made ten judgments, all data points represent multiples of 10 percent. Therefore, to avoid overlapping the values

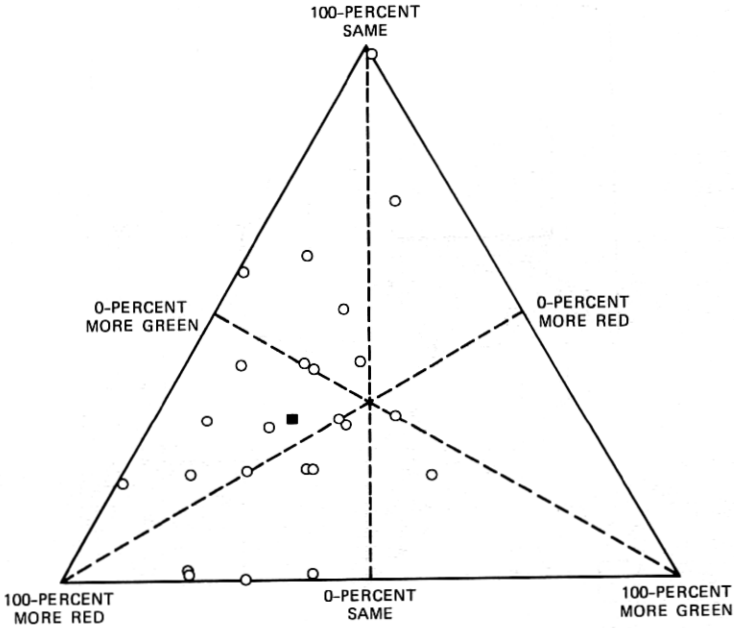

Fig. 2—Area judgments of highly saturated red and green regions.

were jittered: a randomly generated number from the interval −1 percent to +1 percent was added to each value.

First note that most of the 24 subjects' points fall to the left of the vertical axis. This indicates that most subjects judged the red area to be larger more often than they judged the green area to be larger. The filled square is the "robust center of gravity" for the data points. This statistic is calculated by an iterative procedure called "bisquare", using the distances of the data points from the current estimate as input for the next estimate. Using the bisquare[14] produces a statistic that is robust;[15] that is, unlike an arithmetic mean, it is insensitive to outliers, and is highly efficient for a broad class of distributions. The coordinates of this filled square give us another way to summarize the red/green illusion: it falls at about 50 percent "more red" judgments, 20 percent "more green," and 30 percent "same."

To estimate standard errors for the robust center of gravity, the bootstrap method[16] was used. Bisquare estimates were computed on 1000 24-point samples, drawn with replacement. The difference between the green and red coordinates is 49 percent − 22 percent = 26 percent, with a bootstrap standard error of 5.3 percent. Since the bootstrap distribution of the difference was well fit by a normal distribution, the percentage of red-larger judgments is very significantly higher than the percentage of green-larger.

Thus, the likely result of accepting the saturated colors that the PRISM printers normally produce is to make the red areas look too big or the green areas too small. Is there a way around this perceptual distortion? The printer was modified to fill the areas with two unsaturated pastel colors: a half-tone screen composed of red and pale yellow dots and a half-tone screen composed of green and pale yellow dots. The filled square for the robust center of gravity in Fig. 3 shows that for the group as a whole, the percentages of judgments of "more red" and "more green" are nearly equal. Using pastel instead of saturated colors eliminated the overall tendency to call the red area larger than the green.

Additional research could determine whether these findings hold for other pairs of hues, other combinations of brightness and saturation, and other display media (e.g., video displays). For the PRISM printer, it is clear that for accurate judgments of relative area, pastel colors are preferable to the standard saturated red and green.

III. JUDGMENTS OF CIRCLE SIZES

In most of the perceptions that we label "illusions," some extraneous aspect of the stimulus seems to be distorting perception of another aspect. For example, in the familiar Müller-Lyer illusion the arrowheads added to the ends of two horizontal lines distort perception of

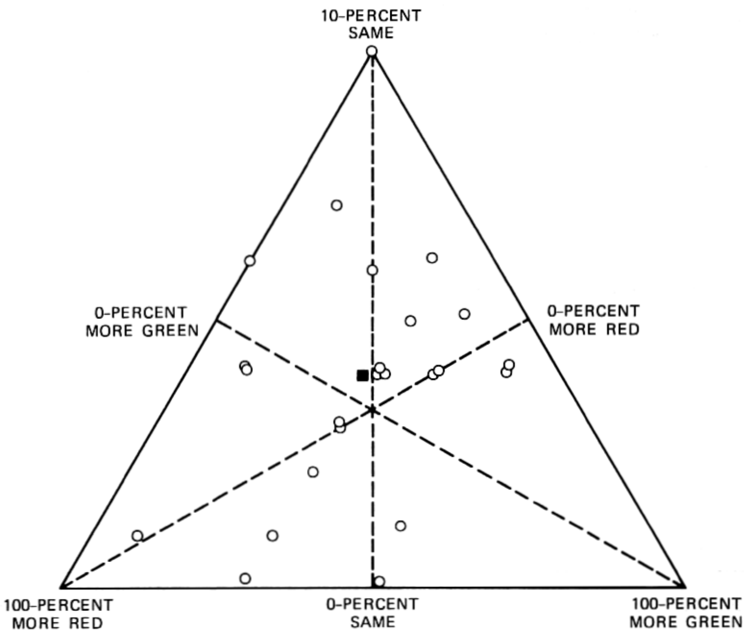

Fig. 3—Area judgments of unsaturated red and green regions.

the lengths of those horizontal lines. Similarly, in the experiment with maps of Nevada, the colors distorted perception of areas. But, in fact, many of our perceptions—perhaps even most of them—are illusions, in the sense that what we see does not match the usual measurements obtained with instruments other than the human observer. For a wide variety of attributes (such as loudness, brightness, tactual roughness, and weight), laboratory studies have found that people's judgments of quantities or intensities seldom vary in direct proportion to these measurements of the physical stimulus. In most cases the relation between subjective judgments and physical measurements can be described by a power function of the form $s = kp^e + c$, where s is the subjective magnitude, p is the physical magnitude, k and c are scaling constants, and e is an exponent that depends on what is being judged, ranging from 0.3 for the brightness of an isolated disk of light to 3.5 for the intensity of an electric shock.[5]

What does this imply for the communication of quantitative information by means of graphs or maps? Consider a common kind of statistical map in which circles of different sizes are used to represent, for example, the number of toll calls from various cities. For laboratory studies of the judgment of area, Stevens (see Ref. 5) gives 0.7 as a typical value for the exponent in the psychophysical function. Such a low exponent would mean that a circle that is double the area of another would be called only 1.6 times as large, and one with five times the area would be called only three times as large, whereas one with 25 percent of the area would be called almost 40 percent as large.

Since the purpose of a map or graph is to convey quantitative relations quickly and accurately, shouldn't areas be scaled to give the correct subjective impressions rather than to be physically correct? Some writers[17,18] have suggested just such a procedure. They recommended scaling the plotted areas to compensate for the low-exponent function found in psychophysical experiments.

However, the exponent found in studies of area judgments might not apply to an actual statistical map. We therefore prepared 24 maplike pages that depicted the average daily telephone toll charges for businesses at different locations in a city.[7] An example is shown in Fig. 4. Fourteen scientists from Bell Laboratories were told that the circle marked with an X represented toll charges of $100. They were to write down, for each of the three circles marked with a dot, their estimate of the dollar amount represented. The word "area" was not used.

The results were quite different from the usual laboratory findings on judgments of area: Most of the exponents in the fitted power functions for individual subjects were closer to 1.0 than to 0.7. There was considerable variation from person to person, with exponents

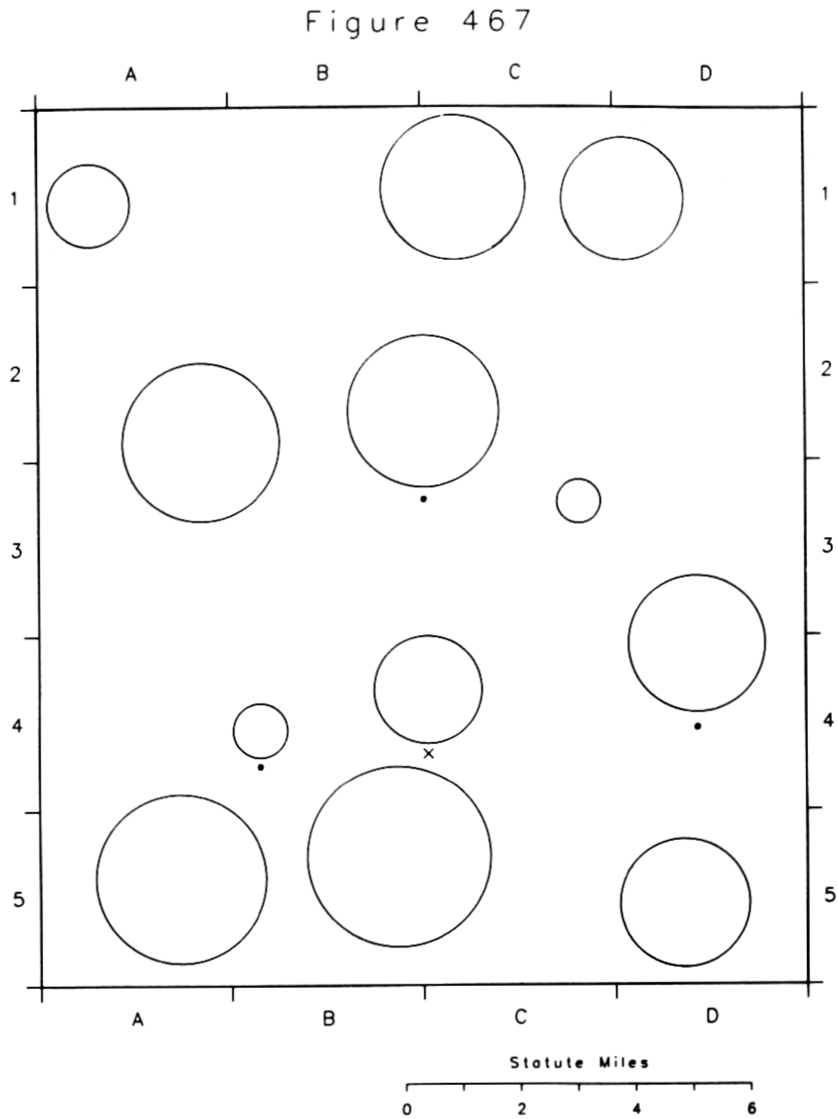

Fig. 4—Maplike display of circles.

ranging from 0.6 to 1.3. This implies that for an actual chart-reading situation, no single form of compensatory scaling of area would give everyone correct impressions. In fact, since the median exponent (0.94) is close to 1.0, these data suggest that the best scale is the simple one with circle areas directly proportional to the represented quantities. (Note that judging diameter or circumference instead of area would yield an exponent of 0.5 for judgments as a function of area. None of our subjects had an exponent that low.)

Why do our exponents differ from those typically found in psychophysical studies? Is it because we asked for a different kind of judgment—dollars represented, rather than area—and displayed a maplike frame and tick marks? We can answer this question because we had the same subjects make area judgments of the same 24 sets of circles on plain pages, without maplike markings (Fig. 5). All of them did this perceptual task after judging the maplike stimuli. The instruction was to judge the areas of the circles, calling the one marked X 100 units of area. We found no difference between the two types of displays and tasks.

Figure 179

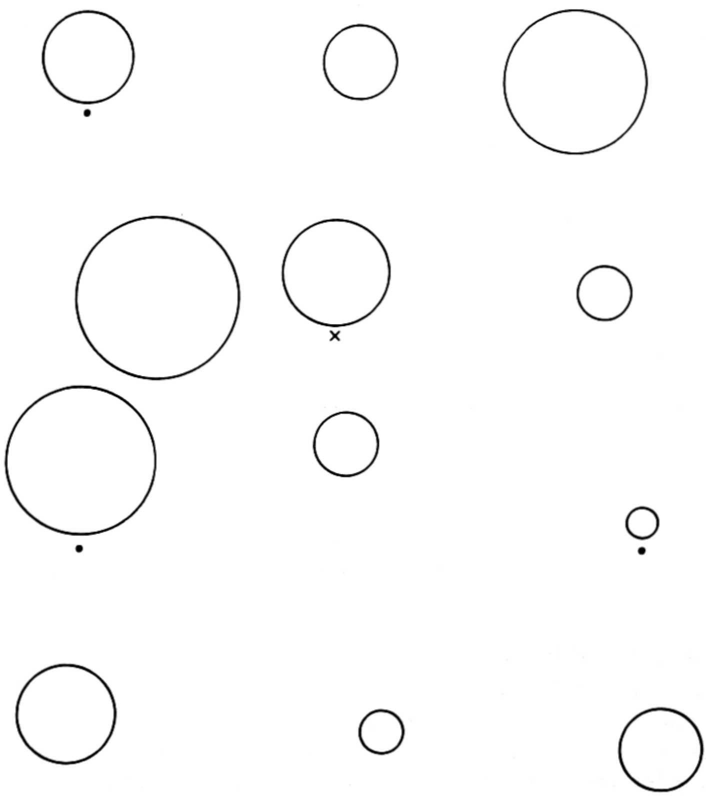

Fig. 5—Display without maplike markings.

Our next thought was that our subjects' scientific training might enable them to judge areas more accurately, on the average, than subjects in previous laboratory experiments. We therefore asked 10 high school students to carry out the same two tasks. Figure 6 presents a comparison of the exponents estimated for the scientists (on the left) and 10 high school students (on the right). Functions of the form $s = kp^e + c$, relating subjective judgment (s) to physical area (p), were fitted to each subject's data. The digits to the left of the colons are the first digits of the estimated exponents (e). Each digit to the right of the colon is the next digit in the exponent for a single subject. The distributions are similar for the scientists and the high school students, with medians at 0.94 and 0.97, respectively. These stem-and-leaf diagrams[19] can be interpreted as enriched histograms. The digits to the left of the colons are the first one or two digits of the estimated exponent, while each single digit to the right of the colon is the next digit in the exponent for a single subject. For example, the stem-and-leaf diagram tells us not only that four scientists had exponents in the range from 1.00 to 1.09, but also that the specific values were 1.00, 1.04, 1.05, and 1.07. There is not much difference between the distributions for the scientists (median = 0.94) and for the students (median = 0.97). Scientific training doesn't have much influence.

The students' data also rule out another possible explanation of the difference from previous psychophysical studies. The scientists judged all of the maplike displays before seeing the ones that contained only circles. When estimating the toll charges, they were free to adopt any strategy they cared to, unlike subjects in psychophysical studies who are told to judge area. Even though the scientists were then asked to judge areas in the second half of the experiment, they might have simply persisted in using the same judgmental strategy as in the first half. This choice of strategy could explain why judgments were the same for maplike displays and plain circles, and also why the results differ from earlier experiments on area estimation. However, half of the students made their area judgments *before* making dollar estimates, and their estimates were quite similar to those made by the students who completed the tasks in the opposite order. The agreement between

EXPONENTS FOR 14 SCIENTISTS	EXPONENTS FOR 10 HIGH SCHOOL STUDENTS
0.5 :	0.5 : 8
0.6 : 3	0.6 : 9
0.7 : 29	0.7 : 4
0.8 : 8	0.8 : 8
0.9 : 01349	0.9 : 7799
1.0 : 0457	1.0 : 3
1.1 :	1.1 : 8
1.2 : 7	1.2 :

Fig. 6—Power-function exponents for judgment of circle sizes.

area and dollar estimates shows that neither the type of judgments, nor the order in which they are made, nor the presence of maplike frames is responsible for the high exponents.

One obvious difference between our experiments and many previous ones is that our subjects always judged circles within a simultaneously visible set, whereas in most psychophysical studies the stimuli are presented one at a time. It may be that higher exponents are obtained when the standard circle is visible along with the circles that are compared with it, instead of being displayed and then removed before the judgments are made. Examination of previous studies offers some support for this hypothesis; for several comparable studies (mostly by psychologists) in which the standard was not visible during the area judgments, the median exponent was 0.7, whereas in several studies (largely by cartographers) in which the standard was always present, the median exponent was 0.9.[7]

These findings suggest that when one is preparing statistical maps, it is probably better simply to make circle areas directly proportional to the quantities represented (exponent = 1) rather than to scale with a very different exponent, as a direct application of psychophysical studies of area judgments might suggest. Future research may enable us to specify what variables contribute to higher or lower exponents, but certainly our multiple-circle displays resemble statistical maps more closely than do circles viewed one at a time.

IV. JUDGED ASSOCIATION IN SCATTERPLOTS

The judgments discussed so far—areas of circles or colored regions on maps—do not require technical training; indeed high school students' judgments proved to be quite similar to those made by scientists. The experiments to be described now, on the other hand, called for subjects with some statistical training. In these experiments[6] all subjects had at least a basic knowledge of statistics (university statistics students and faculty, and practicing statisticians). They were asked to assess the degree of linear association between two variables portrayed by a scatterplot. The most frequently encountered measure of linear association is the correlation coefficient, r. The value of r ranges from 0, when there is no linear association, to $+1$ or -1, when the linear association is perfect and the plotted points fall on a straight line. In many basic statistics courses, r is the only measure of correlation that is taught.

In the study with maps of Nevada, discussed earlier, perceptions of one aspect of the displays—area—was found to be strongly influenced by another aspect—color—which should have been irrelevant. A similar influence was found with judgments of association in scatterplots. The two scatterplots in Fig. 7, projected alternately on a screen, were

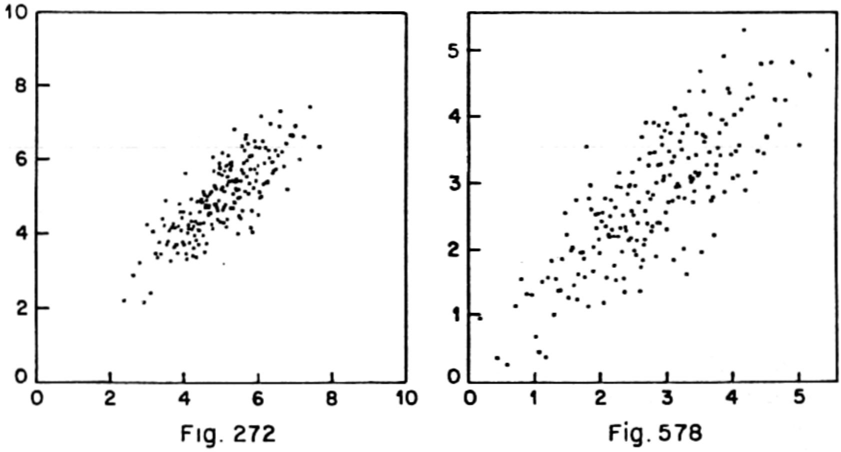

Fig. 7—(Left panel) Scatterplot with correlation r = 0.8. (Right panel) Scatterplot with same correlation as in left panel, but with x and y scales expanded.

shown to 109 subjects. They indicated the degree of association of the two variables by assigning a number from 0 to 100 to each plot. In fact, the correlation is identical for the two plots (r = 0.8); only the scale of the axes, and hence the size of the point cloud, is different. Judgments of the amount of association on the two plots should therefore be identical also. But they are not; judgments of the association portrayed by the left panel of Fig. 7 were generally higher than for the right panel.

To quantify this difference, each subject's estimate for the right panel of Fig. 7 was subtracted from the estimate for the left panel, and the 10-percent trimmed mean was calculated by dropping the largest 10 percent of the differences and the smallest 10 percent and taking the arithmetic mean of the remaining values. (Unlike means, 10-percent trimmed means are robust measures that are not influenced by extreme outliers. Ten-percent trimmed means can be thought of as a compromise between medians, which are nearly 50-percent trimmed means, and means, which are 0-percent trimmed.) The result (after dividing by 100 to bring the judgments into the range 0 to 1) was a difference of 0.068 between the panels of Fig. 7, with a standard error of 0.011. Thus, the estimated association was significantly higher for the smaller point cloud than for the larger one.

A second experiment corroborated this finding. The scatterplots in Fig. 7 were shown to 32 subjects who were told that the correlation coefficients were identical. The subjects were asked whether one plot *looked* more correlated than the other and if so, which one. Sixty-six percent of the subjects reported that the left panel looked more correlated than the right, 22 percent that they looked the same, and only 13 percent that the right panel appeared more correlated. So even explicit information that the two correlations are identical does not dispel the illusory impression that the smaller point cloud depicts a higher correlation.

To obtain more detailed information on judgments of association another study was carried out. In this study 74 subjects each judged 19 scatterplots, assigning a number from 0 to 100 to denote the linear association. The 19 stimuli included 10 with the same axis scales, representing 10 different levels of association, plus an additional 9 with 3 different axis scales (that is, different point-cloud sizes) at 3 of those levels of association. A number of other attributes (such as number of points, standard deviations of the variables, sign of the correlation, and size of the square frame) were identical for all of the scatterplots.

The 10-percent trimmed means across subjects are plotted against actual correlation in Fig. 8. The circle radii portray the standard errors of the trimmed means. Thus the circle areas are proportional to the estimated variances of the trimmed means. The numbers to the left of the circles indicate the point-cloud sizes (1 is the smallest). When

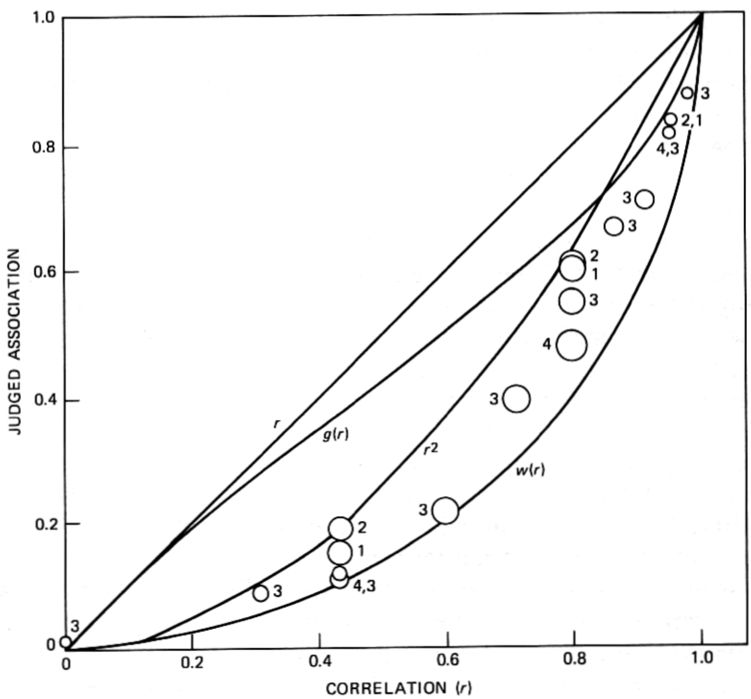

Fig. 8—The 10-percent trimmed means across subjects of judged association divided by 100 for 19 scatterplots are plotted (by the circle centers) against the values of r, the correlation coefficient, of the scatterplots.

two numbers are shown, separated by a comma, two circles are nearly coincident and the first number refers to the circle with the smaller trimmed mean. The information on the plot leads to two conclusions: judged association tends to be higher for smaller point-cloud sizes; judged association is not proportional to any of four standard numerical measures of association [the points do not fall on any of the curves, which plot r, $g(r)$, r^2, and $w(r)$ as a function of r]. Thus, the circles for r near 0.4 and 0.8 corroborate the finding that smaller point clouds tend to elicit higher judgments of association.

The new information is that subjects' judgments are far from proportional to the usual measure of linear association, r. Since the actual correlation coefficient r is given by the horizontal axis, if the subjects had been judging r accurately, the data would have fallen on the 45-degree line. Instead, the points fall well below it. On the average a correlation of 0.4 was judged to be less than 0.2, a difference of more than a factor of 2.

Two other measures of linear association come closer to fitting the data: $g(r) = 1-\sqrt{(1-r)/(1+r)}$ and $w(r) = 1-\sqrt{1-r^2}$ (see Ref. 20). Unlike r, both $g(r)$ and $w(r)$ offer intuitively plausible geometric bases for visual judgments. If we draw the ellipse of the bivariate normal distribution that generated a scatterplot, as in Fig. 9, the ratio of the minor axis to the major axis is $(1-r)/(1+r)$. The smaller this ratio is, the higher the association; so $g(r)$ tells how narrow the ellipse is relative to a zero-correlation circle. The ratio of the area of the ellipse to the area of the rectangle circumscribed around it is $\pi/4\sqrt{1-r^2}$; so $w(r)$ tells how far the ellipse is from filling the rectangle. Unfortunately, neither $g(r)$ nor $w(r)$ fits the data very well, as can be seen from the departure of the data points from both curves in Fig. 8. Another measure, r^2, does not fit the data particularly well either. However, there is a two-parameter family of curves, $1-(1-r)^\alpha(1+r)^\beta$, that includes all four measures of association that we have mentioned (and many others as well). If such a curve is fitted to the data in Fig. 8, the estimates of α and β and their standard errors are 0.71 ± 0.04 and 0.66 ± 0.11, respectively. These estimated parameter values fall between (and are significantly different from) those for $w(r)$ ($\alpha = 0.5$, $\beta = 0.5$) and for r^2 ($\alpha = 1$, $\beta = 1$). [For r, $\alpha = 1$ and $\beta = 0$; for $g(r)$, $\alpha = 0.5$ and $\beta = -0.5$.] This middle position of the parameter estimates is what we would expect since the circles in Fig. 8 lie between r^2 and $w(r)$.

Thus the parameters of the best-fitting curve have values that are significantly different from those of any of the four measures of association that we have considered.

On the face of it, the poor fit of the subjects' judgments to r, r^2, $g(r)$, and $w(r)$ seems to imply that these highly trained subjects do not base

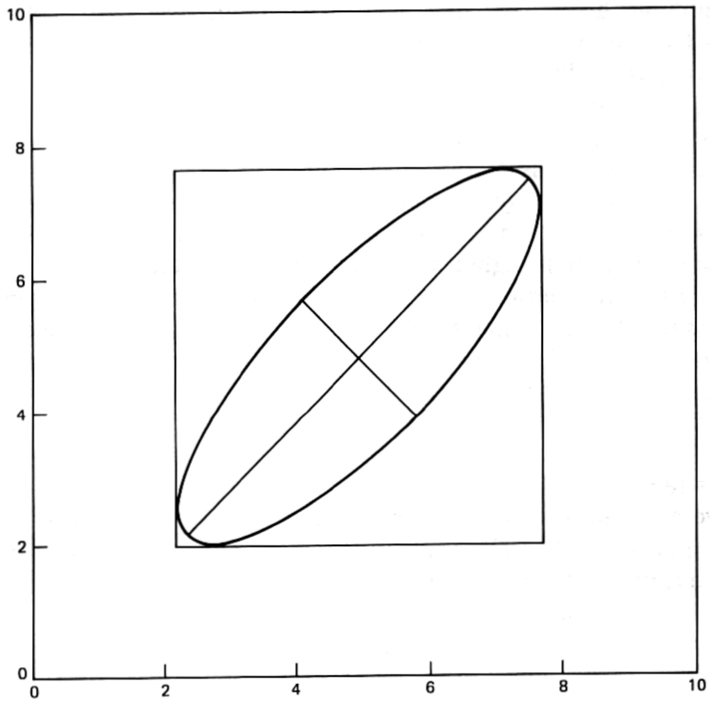

Fig. 9—Geometrical bases for $g(r)$ and $w(r)$.

their estimates of linear association on the visual counterparts of any of these standard measures. However, there is an alternative that remains to be explored. Perhaps subjects do make judgments (for instance, of the ratio of minor to major axis) that are appropriate for one of these measures of association, but misperceive the basic visual attributes on which the judgments are based. For example, if judgments of length were not directly proportional to actual length, then judgments of the ratio of minor to major axis of an ellipse would not be proportional to the actual ratio.

The finding that the size of the point-cloud can have a big effect on judged association means that the axis scaling that many statistical plotting programs apply automatically is not optimal for all purposes. To facilitate comparisons of the degree of association in different plots, it would be wise to make the point-cloud sizes similar; Cleveland and McGill (see Ref. 21) propose a way to do so. The size effect also suggests that when estimates of degree of association are required, the numerical value of a measure of association should also be computed. Even experienced statisticians can have judgments of association affected by extraneous factors.

V. CONCLUSIONS

Behavioral studies of the kind summarized here are essential for devising guidelines for effective communication of quantitative information. These studies confirm that people can make consistent judgments of a wide variety of visual displays. However, those judgments may not match the usual physical measurements of stimulus attributes: People overestimated bright red areas on maps, relative to bright green areas, and judgments of the degrees of linear association in a scatterplot did not agree closely with any of four standard statistical measures of association [r, $w(r)$, $g(r)$, and r^2]. Other findings lead to some recommendations about how to convey quantitative relations more accurately. For example, for output from PRISM plotters, substituting pastel colors for saturated red and green reduced the biasing of area judgments. The finding that smaller point-clouds in a scatterplot are judged to portray higher linear association implies that to permit accurate comparisons of association, axes should be scaled to produce similar point-cloud sizes. And finally, our subjects' judgments of arrays of circles suggest that in statistical maps one should not scale to compensate for the distortions in area judgments that are found when stimuli are viewed one at a time; instead, symbol sizes should be directly proportional to the quantities represented.

With reliance on both old and new forms of visual display becoming increasingly widespread, we will need more behavioral studies like the ones summarized here to guide effective communication of quantitative information.

VI. ACKNOWLEDGMENTS

We acnowledge helpful comments on earlier drafts by Francine Frome, Ram Gnanadesikan, Tom Landauer, Saul Sternberg, Paul Tukey, and several anonymous reviewers.

REFERENCES

1. J. M. Chambers, W. S. Cleveland, B. Kleiner, and P. A. Tukey, *Graphical Methods for Data Analysis*, Belmont, CA: Wadsworth, 1983 (hardcover); and Boston: Duxbury Press, 1983 (softcover).
2. C. F. Schmid and S. E. Schmid, *Handbook of Graphical Presentation*, New York: Wiley & Sons, 1979.
3. S. Coren and J. S. Girgus, *Seeing is Deceiving: The Psychology of Visual Illusions*, Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1978.
4. C. S. Harris (Ed.), *Visual Coding and Adaptability*, Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1980.
5. S. S. Stevens, *Psychophysics—Introduction to its Perceptual, Neural and Social Prospects*, New York: Wiley, 1975.
6. W. S. Cleveland, P. Diaconis, and R. McGill, "Variables on Scatterplots Look More Highly Correlated When the Scales are Increased," Science, *216* (June 1982), pp. 1138–41.
7. W. S. Cleveland, C. S. Harris, and R. McGill, "Judgments of Circle Sizes on Statistical Maps," J. Amer. Stat. Assn., *77* (September 1982), pp. 541–7.

8. H. Wainer and D. Thissen, "Graphical Data Analysis," Ann. Rev. Psychol., *32* (1981), pp. 191–241.
9. L. Wilkinson, "An Experimental Evaluation of Multivariate Graphical Point Representations," Proc. Conf. on Human Factors in Computer Systems, Gaithersburg, Maryland, March 15–17, 1982, pp. 202–9.
10. W. Kruskal, "Visions of Maps and Graphs," Proc. Int. Symp. on Computer-Assisted Cartography, September 21–25, 1975, Amer. Congress on Survey and Mapping and the U.S. Bureau of the Census (1977), pp. 27–36.
11. W. Kruskal, "Criteria for Judging Statistical Graphics," Utilitas Mathematica, *21B* (May 1982), pp. 283–310.
12. W. S. Cleveland and R. McGill, "A Color-Caused Optical Illusion on a Statistical Graph," *The Amer. Statistician*, (1983a), in press.
13. A. H. Munsell, *The Munsell Book of Color—Glossy Finish Collection*, Baltimore: Munsell Color Co., 1966.
14. F. Mosteller and J. W. Tukey, *Data Analysis and Regression*, Reading, Massachusetts: Addison-Wesley, 1977.
15. P. J. Huber, "Robust Estimation of a Location Parameter," Ann. Math. Stat., *35* (February 1964), pp. 73–101.
16. B. Efron, "Bootstrap Methods: Another Look at the Jackknife," Ann. of Stat., *7* (January 1979), pp. 1–26.
17. J. J. Flannery, "The Relative Effectiveness of Some Common Graduated Point Symbols in the Presentation of Quantitative Data," Canadian Cartographer, *8* (December 1971), pp. 96–109.
18. G. F. McCleary, Jr., "In Pursuit of the Map User," Proc. Int. Symp. on Computer-Assisted Cartography, Auto-Carto II, (1975), pp. 238–50.
19. J. W. Tukey, *Exploratory Data Analysis*, Reading, Massachusetts: Addison-Wesley, 1977.
20. M. B. Wilk, unpublished work.
21. W. S. Cleveland and R. McGill, unpublished work.

AUTHORS

William S. Cleveland, A.B., 1965 (Mathematics), Princeton; M.S. and Ph.D., 1969 (Statistics), Yale; Bell Laboratories, 1972—. Mr. Cleveland works on the development of statistical methodology. His current research areas include time series analysis, seasonal adjustment, economic forecasting, atmospheric chemistry, graphical methods for data analysis, software systems for presentation graphs, and graphical perception. Mr. Cleveland is co-author of a book *Graphical Methods for Data Analysis* that was published in 1983. Member, AAAS, National Computer Graphics Association, American Statistical Association (Fellow, Chairman of Statistical Graphics Committee); elected member, International Statistical Institute.

Robert McGill, Bell Laboratories, 1969—. Mr. McGill is a Member of Technical Staff in the Statistics and Data Analysis Research Department. Currently, he is working on graphical techniques and perception of statistical graphics. Member, American Statistical Association, Association for Computing Machinery, National Computer Graphics Association, International Association for Statistical Computing, American Association for Advancement of Science.

Charles S. Harris, B.A., 1959, Swarthmore College; M.S., 1961, Yale University; Ph.D. (Experimental Psychology), 1963, Harvard University; Assistant Professor, University of Pennsylvania, 1964–1966; Bell Laboratories, 1966—. Mr. Harris has carried out research on various phenomena of visual and kinesthetic perception, including motion illusions, adaptation to visual distortions, selective attention, color and spatial-frequency aftereffects, and perceptual facilitation by object-like context. He has edited a book entitled *Visual Coding and Adaptability.* He is currently a member of the Human Information-Processing Research Department. Member, Sigma Xi, Phi Beta Kappa, Psychonomic Society, Eastern Psychological Association.