DESIGNING EFFECTIVE GRAPHS*

Edward W. Frees† and Robert B. Miller‡

ABSTRACT

Actuaries, like other business professionals, communicate quantitative ideas graphically. Because the process of reading, or decoding, graphs is more complex than reading text, graphs are vulnerable to abuse. To underscore this vulnerability, we give several examples of commonly encountered graphs that mislead and hide information. To help creators design more effective graphs and to help viewers recognize misleading graphs, this article summarizes guidelines for designing graphs that show important numerical information. When designing graphs, creators should:

(1) Avoid chartjunk
(2) Use small multiples to promote comparisons and assess change
(3) Use complex graphs to portray complex patterns
(4) Relate graph size to information content
(5) Use graphical forms that promote comparisons
(6) Integrate graphs and text
(7) Demonstrate an important message
(8) Know the audience.

Some of these guidelines for designing effective graphs, such as (6), (7) and (8), are drawn directly from principles for effective writing. Others, such as guidelines (3), (4) and (5), come from cognitive psychology, the science of perception. Guidelines (1) and (2) have roots both in effective writing and in graphical perception. For example, the writing principle of brevity demonstrates how eliminating pseudo three-dimensional perspectives and other forms of chartjunk improve graphs. As another example, the writing principle of parallel structure suggests using small multiple variations of a basic graphical form to visualize complex relationships across different groups and over time.

To underscore the scientific aspect of graphical perception, we examine the process of communicating with a graph, beginning with a sender's interpretation of data and ending with a receiver's interpretation of the graph. In keeping with scientific tradition, this article discusses several studies in the literature on the effectiveness of graphs.

We conclude that the actuarial profession has many opportunities to improve its practice, making communication more efficient and precise.

1. INTRODUCTION

Like other business professionals, actuaries communicate ideas orally and in writing, as well as through presentations, which are interactive forms of communication that encompass oral and written messages. Actuaries, as well as other financial analysts, communicate ideas with important quantitative components. Writers express quantitative ideas as (1) numbers within paragraphs, (2) numbers within tabular forms, (3) functional relationships such as equations, and (4) data or equations as graphs.

Graphs are a simple yet powerful medium for written communication of quantitative ideas. Graphs can present a large amount of data in a small space, express important relationships between quantities, compare different sets of data, and describe data, thus providing a coherent picture of complex systems.

*This paper was originally presented at the 32nd Actuarial Research Conference, held August 6–8, 1997, at the University of Calgary, Calgary, Alberta, Canada.

†Edward W. (Jed) Frees, F.S.A., Ph.D., is Time Insurance Professor of Actuarial Science, School of Business, University of Wisconsin–Madison, 975 University Avenue, Madison, Wisconsin 53706, e-mail, jfrees@bus.wisc.edu.

‡Robert B. Miller, Ph.D., is Professor of Business and Statistics, University of Wisconsin–Madison, 975 University Avenue, Madison, Wisconsin 53706, e-mail, rbmiller@facstaff.wisc.edu.

Graphs do more than merely state an idea; they demonstrate it.

Graphs are powerful because they are flexible, but flexibility can be a disadvantage because of the potential for abuse. Well-accepted references dealing with methods of quantitative data presentation mitigate opportunities for abuse. *The Chicago Manual of Style* (1993), a standard reference, discusses presentation of in-text data, and Ehrenberg (1977) and Tufte (1983) discuss presentation of tabular data. In contrast, we focus on data presentation through *graphical* displays.

This article seeks to improve actuarial practice as it relates to graphical displays. We intend to: (1) demonstrate the importance of graphical displays, (2) provide guidelines to improve graphical practice, and (3) introduce some of the scientific underpinnings of good graphical practice. The agenda is ambitious, yet the goal of this article is to provide practicing actuaries with basic tools that they can use to become critical consumers and effective producers of graphs. We also hope that readers will adopt our enthusiasm and wish to explore the graphical design literature on their own.

An important theme of this article is that principles of vigorous writing can and should be applied to the practice of making effective graphs. The *Elements of Style* (Strunk and White 1979, p. xiv) summarizes vigorous writing:

> Vigorous writing is concise. A sentence should contain no unnecessary words, a paragraph no unnecessary sentences, for the same reason that a drawing should have no unnecessary lines and a machine no unnecessary parts. This requires not that the writer make all his sentences short, or that he avoid all detail and treat his subjects only in outline, but that every word tell.

White attributes this quotation to William Strunk. White calls it "a short, valuable essay on the nature and beauty of brevity—sixty-three words that could change the world." We argue that brevity is especially important when making effective graphs. This was also understood by Strunk; as noted above, he said "a drawing should contain no unnecessary lines. . . ." We use the term *chartjunk*, introduced by Tufte (1983), for any unnecessary appendage in a graph.

Vigorous writing principles other than brevity also apply to the practice of making effective graphs. Just as with writing, effective graphs are the result of repeated revising and editing. Poorly designed graphs can and do hide information and mislead. Fancy or pretentious graphs are distracting when simpler graphs suffice.

Although the principles of effective writing are valuable, they are not sufficient for producing effective graphs. Writing is processed in a serial manner, word by word, sentence by sentence, with a beginning and an ending. The process of "reading," or *decoding*, a graph is nonlinear and more complex. The additional complexities mean that even authors who follow effective writing practices may produce ineffective graphs. Often the form of written prose is the sole determinant of its value, whereas in graphics the communication process plays the dominant role. We assume that readers are familiar with effective writing forms. Thus, we first review the communication process in which a graph plays a crucial role.

To underscore the importance of effective graphical design, Section 2 provides several illustrations of graphs that hide information and are misleading; the defects illustrated are more serious drawbacks than mere chartjunk. The Section 2 illustrations motivate the need for additional guidelines and methods for constructing effective graphs.

Section 3 introduces eight important guidelines for creating and viewing graphs. Although the guidelines do not provide a panacea for all graphical defects, they do provide business professionals such as actuaries with a key checklist for creating effective graphs. The guidelines are organized so that the first two, on chartjunk and the use of multiples, are based on both effective writing and graphical perception perspectives. Guidelines Three, Four and Five are related primarily to the graphical perception literature, whereas Guidelines Six, Seven and Eight are based primarily on effective writing principles.

As with effective writing, questions of style enter into the discussion of what is and what is not an effective graph. Many style decisions are based upon accepted practices without a firm scientific foundation. However, the process of perceiving graphs has been the subject of inquiry in several scientific disciplines, including psychophysics, cognitive psychology, and computational visions (Cleveland 1995, ch. 4). Section 4 illustrates some types of experimental evidence for determining an effective graphical form based on both the receiver and the graph itself as units of study. Section 4 also illustrates how such mainstays of business publications as bar charts and pie charts are poor communicators of numerical information.

Section 5 contains concluding remarks and descriptions of some resources for actuaries who wish to learn more about designing effective graphs.

Most readers are removed from the detailed data summarized by a graph. Several difficulties and misconceptions can arise owing to the distance between the original data and a viewer's interpretation of the graph. Figure 1 illustrates the challenge of communicating with a graph. The sender (and creator) of the graph has a message derived from an interpretation of data. Although a few graphs communicate raw data, the primary purpose of most graphs is to communicate the sender's interpretation. The message the sender intends is encoded in a graph and passed on to the receiver.

In general, the receiver is party to neither the exact interpretation intended by the sender nor the raw data. Thus the receiver must decode the graph and develop an interpretation of its message. Two issues arise:

- Whether the interpretation constructed by the receiver is congruent to the interpretation of the sender
- Whether the receiver's interpretation is consistent with and supported by the data.

The first issue depends on the skill with which the sender constructs the graph and the skill with which the receiver decodes it. A poorly constructed graph can hide or distort the sender's message. A graph that is hard to read can discourage the receiver from spending the time necessary to decode the message correctly. The receiver can ignore or misinterpret a graph that is not constructed with care.

The second issue depends not only on the skills mentioned above but also on the skill with which the sender draws meaning from the data. How carefully does the sender document the process of interpretation? Is this communicated to the receiver? Is the receiver capable of assessing the extent to which the graph is a credible summary of the data? Failure at any of these points could result in the receiver ignoring or misinterpreting the graph.

This article assumes that the graphs included in business communications are the subject of scrutiny by serious readers. Graphs that appear quickly on the television screen, a flip chart or presentation package are designed to attract attention and to entertain the viewer. Design, rather than information, considerations dominate these media. We focus instead on graphs that are part of professional writing and are designed to inform. As with effective writing, we assume that in creating graphs ". . . one must believe—in the truth and worth of the scrawl, in the ability of the reader to receive and decode the message" (Strunk and White 1979, p. 84).

We now turn to examples of graphs that mislead.

2. GRAPHIC DESIGN CHOICES MAKE A DIFFERENCE

As noted by Schmid (1992), the ancient proverb "One picture is worth ten thousand words," when applied to graphs might well read, "One picture *can be* worth ten thousand words *or figures*." Graphic potential is not easily realized. Because of their flexibility, graphs too easily render visual displays of quantitative information that are uninformative, confusing or even misleading.

Examples 2.1 through 2.5 illustrate five different types of deceptive graphs. In each case, the data were not altered nor were different dimensions of the data portrayed. The common theme of the examples is that, by altering only the data scales, the creator can alter dramatically a viewer's interpretation.

Example 2.1 Including Zero To Compress Data

Figure 2 shows a time series of the percentage of full-time-equivalent workers employed in the insurance industry. The annual data, 1948–1993, are from the National Income and Product Accounts produced by

Figure 1

Flow Chart of the Process of Communicating with a Graph. The graph is a crucial intermediary in the process of communicating data interpretation to the receiver.

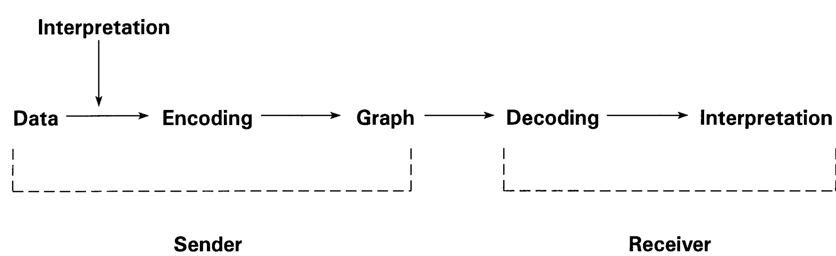
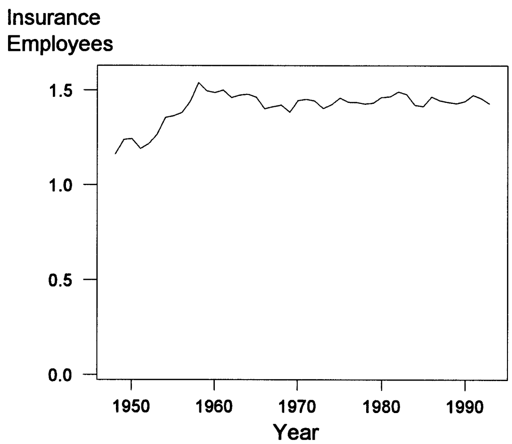


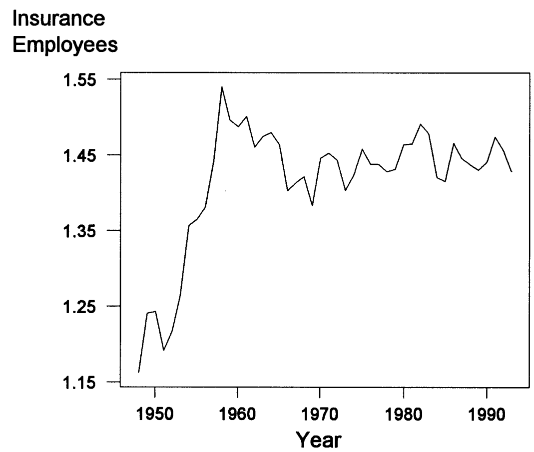
Figure 2

Annual Insurance Employees, 1948–1993.

"Insurance employees" is the percentage of full-time-equivalent employees who are working for insurance carriers. Allowing the data to determine the scale ranges reveals interesting aspects of the data.



(a) A stable insurance industry



(b) The insurance industry work force increased dramatically in the 1950s

the Bureau of Labor Statistics. The left-hand panel, Figure 2a, provides the impression of a stable employment environment for the insurance industry. Including zero on the vertical axis produces this seeming stability. By doing this, most of the graph is devoted to white space that does not show the variability in the data. In contrast, the right-hand panel, Figure 2b, uses the data to set the range on the axes. This panel clearly shows the large employment increases in the years following the Korean War, circa 1952. It also allows the reader to see the employment declines that the insurance industry has suffered in the last three years.

Example 2.1 is similar to a popular illustration from Huff's well-known *How to Lie with Statistics* (Huff 1954). The point is that motivation external to the data, such as including zero on an axis, can invite us to alter the data scale and change a viewer's interpretation of the data. As Example 2.2 shows, creators of graphs can also alter a viewer's interpretation by changing both scales of a two-dimensional graph.

Example 2.2 Perception of Correlation

Figure 3 relates risk management cost effectiveness to firm size. These data are from a survey of 73 risk managers of large, U.S.-based, international firms that was originally reported in Schmit and Roth (1990). The data are analyzed in Frees (1996, ch. 7). Here, the measure of risk management cost effectiveness, firm cost, is defined to be the logarithm of the firm's total property and casualty premiums and uninsured losses as a percentage of total assets. The firm size measure is total assets in logarithmic units.

The left-hand panel, Figure 3a, shows a negative relationship between firm costs and firm size, as anticipated by Schmit and Roth. The correlation coefficient between the two variables is -0.64. The data are in a small center portion of Figure 3b when compared to the left-hand panel, Figure 3a. Figure 3a uses the data to determine the axes and thus shows more patterns in the data. As Cleveland, Diaconis, and McGill (1982) show, the scaling makes the data in the right-hand panel appear more correlated than in the left-hand panel.

Change of scales can also alter the viewer's perception of trend in time series data, as illustrated in Example 2.3.

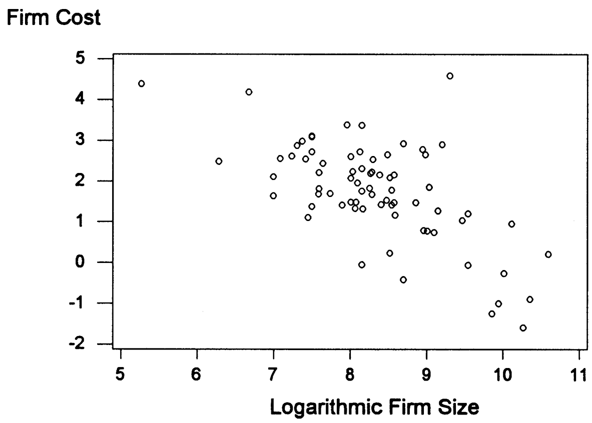
Example 2.3 Transforming to a Logarithmic Scale

Figure 4 exhibits a time series of the U.S. credit insurance market over 1950–1989. These data are analyzed in Frees (1996) and are originally from the *Life Insurance Fact Book* (1990). When the amount of

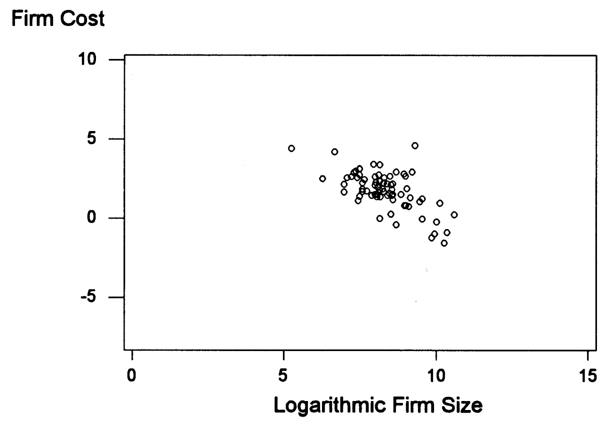
Figure 3

Scatterplot of Cost Effectiveness of a Firm’s Risk Management Practices Versus Firm Size.

The data represented in each figure are the same. However, the wider scales in Figure 2.2b suggest that the data are more highly correlated.



(a) The data in this figure appear less correlated.

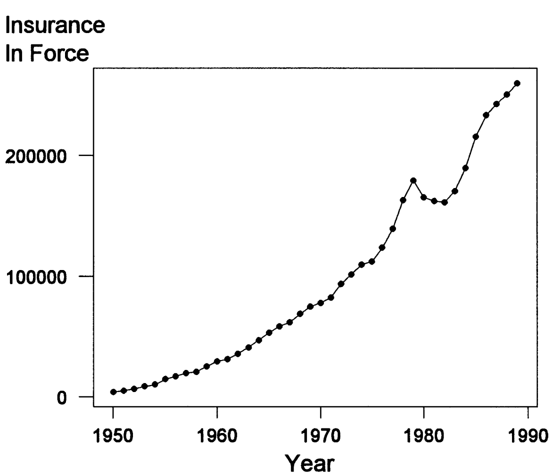


(b) The data in this figure appear more correlated.

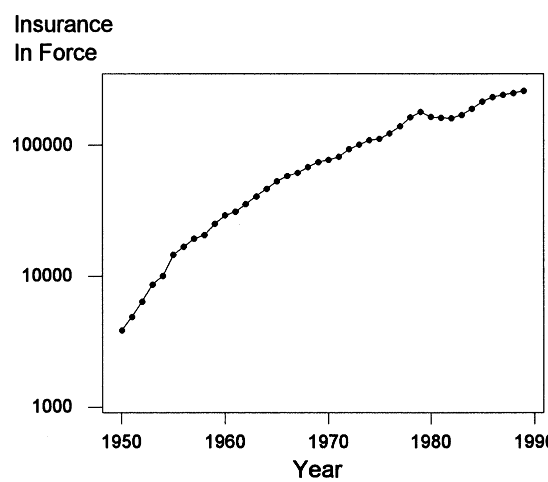
Figure 4

Annual U.S. Credit Life Insurance in Force, 1950–1989.

Different vertical scales give different impressions of the rate of growth over time.



(a) U.S. credit life insurance market exploding.



(b) U.S. credit life insurance market leveling off.

insurance is examined on a linear scale in Figure 4a, the credit insurance market appears to be expanding rapidly. However, Figure 4b shows that, when examined on a logarithmic scale, the market is leveling off. As discussed, for example, in Frees (1996, ch. 6), changes on a logarithmic scale can be interpreted as proportional changes. Thus, Figure 4a shows the market is increasing rapidly, and Figure 4b shows that the rate of increase is leveling off. These messages are not contradictory, but viewers must interpret each graph critically to understand the intended message.

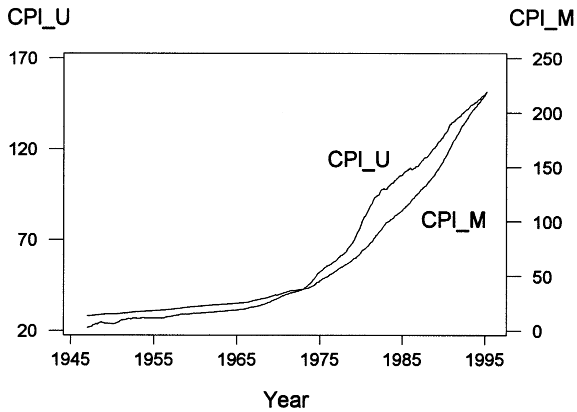
Example 2.4 Double Y-Axes

Figure 5 displays two measures of inflation that are produced by the Bureau of Labor Statistics. On the left-hand axes are CPI_U, the consumer price index for urban consumers. On the right-hand axes are

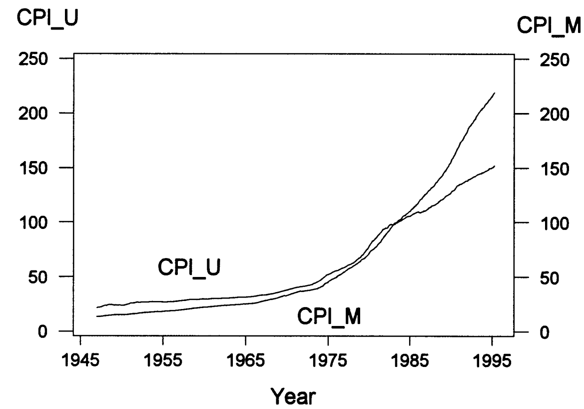
Figure 5

Monthly Values of the Overall Consumer Price Index (CPI) and the Medical Component of the CPI, January 1947 through April 1995.

Different scale ranges alter the appearances of relative growth of the two series.



(a) Overall Consumer Price Index (CPI) is similar to the medical component of the CPI.



(b) Overall Consumer Price Index (CPI) is increasing more slowly than the medical component of the CPI.

CPI_M, the consumer price index for medical components of the overall index. Each series consists of monthly values ranging from January 1947 through April 1995.

The left-hand panel, Figure 5a, suggests that the CPI_U and the CPI_M begin and end in approximately the same position, thus implying that they have increased at about the same rate over the period. The creator could argue that each index measures the value of a standard bundle of goods, thus justifying the argument for using a different scale for each series.

The right-hand panel, Figure 5b, provides a more useful representation of the data by using the same scale for each series. Here, CPI_M begins lower than CPI_U and ends higher. That is, the medical component index has increased more quickly than the index of prices for urban consumers. Other patterns are also evident in Figure 6: each series increased at roughly the same rate over 1979–1983 and CPI_M increased much more quickly from 1983 to 1994 when compared to 1948–1979.

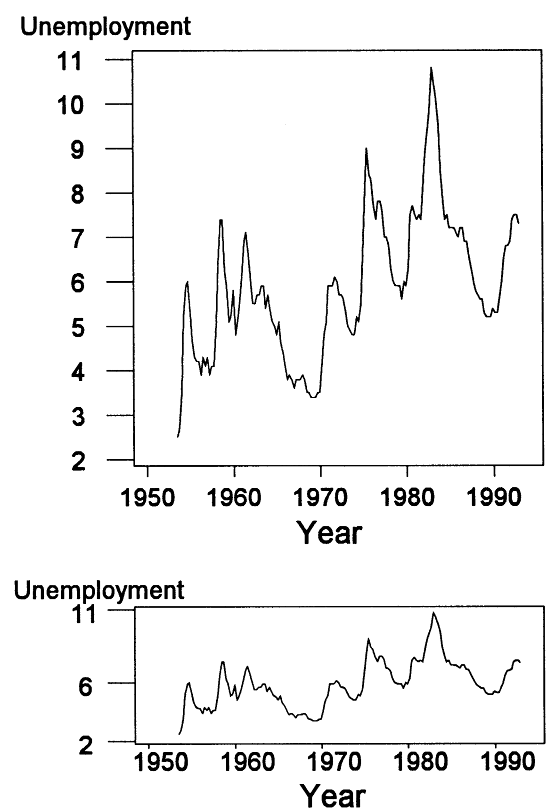
Example 2.5 Aspect Ratio

Figure 6 shows a time series plot of the monthly unemployment rate, April 1953 through December 1992. The unemployment rate is the percentage of unemployed civilian labor force, seasonally adjusted. It is part of the Household Survey produced by the

Figure 6

Time Series Plot of Quarterly Values of the U.S. Unemployment Rate, 1953–1992.

The lower panel displays a feature that is not evident in the upper panel; unemployment declines more slowly than it rises.



Bureau of Labor Statistics, Department of Labor. This series was analyzed in Frees et al. (1997). The top panel of Figure 6 shows that the unemployment rate averaged 5.9% with a peak of 10.8% in the fourth quarter of 1982 and a minimum of 2.7% in the third quarter of 1953.

The two panels in Figure 6 differ only in their shape, not in the scaling of either variable or in the relative amount of space that the data take within the figure frame. To differentiate these two shapes, we can use the concept of a figure's *aspect ratio*, defined to be the height of the data frame divided by its width (some sources use the reciprocal of this value for the aspect ratio). The data frame is simply a rectangle whose height and width just allow the graph *of the data* to fit inside. To illustrate, in the upper panel in Figure 6, the length of the vertical side is equal to the length of the horizontal side. In the lower panel, the vertical side is only 25% of the horizontal side.

Both panels show that the unemployment series oscillated widely over this 39-year period. The lower panel, however, displays a feature that is not apparent in the upper panel; the rise to the peak of an unemployment cycle is steeper than the descent from the peak. Within each unemployment cycle, the percentage of workers unemployed tends to rise quickly to a maximum and then to fall gradually to a minimum. This behavior is surprisingly regular over the almost 39-year period displayed in the plot.

Different aspect ratios can leave substantially different impressions on the eye, as Figure 6 illustrates. Thus, the aspect ratio can be chosen to emphasize different features of the data.

3. DESIGN GUIDELINES

Understanding the issues illustrated in Section 2 can help actuaries and other business professionals create and interpret graphs. This section presents eight guidelines for designing effective graphs. One of our main points is that current practice is not in accord with these guidelines. Thus, we anticipate that not all of our readers will find the demonstrations of the guidelines visually appealing, but, as stated in Section 1, many of the guidelines are based on a scientific foundation outlined in Section 4. "Intuition" is something we learn and cultivate; progress in science does not always conform to current intuition. It was widely believed at one time that the earth was flat and that the sun revolved about the earth. The demonstrations of this section may or may not be immediately intuitive, but they are logical conclusions from the design guidelines advocated here.

Guideline One: Avoid Chartjunk

In Section 1, we defined chartjunk to be any unnecessary appendage in a graph. Creators of graphs who use chartjunk lower their credibility with serious receivers. Even when senders convey a correct interpretation accompanied by chartjunk, they ask receivers to process and properly ignore the chartjunk. If chartjunk is part of the default, or easily used, options of a software package, then the sender can clutter a graph, or even make a graph misleading, simply by punching a button.

Senders who avoid chartjunk raise their credibility. They ask receivers to look only at meaningful characters and marks. Senders may have to spend considerable time with their software to make effective graphs, but the respect and attention of their receivers reward them. Another way to avoid chartjunk is not to use a graph at all if a few words will do. If the message in a graph can be summarized in a few words, then the graph is not needed. Avoid pictures that are not worth ten thousand words!

Avoiding chartjunk is based in part on the concept of brevity in vigorous writing principles. From the graphical perception viewpoint, avoiding chartjunk reduces the noise when communicating between the graph's sender and receiver. Thus, this guideline is important because it has roots in both writing and perception principles.

Example 3.1 Premium Receipts of Life Insurance Companies

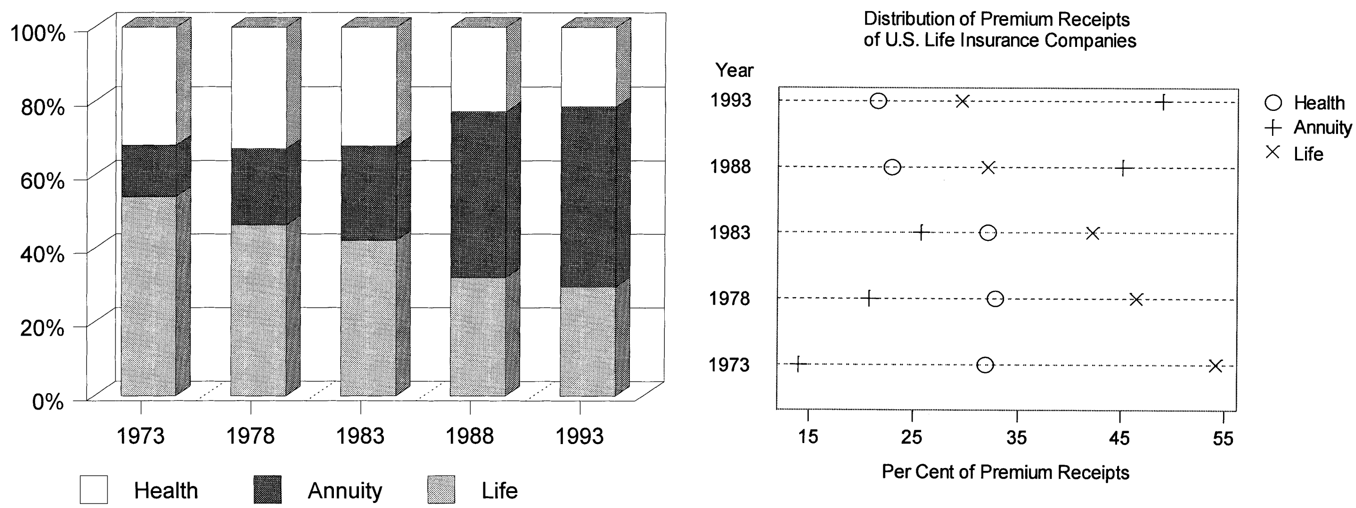
Figure 7a is an adaptation of a graph on page 69 of the *Life Insurance Fact Book* (1994). The graph reports 15 bits of information: 5 years and 2 percentages for each year (a third percentage is found by subtraction). A three-dimensional box represents each percentage, and each box displays different shadings to represent the three lines of business: health, annuity and life. These figures could be reported compactly in a small table. However, granting that a graph may help the receiver appreciate trends in the figures, the graph's simplicity should reflect the simplicity of the information available in the figures. In particular, a small plotting symbol suffices to report a percentage. A three-dimensional, shaded box is hardly called for. It is interesting that the three-dimensional box was an "innovation" in 1994. Earlier editions of the *Fact Book* used two-dimensional boxes. The volume of chartjunk took a big jump in 1994.

Figure 7b is a *dot plot*, discussed by Cleveland (1994). Different plotting symbols show the different

Figure 7

Distribution of Premium Receipts, 1973–1993.

The excessive chartjunk of (a) hides the large change in distribution types between 1983 and 1988.



(a) The three-dimensional stacked bar chart is a poor graphical form for making comparisons over time and across lines of business.

(b) The dot plot allows for direct comparison over time and across lines of business.

lines of business. The tick marks on the lower horizontal axes help us estimate the percentages, and the light, dotted grid lines help us scan across the graph to the plotting symbols of interest. The major shifts, and the approximate magnitudes of the shifts, that happened between 1983 and 1988 are clear here.

Guideline Two: Use Small Multiples to Promote Comparisons and Assess Change

Statistical thinking is directed towards comparing measurements of different entities and assessing the change of a measurement over time or some other unit of measurement. Graphical displays are inherently limited when portraying comparisons or assessing changes because they are static, two-dimensional media. Graphs that contain multiple versions of a basic graphical form, each version portraying a variation of the basic theme, promote comparisons and assessments of change. By repeating a basic graphical form, we promote the process of communication.

Tufte (1997) states that using *small multiples* in graphical displays achieves the same desirable effects as using parallel structure in writing. Parallel structure in writing is successful because it allows readers to identify a sentence relationship only once and then focus on the meaning of each individual sentence element, such as a word, phrase or clause. Parallel structure helps achieve economy of expression and draw together related ideas for comparison and contrast. Similarly, small multiples in graphs allow us to visualize complex relationships across different groups and over time.

The Section 2 figures illustrated the use of small multiples. In each figure, the two plots portrayed were identical except for the change in scale; this use of parallel structure allowed us to demonstrate the importance of scaling when interpreting graphs. Example 3.2 below illustrates another application of small multiples in graphical displays, Cleveland's (1993) multiway dot plot.

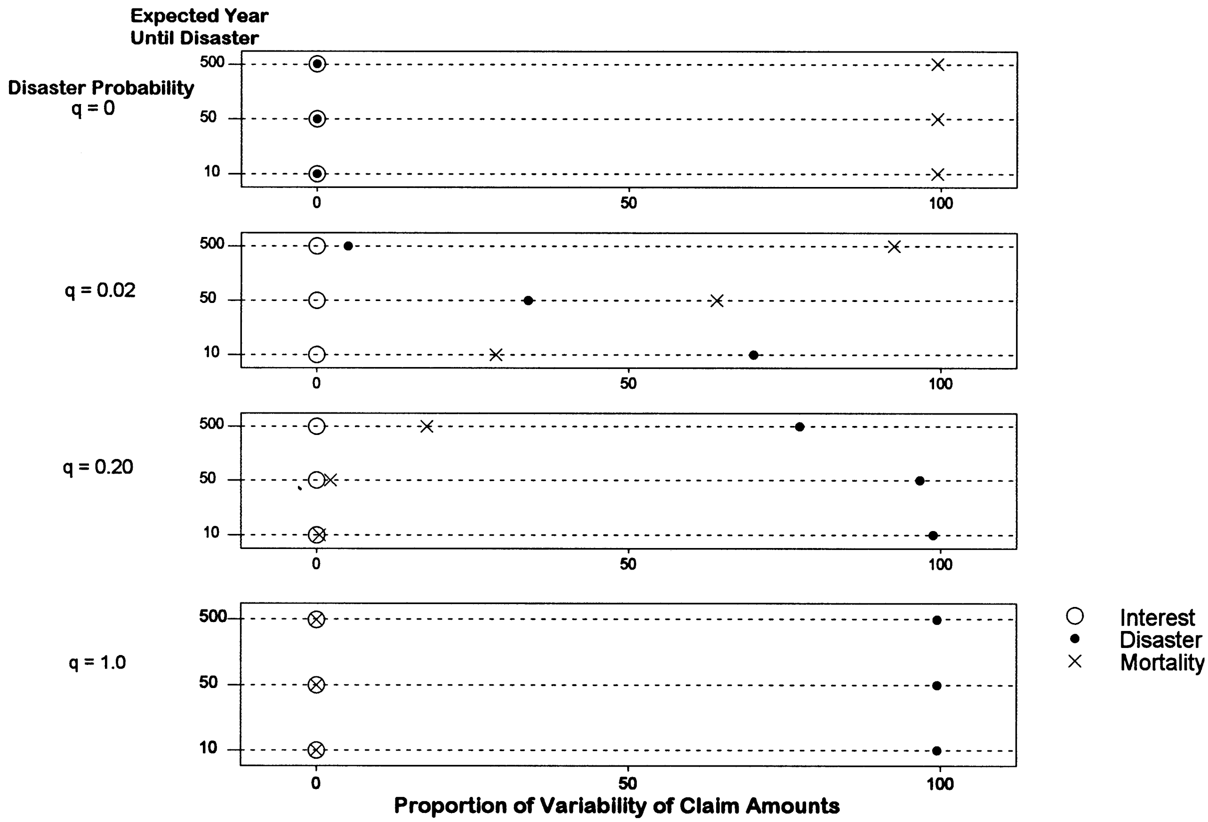
Example 3.2 Relative Importance of Risk Source

Figure 8, called a *multiway dot plot*, demonstrates conclusions reached by using a model introduced in Frees (1998) concerning the relative importance of risk sources within a block of short-term insurance contracts. The risk sources are the stochastic interest environment, the frequency of claims (mortality), and the possibility of a catastrophic event (disaster) occurring. The relative importance of these three risk

Figure 8

The Relative Importance of Risk Sources.

This complex graph allows us to visualize differences over sources of risk (interest, disaster and mortality), expected year until disaster, and probability of disaster. The multiway dot plot demonstrates how quickly the importance of the disaster component increases as the probability of disaster increases.



sources is considered by letting two parameters of interest vary. These parameters are the expected year until disaster and, in the event of disaster, the expected proportion (probability) of policyholders that will succumb to disaster.

Figure 8 shows that when no policyholders succumb to disaster ($q = 0$), then the frequency component, mortality, dominates the other risk sources. At the opposite extreme, when all policyholders succumb to disaster ($q = 1$), then the disaster component dominates the other risk factors. This is true even when the expected time until disaster is 500 years! For the intermediate cases, when either the expected proportion of policyholders succumbing to disaster increases or the expected year until disaster decreases, the importance of the disaster component increases at the expense of the mortality component. Because of the short-term nature of the contract considered, the interest component does not play an important role in Figure 8.

This story of relative importance could not be told using analytic expressions because of the complexity of the underlying models. The story behind Figure 8 could be told, however, using tabular displays. The advantage of Figure 8 is that it allows the viewer to make comparisons over three different risk sources when two parameters of interest vary. Although such comparisons are possible with tabular displays, graphical displays are more effective devices.

Guideline Three: Use Complex Graphs to Portray Complex Patterns

Many authors believe that a graph should be simple and immediately understood by the viewer. Simple graphs are desirable because they can deliver their message to a broad audience and can be shown quickly and digested immediately. Although this notion may be appropriate for popular writing, for professional writing the concept of instant understanding is limiting in that it precludes the notion that graphs demonstrate complex ideas. Complex patterns should be portrayed as simply as possible, although the patterns themselves should not be unnecessarily simplified.

One way for a graph to represent complex patterns is for some of its basic elements to serve more than one purpose. Tufte (1983) called such elements *multifunctioning*. For example, we can use plotting symbols to represent not only elements corresponding to the horizontal and vertical scales but also a level of a categorical variable.

Example 3.3 Frequency and Severity of Hospital Costs

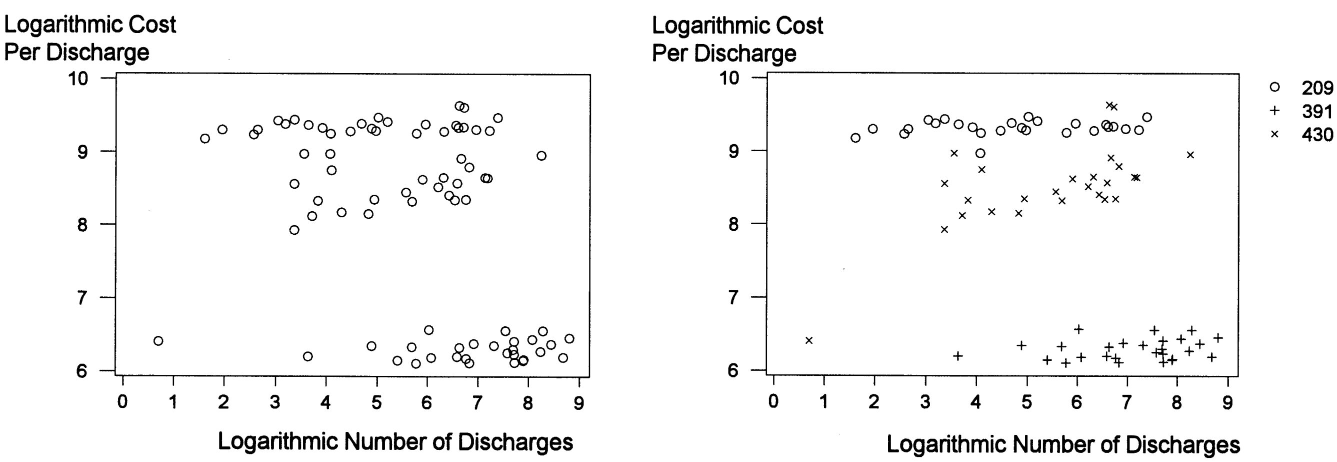
Figure 9 displays the relationship between average hospital costs and frequency of hospital usage. These data for the year 1989 were obtained from the Office of Health Care Information, Wisconsin's Department of Health and Human Services, and are further analyzed in Frees (1996). The data represent averages over the state of Wisconsin, broken down by nine health service areas, three types of payer (fee for service, health maintenance organization, and other) and three types of diagnosis-related groups (DRGs). The three DRGs, numbers 209, 391 and 430, represent major joint and limb reattachment, normal newborns, and psychoses, respectively. Each plotting symbol in Figure 9 represents a combination of health service area, type of payer, and type of DRG. The horizontal axis provides the number of patients admitted in 1989 for each combination, in natural logarithmic units. The vertical scale provides the average hospital cost per discharge for each combination, in natural logarithmic units.

The story in the left-hand panel, Figure 9a, is one of increased economies of scale. That is, combinations of health service areas, type of payer, and DRG that have a larger number of patients, measured by discharges, have lower costs. A substantial negative relationship is evident in Figure 9a; the correlation coefficient is -0.43. This is true despite the aberrant point in the lower left-hand region of Figure 9a. The aberrant point is less important economically than

Figure 9

Logarithmic Cost per Discharge Versus the Logarithmic Number of Discharges.

By adding a plotting symbol code for the level of DRG, the three distinct groups are evident. The three DRGs, 209, 391, and 430, represent major joint and limb reattachment, normal newborns and psychoses, respectively.



(a) With the exception of one outlying observation in the lower left-hand region, there appears to be a significant negative relationship between cost and number of hospital discharges.

(b) By introducing the DRG codes, we see a small positive relationship between cost and number of hospital discharges within each group.

the others; it represents a combination with only two discharges. When the point is removed, the correlation becomes -0.50, thus representing an even stronger negative relationship.

Despite its simplicity, Figure 9a hides an important relationship. The right-hand panel, Figure 9b, is a redrawing of Figure 9a that includes different plotting symbols for different DRGs. Here, the story is the opposite to the one of increased economies of scale. For combinations representing major joint and limb reattachments and normal newborns, the relationship between frequency and cost is fairly flat. For these DRGs there are few economies of scale. For the psychoses DRG, number 430, Figure 9b shows a small positive relationship between frequency and cost, even discounting for the combination with only two patients discharged.

The two panels illustrate a phenomenon in statistics referred to as *Simpson's paradox*, or a problem of *aggregation of data*. See Frees (1996) for further discussion. The important point for this article is that sometimes simple graphs are misleading. Complex graphs may take more time for viewers to interpret, but they more effectively summarize complex relationships.

Guideline Four: Relate Graph Size to Information Content

"How large should the graph be?" is an important question. The bounds on size are clear. Graphs should not be so small that they are not clearly legible, particularly upon reproduction that degrades an image, nor should they be so large that they exceed a page. With large graphs, it is difficult to compare elements within the graph, thus defeating a primary purpose of graphs.

Within these bounds, a graph should be proportional to the amount of information that it contains. To discuss the proportion of information content, Tufte (1983) introduced the *data density of a graph*. This is defined to be the number of data entries per unit area of the graph. For comparing graph size and information, the data density is a quantity to be maximized, either by increasing the number of data entries or reducing the size of the graph. By examining this density over a number of popular publications, Tufte concluded that most graphs could be effectively shrunk.

For example, Figure 7a is a chart with a low data density. This chart represents only 15 numbers. With an area of approximately 9 square inches, this graph's data density is roughly 15/9. For comparison, Figure 10 shows approximately 600 numbers. Although Figure 10's area is about twice as large as that of Figure 7a, the data density is much larger in Figure 10 than in Figure 7a.

Example 3.4 Inflation Rate Forecasts

Figure 10 is a complex graph that contains much information about a complex subject, forecasting the inflation rate (CPI) for projections of Social Security funds (Frees et al. 1997). The graph shows actual experience of quarterly inflation rates up through the first quarter of 1995. Experience up through 1992 was used to fit a time series model described in Frees et al. (1997), and this model was used to generate prediction intervals (PIs) of the inflation rate. These prediction intervals can be compared to held-out experience that was not used to fit the model (1993–1995) as well as projections of inflation by Social Security experts. The thick lines represent high-, intermediate-, and low-cost inflation projections determined by Social Security experts.

Figure 10 is complex and may not be immediately understood by the viewer. However, almost every stroke within the data region represents numerical information. Although complex, Figure 10 allows the viewer to compare (1) 20 years of experience to a 10-year forecast, (2) recent held-out experience to forecasts, and (3) expert projections to forecasts generated by a time series model. The graph's complexity reflects the complexity of forecasting inflation rates; this complexity is not due to unnecessary elements that distract viewers and make them more "interested" in the graph.

Guideline Five: Use Graphical Forms That Promote Comparisons

Creators of graphs are often faced with the choice of several graphical forms that could be used to represent a feature of the data. As we describe in Guideline Eight, the receiver's knowledge of graphical forms can influence the choice. Graphical perception is also an important determinant. In Section 4, we discuss this issue in detail. We include it here as part of the Guidelines Section for completeness.

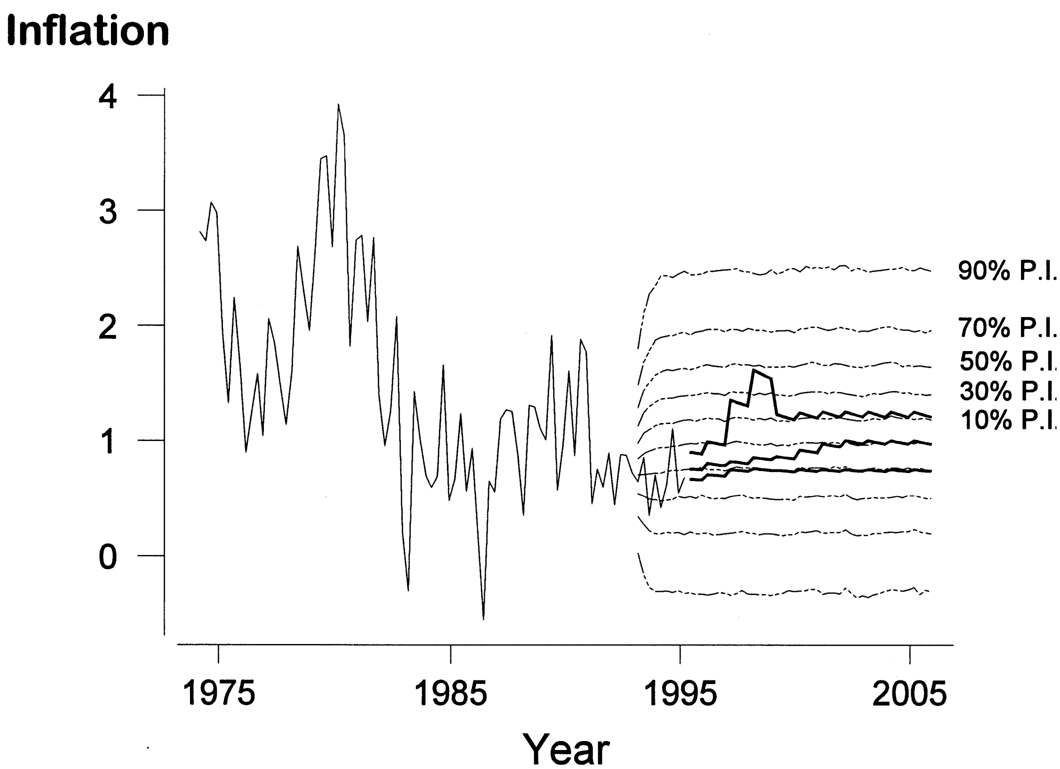
Guideline Six: Integrate Graphs and Text

Data graphics should be carefully integrated with text, tables, and other graphs. A legend summarizes the graph and its main message, but the surrounding text

Figure 10

Comparison of Stochastic Prediction Intervals to Held-out Actual Experience and to Social Security's Assumptions.

The thin solid lines represent actual inflation rates, and the thick solid lines represent projections by Social Security experts. The dotted lines represent prediction intervals generated by a stochastic time series model. This complex graph allows viewers to make comparisons based on approximately 600 points.



develops the theme leading up to the message and discusses its impact. Although "a picture is worth ten thousand words," a graph needs supporting text. Tufte (1983) encourages readers and writers to think of data graphics as paragraphs and to treat them as such.

Data graphics can be complemented by a tabular presentation of data: graphics can highlight relationships among the data, and tables can present precise numerical descriptions of the data. The two modes are complementary. A good writing device is to place a graphical display in the main body of the report and to reinforce the graph with a tabular display in an appendix.

The American Statistical Association, in its *Style Guide* for journal publications, reminds us that a detailed legend is helpful when interpreting graphs. The *Style Guide* recommends that a legend describe a graph, draw attention to the graph's important features, and explain this importance.

Guideline Seven: Demonstrate an Important Message

Detailed legends and graphs should reinforce messages that are developed in the main body of the text. To illustrate, when considering ways of portraying a complex dataset, choose a graphical form that highlights an important message. All too often, creators of graphs display data features that are not part of the theme that is being developed.

Cleveland (1994) recommends that we "put major conclusions in a graphical form." In regression data analysis, major conclusions are about patterns in the data that are summarized using models. Usually major conclusions are best presented graphically. Graphs display a large amount of information that is retained by the viewer because it is visualized. Graphs communicate patterns directly to a viewer, without using an equation to represent the patterns. In this way, a wider audience can be reached than if the presentation relies solely on a model-based interpretation of the data. Further, patterns suggested by a graph reinforce those represented by a model, and vice versa. Thus the two tools, graphs and models, reinforce and strengthen one another.

Tukey (1977) states that "The greatest value of a picture is when it forces us to notice what we never expected to see." Unexpected phenomena are usually memorable events; viewers of graphs remember these results, which makes them powerful. In writing this article, we did not expect the results of Figure 6. This figure demonstrates that unemployment rises much more quickly than it declines; it is a powerful example of the use of aspect ratios.

Guideline Eight: Know Your Audience

A basic precept of effective writing, familiarity with one's audience, is also valid for designing effective graphs. As stated in the Introduction, our primary motivation in developing guidelines is to encourage the precise and concise communication of quantitative ideas to a scientific audience using a written medium. As discussed in Section 4, the graphical form is subservient to the real role of the graphical display, *communicating* quantitative ideas of the creator to the viewer of a graph. If the audience does not have an understanding of the graphical form, then the form will hinder the communication flow rather than aid it. Thus, each of the seven guidelines already discussed can be modified or even ignored upon occasion, depending on the audience for the graph. To illustrate, in Example 3.1 we argued that the dot plot was superior to the three-dimensional stacked bar chart. As another example, in Section 4 we argue that pie charts are ineffective communicators of information based on the science of cognitive perception. However, for some audiences, creators of graphs will prefer the less effective forms based on the level of audience familiarity. We hope that practice will eventually shift from these ineffective modes of communication. Still, it is important to recognize the background of the audience of the graph. We recommend that creators of graphs not so much swim against the tide of poor graphic design as bend their course towards more effective modes of communication.

4. EMPIRICAL FOUNDATIONS FOR GUIDELINES

This section consists of two different scientific aspects of graphical studies: science of perception and surveys of graphical practice.

This article does not include a number of graphical forms that are mainstays in business publications and the popular press, such as pie charts, pictographs, and stacked bar charts. In fact, we have shown stacked bar charts in Section 3.1 only as an example of how *not* to draw figures. Why are these widely used graphical forms not adopted in an article emphasizing data graphics? The reasons lie in how graphical forms communicate information and how we perceive graphical information. We demonstrate that, given how we perceive information, pie and stacked bar charts are poor communicators of numerical information.

As described in Section 1, data graphics encode information, and we, as viewers, decode this information when viewing a graph. The efficiency of this transmission can be considered in the context of cognitive psychology, the science of perception. This discipline provides a framework for distinguishing among different types of information processing that we do when decoding graphs. Identifying different types of information processing will help us decide what are effective, and ineffective, graphical forms.

4.1 Viewers as Units of Study

Table 1 is an ordered list of basic graphical perception tasks, according to Cleveland (1994). Here, the ordering begins with a set of tasks that is least difficult for a viewer to perform and ends with a set that is most difficult. Thus, for example, judging position along a common scale is the least difficult for viewers and judging relative shadings of colors and density (the amount of ink) is the most difficult.

Table 1

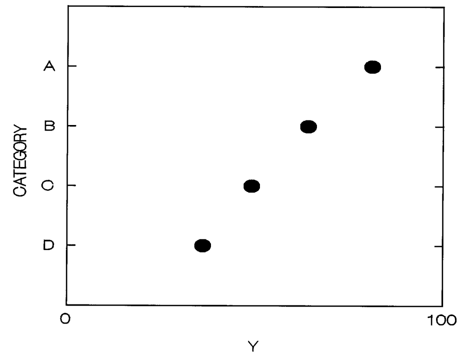
Basic Graphical Perception Tasks

1. Position along a common scale
2. Position along identical, nonaligned scales
3. Length
4. Angles and slopes
5. Area
6. Volume
7. Color and density

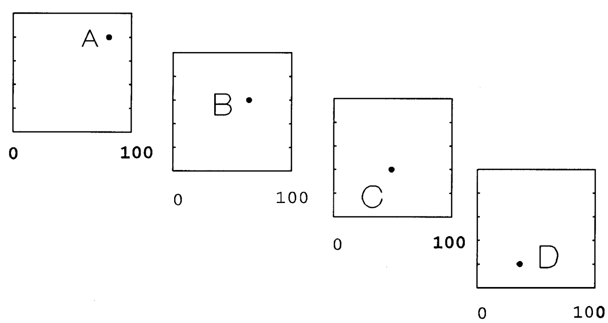
To understand the relative difficulty of the tasks, Cleveland and McGill (1984) performed a series of tests on many experimental subjects. To illustrate, Figure 11 presents a series of tests that are analogous to the first five tasks. Cleveland and McGill summarized the performance of the experimental subjects by calculating the accuracy with which the subjects performed each set of tasks. Through these measures of relative accuracy, and arguments from cognitive psychology, Cleveland and McGill developed the ordering presented in Table 1.

Figure 11

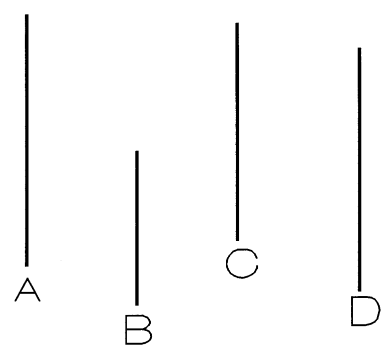
Experiments in Judgments about Graphical Perception



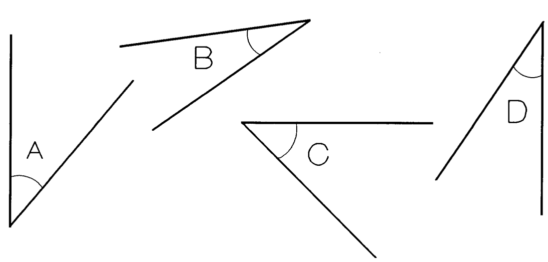
(a) Experiment to Judge Position along a Common Scale. Assess the relative values of A, B, C and D along this 100-point scale.



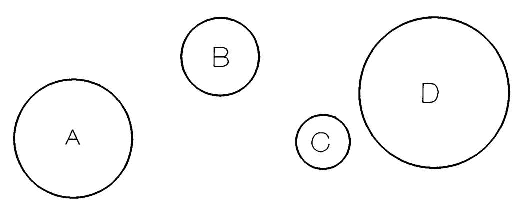
(b) Experiment to Judge Position along Identical, Nonaligned Scales. Assess the relative values of A, B, C and D on a common 100-point scale.



(c) Experiment to Understand Length Judgments. Suppose line A is 100 units long. Assess the relative lengths of lines B, C and D.



(d) Experiment to Understand Angle Judgments. Suppose angle A is 100 units. Assess the relative values of angles B, C and D.



(e) Experiment to Understand Area Judgments. Suppose circle A has area 100 units. Assess the relative areas of circles B, C and D.

This article does not discuss the use of color because of the complexities of coding and decoding it effectively. We refer interested readers to Cleveland (1994, sect. 3.13) and Tufte (1990, ch. 5) for further information.

The ordered list of graphical perception tasks can help the creator choose the appropriate graphical form to portray a dataset. When confronted with a choice of two graphical forms, a creator should select the form that is least difficult for the viewer. Other things being equal, a task that can be performed with little difficulty by the viewer means that information can be transmitted more reliably. To illustrate, we discuss two examples in which Table 1 can help you decide on the appropriate graphical form for portraying a dataset.

Example 4.1 Distribution of Premium Income

The first example demonstrates some shortcomings of the stacked bar chart. For this discussion, we return to Example 3.1. Figure 7a is a three-dimensional stacked bar chart. We have already discussed the substantial amount of chartjunk in this figure. Even without the useless pseudo third dimension, the stacked bar chart requires the viewer to make length judgments to understand, for example, the distribution of annuity receipts over time. In contrast, the dot plot in Figure 7b requires the viewer to make comparisons only according to positions along a common scale. As described in Table 1, the latter is an easier task, resulting in more reliable information for the viewer. Thus, we conclude that the dot plot is preferred to the stacked bar chart.

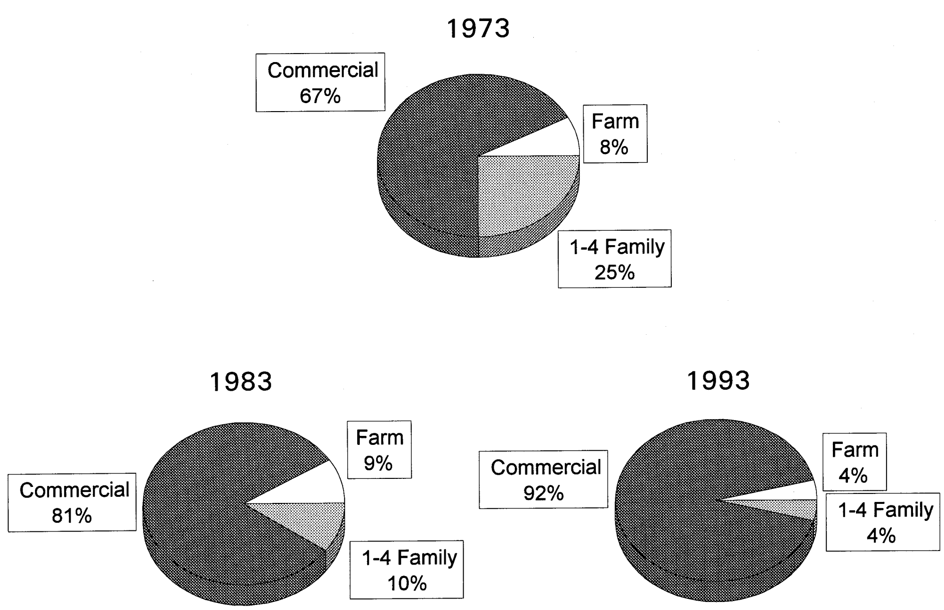
Example 4.2 Distribution of Mortgages

Our second example demonstrates the inadequacy of pie charts. Figure 12 is an adaptation of the figure on page 100 of the *Life Insurance Fact Book* (1994). It reports, for the years 1973, 1983 and 1993, commercial, 1- to 4-family, and farm mortgages as percentages of total mortgages. Pie charts make comparisons difficult. For example, the graph makes it difficult to detect whether farm mortgages are more prevalent than 1- to 4-family mortgages in 1983, or whether farm mortgage percentages increased or decreased from 1973 to 1983. The comparison of percentages across years is a linear operation, yet the pie charts require us to decode angles, a difficult task according to the ordering in Table 1. As with Example 3.1, the charts in Figure 12 make things worse by reporting in three dimensions; these figures not only require us to decode volumes but also add substantially to the chartjunk in the graphic. *Only nine numbers are reported in this graphic*, three years and two percentages

Figure 12

Distribution of Mortgages for the Years 1973, 1983 and 1993.

The three-dimensional pie chart is a poor graphical form for making comparisons over time and across types of mortgages.



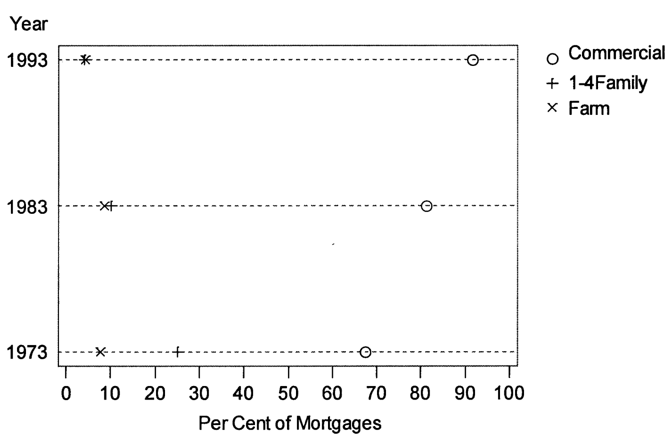
in each year. (The third percentage can be computed by subtraction.)

If a graphic is needed, then the dot plot in Figure 13 is more than sufficient. Here, comparisons are made according to positions along a common scale, a task easier than comparing angles. Pie charts require us to make comparisons using angles, which are more difficult and less reliable than comparisons using other graphical forms.

Figure 13

Commercial, 1- to 4-Family, and Farm Mortgages as Percentages of Total Mortgages for 1973, 1983 and 1993.

A negative aspect of this graph is the overlap of the 1- to 4-family and farm plotting symbols in 1983 and 1993.



Although Figure 13 is a more effective graph than Figure 12, for these data we recommend a tabular display (Table 2), which allows for clear comparisons across mortgage types and across years. Further, more detailed information about mortgage percentages is available in Table 2 than in Figure 12 or 13. Of course, we can always superimpose the actual percentages, as is often done with pie charts and as illustrated in Figure 12. Our response to this approach is to question the worth of the entire graph. As with writing, each stroke should offer new information; let creators of graphs make each stroke tell!

4.2. Graphs as Units of Study

Surveys of graphical practice in professional publications provide an important database with which to assess prevalence of good and bad practice and changes in practice over time. Tufte (1983, pp. 82–86) discusses a survey of approximately 4,000 graphs

Table 2

Commercial, 1- to 4-Family, and Farm Mortgages as Percentages of Total Mortgages for 1973, 1983, 1993

	Year		
Mortgage Type	**1973**	**1983**	**1993**
Commercial	67.5	81.3	91.7
1–4 Family	24.9	10.1	4.1
Farm	7.6	8.6	4.2

randomly selected from 15 news publications for the years 1974 to 1980. The graphs were assessed for "sophistication," defined as presentation of relationship between variables, excluding time series or maps. Cleveland and McGill (1985) report a similar survey of scientific publications, assessing the prevalence of graphical errors.

Harbert (1995) assessed every graph and table in the 1993 issues of four psychology journals on 34 measures of quality. The measures of quality were gleaned from the current research literature on graphic quality. They were converted into a check sheet, and a check sheet was filled out for each graph and table in the selected psychology journals. Harbert's study yielded data on 439 graphs and tables. We summarize the analysis of the 212 graphs.

Harbert assigned letter grades to the graphics: A, AB, B, BC, C, CD, D, DF and F. These grades reflected her overall evaluation of the graphs as communicators of statistical information. The grades were converted to numerical values: 4.0, 3.5, 3.0, 2.5, 2.0, 1.5, 1.0, 0.5 and 0.0. The numerical values were the dependent variable in a regression. The independent variables were the 34 measures of quality, suitably coded. The purpose of the study was to determine which factors were statistically significant predictors of the grades assigned by an "expert" evaluator of graphics. By trial and error, Harbert selected a multiple linear regression equation in which all the predictors were statistically significant (5% level) and no other predictors achieved this level of significance when added to the equation. Table 3 shows the variables included in the regression equation ($R^2 = 0.612$).

Data-ink ratio was defined by Tufte (1983, p. 93) as the "proportion of the graphics ink devoted to the nonredundant display of data-information" or equivalently as "1.0 − proportion of a graphic that can be erased without loss of data-information." The data-ink

Table 3
Factors Affecting Assessment
of Graphic Quality, Harbert Study

Variables with Positive Coefficient	Variables with Negative Coefficient
Data-ink ratio Comparisons made easy Sufficient data to make a rich graphic	Proportion of page used by graphic Vertical labels on Y-axis Abbreviations used Optical art used Comparisons using areas or volumes

ratio is more readily calculated than the data density measure defined in Section 3 of this paper. Optical art is decoration that does not tell the viewer anything new.

One variable that had been anticipated as very significant was data density, which is difficult and time-consuming to measure. An important finding of the study was that the easier-to-measure data-ink ratio and proportion of page variables were sufficient to predict the grades. A quotation from Harbert's thesis sums up the finding: "The highest grades were given to those graphics that take up small proportions of the page, have a large data-ink ratio, make comparisons easy, have enough data points, have horizontally printed labels, do not have abbreviations, do not have optical art, and do not use volume or 3-D comparisons" (Harbert 1995, p. 56).

As a small follow-up study to Harbert's work, we examined each of the 19 non-table graphics in the *Life Insurance Fact Book* (1994), assessing them on seven negative factors. Table 4 shows the percentage of graphs that displayed each of the negative factors.

Table 4
Percentage of Graphs Displaying Negative
Factors in ***Life Insurance Fact Book 1994***

Negative Factor	Percentage of Graphics
Use of 3-D bars	79%
Grid lines too dense	79
Making comparison of time series values hard	37
Use of stacked bars	37
Growth displayed poorly	32
Use of lines that are wider than need be	16
Use of pies	5

Our review suggests that every graphic could have been reduced by 50% to 75% without loss of clarity. This observation is in keeping with Harbert's finding about the proportion-of-page variable. In a word, the graphs in the *Life Insurance Fact Book* could be produced much more ably. Doing so would improve the quality of communication and would potentially increase the respect with which knowledgeable professionals in other fields view the insurance industry.

We hope that other investigators will engage in further study of graphic practice in actuarial publications. By using data from such studies, the profession can improve its practice, making communications efficient and precise.

5. CONCLUDING REMARKS

The Society of Actuaries motto is a quotation of Ruskin: "The work of science is to substitute facts for appearances and demonstrations for impressions." Armed with the guidelines outlined in this paper and discussed further in the references, actuaries can be leaders in presenting data graphically, thus substituting demonstrations for impressions. Surveys of recent actuarial literature should be the basis for assessing current practice. Editors and referees of professional publications can be especially influential in bringing about a rapid improvement in standards of practice. Moreover, actuaries can recommend and use statistics textbooks that pay attention to graphic quality, such as Cryer and Miller (1994) and Frees (1996).

Because actuaries read material that contains graphs, they are consumers. They should become tough customers! All too often the defaults in spreadsheet and statistical graphics software become the norm. Actuaries should not allow the choices made by software programmers to drive graphic quality or standards. Although it is easy to create graphs using defaults in the graphics software, the resulting graphs are seldom fully satisfactory. If a graph is not worth doing well, let's leave it out of our publications.

In addition to the references listed, other resources are available to actuaries interested in improving their graphic design skills. Like the Society of Actuaries, another professional organization, the American Statistical Association (ASA), has special interest sections. In particular, the ASA now has a section on statistical graphics. Interested actuaries can join ASA and that section to get the newsletter *Statistical Computing & Graphics*. This publication has examples of excellent graphical practice in the context of scientific discovery and application. Membership information is available on the Internet: asainfo@asa.mhs.compuserve.com.

The technical *Journal of Computational and Graphical Statistics* contains more in-depth information on effective graphs. We also recommend accessing and using the *ASA Style Guide* at http://www.amstat.org/publications/style-guide.html as an aid to effective communication of quantitative ideas.

REFERENCES

AMERICAN COUNCIL OF LIFE INSURANCE. Various years. *Life Insurance Fact Book*. Washington, D.C.: ACLI.

CLEVELAND, W.S. 1994. *The Elements of Graphing Data*. Monterey, Calif.: Wadsworth.

CLEVELAND, W.S. 1993. *Visualizing Data*. Summit, N.J.: Hobart Press.

CLEVELAND, W.S., DIACONIS, P., AND MCGILL. R. 1982. "Variables on Scatterplots Look More Highly Correlated When the Scales Are Increased," *Science* 216:1138–41.

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* 79:531–54.

CLEVELAND, W.S., AND MCGILL, R. 1985. "Graphical Perception and Graphical Methods for Analyzing and Presenting Scientific Data," *Science* 229:828–33.

CRYER, J.D., AND MILLER, R.B. 1995. *Statistics for Business: Data Analysis and Modeling*, 2nd ed. Belmont, Calif.: Duxbury.

EHRENBERG, A.S.C. 1977. "Rudiments of Numeracy," *Journal of the Royal Statistical Society A* 140:277–97.

FREES, E.W. 1996. *Data Analysis Using Regression Models*. Englewood Cliffs, N.J.: Prentice Hall.

FREES, E.W. 1998. "Relative Importance of Risk Sources in Insurance Systems," *NAAJ* 2, no. 2 (April):34–51.

FREES, E.W., Kung, Y.-C., Rosenberg, M.A., Young, V.R., and Lai, S.-W. 1997. "Forecasting Social Security Assumptions," *NAAJ* 1, no. 3 (October):49–82.

HARBERT, D. 1995. "The Quality of Graphics in 1993 Psychology Journals," Senior honors thesis, University of Wisconsin-Madison.

HUFF, D. 1954. *How To Lie with Statistics*. New York: Norton.

SCHMID, C.F. 1992. *Statistical Graphics: Design Principles and Practices* Malabar, Fla.: Krieger Publishing Co.

SCHMIT, J., AND ROTH, K. 1990. "Cost Effectiveness of Risk Management Practices," *Journal of Risk and Insurance* 57: 455–70.

STRUNK, W., AND WHITE, E.B. 1979. *The Elements of Style*. 3rd ed. New York: Macmillan.

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

TUFTE, E.R. 1990. *Envisioning Information*. Cheshire, Conn.: Graphics Press.

TUFTE, E.R. 1997. *Visual Explanations*. Cheshire, Conn.: Graphics Press.

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

UNIVERSITY OF CHICAGO PRESS. 1993. *The Chicago Manual of Style*. 14th ed. Chicago, Ill.

DISCUSSIONS

WILLIAM C. CUTLIP*

Congratulations to Dr. Frees and Dr. Miller. They have done a wonderful job of sharing some communication insights in an area in which we often take too much for granted.

My career practice has been in the business arena. With that in mind, I have several observations I would like to offer.

First, the eight guidelines are very well stated; however, I would suggest that emphasis should be placed on Guideline Seven, "Demonstrate an Important Message," and on Guideline Eight, "Know Your Audience." In fact, I would rank these last two guidelines as the first two.

When developing any presentation, make sure that you have something to say. Know what your subject is, understand the focus of it, and cut away the "presentation junk" so that the hearer is led directly to the point you are trying to make.

Knowing your audience is very important in any communication. You will use different vocabulary, different charts, and different points to help your audience hear what it is you are trying to say to them. A scientific audience can deal with more technical and complex graphics than can a nontechnical audience. Your audience may include scholars, other actuaries, MBA's, business leaders, salespeople, business workers, media, or others. Understanding your audience and their receptivity will help lead to the appropriate graphics.

Second, the dot plot chart favored by the authors was one that was new to me in graphic presentations. I can see its usefulness in trying to present a complex collection of information. However, in several circumstances I found it less enlightening than a bar chart. For example, in Figure 7a a simple (not stacked) bar chart with three areas arranged by type of insurance would, for me, more quickly and clearly show the dramatic changes in premium receipts from year to year.

The dot plot chart may be more familiar to those in the academic and scientific communities than to those of us in the business world. This again

*William C. Cutlip, F.S.A., is President of William C. Cutlip Consulting, 3531 Sabaka Trail, Verona, Wisconsin 53593, e-mail, wcutlip@compuserve.com.

emphasizes the fact that it is important to know the audience and the easiest means of receptivity for them.

Third, though it was not mentioned in the paper, one thing to consider is the value of using alternative forms of charts, particularly in a long oral presentation. This will help keep the audience's attention and will also help the audience focus on the presenter's points. This is especially true with an audience of mixed backgrounds and chart receptivity.

Fourth, in the authors' Guideline Eight, they state that the primary motivation was to "encourage the precise and concise communication of quantitative ideas to a scientific audience." It is also helpful to encourage actuaries to study, understand, and use different forms of graphics to meet the needs of the business and public audiences with whom they will come in contact. Let's face it, most people have been nurtured on the graphics of Ross Perot and *USA Today*.

Fifth, care should be taken when using complex graphs. (For example, see the authors' Figure 10.) It may take more thought and time to encode and decode the message than is valuable. Before a complex graph is undertaken, keep in mind the paragraph on vigorous writing that was quoted from *The Elements of Style*.

A final hurrah to the authors for tackling a subject to improve the quality of communication!

DOUGLAS A. ECKLEY*

Dr. Frees and Dr. Miller have made a valuable contribution to actuarial literature. We all use graphs.

I admire the authors for being bold in stating their conclusions. Their criticism of the pie chart was thought-provoking, since that is a commonly used type of graph.

In Section 3 the authors provide eight guidelines for graphical design; these are highly useful. The last of these, "know your audience," to me is the most important. In Example 4.2 the authors suggest that a table is better than a dot plot, but both are better than a pie chart. I believe the authors are assuming that the audience is technically oriented. Pie charts would be useful to an actuary communicating with unsophisticated pension plan participants about investment mix. The conceptual simplicity would more than offset any difficulty in visually estimating pie slice sizes.

I propose a Guideline Nine: "label the graph fully and clearly." For example, if the data are in dollars, show the dollar signs. This is as important as knowing your audience. If the user is not sure what is being graphed, the message has surely been lost. I must admit that the authors confused me in Example 2.3. What exactly are the numbers on the vertical axes in the two graphs?

In Section 2 the authors illustrate some crucial choices that the designer must make. The examples in this section are excellent. In the first three examples the authors point out ramifications of different choices without recommending a particular choice. In Example 2.4 they submit that scale ranges should be the same on both y-axes of a double y-axis graph. I respond that:

- Double y-axis graphs are difficult to work with in any event
- If the authors are correct that scale ranges should be the same, then double y-axis graphs cease to exist, because when the two axes are the same, there is only one axis
- Logarithmic double y-axes are sometimes useful.

Actuaries have no choice but to use tables in many cases. For example, a graph of annuity rates offered by 50 companies would not be useful to a potential buyer. And the authors recommend a table over a graph in Example 4.2. When a table is used to display an allocation, I prefer to see totals at the bottom. In Table 2, I wondered if the percentages in a year add up to 100%, or if some smaller mortgage classes were excluded.

Since designing tables is less creative than designing graphs, we may not need a paper entitled "Designing Effective Tables." If we do, we would be fortunate to have Mr. Frees and Mr. Miller write it.

Designing graphs is a creative exercise. The authors are wise to guide rather than admonish; the guidelines leave ample room for creativity. I believe that, once the guidelines are followed, graphs are like case studies: there is no one right answer.

GARY S. LANGE*

The authors, Dr. Frees and Dr. Miller, have provided an excellent summary of the topic, communication with graphs. This vital information has been missing from every list of actuarial educational material

*Douglas A. Eckley, F.S.A., is a doctoral candidate in economics at George Mason University and a practicing actuary; e-mail, smvk23a@prodigy.com.

*Gary S. Lange, F.S.A., is Associate Actuary, CNA Life Reinsurance, CNA Plaza 35 S, Chicago, Illinois 60685, e-mail, g.lange@cna.com.

currently in use, at a time during the evolution of the actuary in which the ability to communicate is considered one of the most important tools. The actuary has to communicate complex ideas and information to a world not quite so mathematically oriented. Much of this information can be presented concisely with graphs. Yet after reading this paper and working with attempts at communication via graphs, I realize many of us have achieved only beginner levels of competency in the design of effective graphs.

As chairperson of the *Record* Editorial Board for the past two years I have been exposed to all types of graphs used by actuaries in their presentations at the Society of Actuaries annual and spring meetings. Even before reading this delightful paper, I realized that many of the graphs from these presentations were very difficult to understand. I assumed the problem was that I reviewed them without the benefit of the live presentation. Now I realize that the problem is graph abuse. Any actuary considering using a graph in a presentation at a Society meeting should read this paper before designing that graph!

The major source of the problem is easy to see, and the authors point out that ". . . it is easy to create graphs using defaults in the graphics software." Click on a few icons and suddenly your list of numbers is a beautiful graph. The beauty of the picture can be so dazzling that the goal of communication is overlooked. The software is just too easy to use, but rather than discarding it, read "Designing Effective Graphs." The suggestions and examples will help you use the software properly to improve your efforts at communication.

The authors of this paper challenged me, as an editor, to bring about ". . . rapid improvement in standards of practice." The best way for me to meet their challenge is to entice readers to use this paper to improve their skills in communicating with graphs. Terms used by the authors are easy to remember and will help all of us improve our communication skills. How can one ignore "chartjunk" or ever use that "pseudo third dimension" again? The next time you use a graph, is it a form of communication or is it "optical art?" Finally, all of us are reminded to do a self-evaluation with the statement, "If a graph is not worth doing well, let's leave it out of our publications."

Actuaries need this paper. I hope that this paper will be included in the information sent to speakers before each actuarial meeting. I would like to thank the authors for providing the resources to help us improve our communication skills. Read the paper now—don't wait for the movie!

EDWARD M. MAILANDER*

The authors are to be commended for adding some very practical information to the body of actuarial knowledge. Actuaries would be well advised to follow the authors' guidelines and recommendations in using graphs to communicate technical information. It is not sufficient to solve a technical problem or analyze data; the results of the work must be communicated to others. This paper points the way to improving that communication.

Most of what the authors write in their paper also could be applied to tabular information. For example, there is the tabular equivalent of chartjunk (tablejunk?), such as unnecessary shading and shadowing and an excess of different typefaces.

Guideline Two about using multiples could be expanded to include a comment about the importance of consistent formats. This consistency would be both within an article or presentation and consistency with standard practices (such as displaying the independent variable on the x-axis). An example is the display of years. I believe that the most common way to display time is to show it increasing to the right. (Perhaps I have been conditioned as a student by hours spent with timelines.) I don't like the format for Figure 13. My dislike could come from the years being in a place where I am not used to seeing them (that is, the y-axis) or because time is the independent variable and it is not on the x-axis.

In addition to improving the effectiveness of graphs when they are used, I believe that communication could be improved by using graphs in situations in which they have an advantage over other methods of communicating data. Such areas include using graphs to show the distribution of data and trends over time and to display raw data.

The authors made a brief comment about the use of graphs to display raw data. I would like to see more raw data included in technical articles so that the reader is in a better position to evaluate the conclusions of the author. Graphic inclusion of the raw data would address this without adding significantly to the length of most articles.

*Edward M. Mailander, F.S.A., is a Senior Actuarial Consultant, Ernst & Young LLP, 370 17th Street, Denver, Colorado 80202.

ALEXANDER J. MCNEIL*

As someone who spends a good deal of time creating graphs to present statistical results, I was interested to read the authors' opinions on this subject. I agree largely with their design guidelines and believe they are to be welcomed. As is often the case with style guides, I might be tempted to say that many of the guidelines are common sense, if it were not for the frequency with which the proposed rules are violated and bad graphs are created and published.

The admonition to "avoid chartjunk" (Guideline One) is a useful basic principle. Graphics packages have definitely led to a tendency to create irrelevant graphs as well as to load relevant graphs with irrelevant information. The idea of appropriate complexity of presentation to the complexity of the problem in hand (Guideline Three) is also worth stressing in this context. Slightly less obvious to me was the idea that the physical size of the graph should be proportional to its information content (Guideline Four), but this is something I have now taken on board. It is interesting that the authors can support many of their guidelines by appeals to findings in the science of perception.

Because the focus of the article is on graphs in professional writing and not only in scientific papers, perhaps I may also mention in passing my general dislike for the modern overuse of pseudo-scientific charts, by which I mean diagrams utilizing bubbles, arrows, proliferating axes, and the full armory offered by graphical packages to express the interrelation of concepts. I seldom find such diagrams helpful, and often they are in violation of the chartjunk guideline. In contrast, an example of an effective diagram is the simple flow chart of Figure 1, which is unobtrusive in size and repays a few moments of reflection.

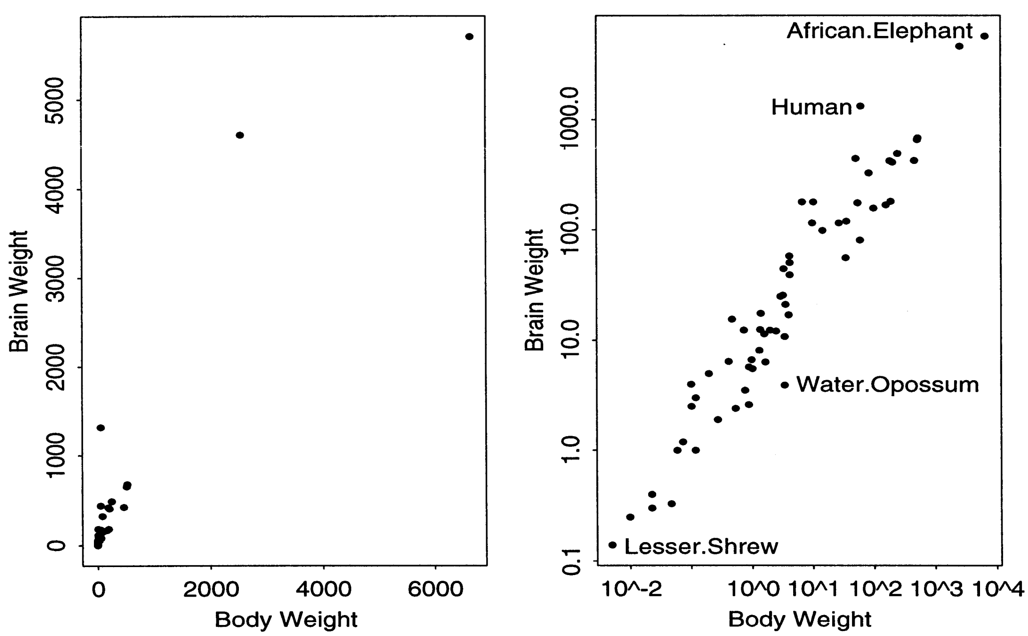
Some specific points in the paper that led me to think more carefully were the issues of scales and types of axes (in relation to the examples of Section 2). In my work on modeling large insurance losses, I have been quite a heavy user of logarithmic axes. I believe the message of my graphs in such applications can barely be seen on a linear scale. However, my feeling about logarithmic axes is that they should be avoided unless absolutely necessary. Logarithmic

*Alexander J. McNeil, Ph.D., is Swiss Re Research Fellow, Department of Mathematics, Swiss Federal Institute of Technology, ETH Zentrum, CH-8092 Zurich, Switzerland, e-mail, mcneil@math.ethz.ch.

Figure 1

Relationship between Body and Brain Weights (Differing Units) for Various Mammals.

Left picture has linear x- and y-axes; right picture has logarithmic x- and y-axes. Identities of four mammals revealed.



axes are terrible for interpolation and extrapolation; they lead frequently to confusion because people seldom think on a logarithmic scale. Only when the true nature of the problem is not visible on the linear scale does the log scale seem useful, for example, when the range of the data is extreme or when a power relationship explains the data.

An example comes from my statistics teaching when I give my students a practical example of linear regression. When the log of average brain weight is plotted against the log of average body weight for a variety of different mammals, a striking linear relationship is observed (with a few conspicuous outliers including the human being). On the linear scale nothing is observed, owing to the distorting effects of having mice and elephants in the same plot; the log-log plot reveals all and allows us to posit a power relationship ($y = ax^b$) between body weight and brain weight in mammals—a nice model with admittedly limited scope for actuarial application.

I also have reservations about double y-axes. This is another technique that leads most frequently to confusion and can be justified only when the second y-axis is making some additional important point, so that the complexity of the ideas to be expressed justifies the complexity of the presentation (the authors' Guideline Three).

In conclusion, I thank the authors for asking us to think about graphical presentation. For those of us who use graphs, it is a useful exercise to calibrate our own practice against the authors' sensible guidelines.

ARNOLD F. SHAPIRO* AND EDWARD B. KLEINMAN†

This paper by Dr. Frees and Dr. Miller deals with the important issue of communicating quantitative ideas graphically, an area where we have often seen abuse. Their overview of the potential problems and guidelines undoubtedly will be included on the suggested reading list in the "Presenter Kit" of many future conferences.

*Arnold F. Shapiro, F.S.A., M.S.P.A., Ph.D., is Professor of Actuarial Science and Insurance, Robert G. Schwartz Faculty Fellow and Director of the Risk Management Research Center, Smeal College of Business, Pennsylvania State University, 409 F Business Administration Building, University Park, Pennsylvania 16802, e-mail, afs1@psu.edu.

†Edward W. Kleinman, Ph.D., is an instructional design specialist, Smeal College of Business, Pennsylvania State University, 409 Business Administration Building, University Park, Pennsylvania 16802, e-mail, ebk@afshapiro.com.

The following discussion addresses two key issues in efforts to avoid complexity, ambiguity, and abuse in visuals that accompany presentations:

- The different levels of intellectual skills that are required at various learning levels
- The design of visuals for consistency and congruency.

Even the most erudite presentations have the potential to be misleading and confusing despite the best intentions of the creators. One key explanation for that potential is that many different levels of learning exist at which comprehension is desired, not just one (Gagne 1985, pp. 132–36; Gagne 1993, pp. 59–61; Dwyer 1978, p. 70; Kleinman 1997, pp. 88–91). Furthermore, achievement or understanding rarely is measured at a specified learning level. Without such specifications, it is difficult to evaluate the effects of visuals (Dwyer 1978, pp. 39–41). Achievement at those different levels of learning can be influenced by the visuals that attend the presentation of information (Dwyer 1994, pp. 398–99; Kleinman 1997, p. 117).

In general, the intellectual skills that operate at the different levels of learning can be subdivided into the following four groups (Gagne, 1985, p. 129; Driscoll and Gagne 1988, p. 85; Kleinman 1997, pp. 6–7):

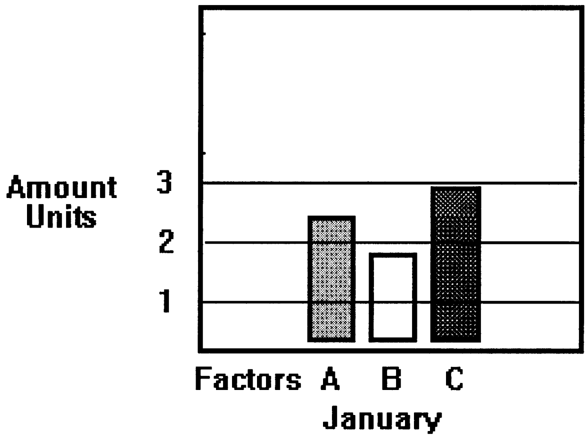
1. Facts: single units or pieces of information that have been learned
2. Concepts: two or more facts that are related
3. Procedure or rule: steps or rules that are followed to aid in the solution of a problem
4. Problem-solving: the capacity to explore for answers to unresolved questions and apply procedures, concepts, or facts.

Figure 1 illustrates a simple vertical bar graph that focuses on the presentation of factual information. Pictorially it demonstrates that there are three factors, A, B and C. In January, it is apparent how many units there are for each factor and how the factors are related. Those facts are elicited easily from the visual representation.

What cannot be elicited easily from the figure is the precise quantity of variation among the factors. The choice of units for a figure affects the ability to interpret the meaning. That is a point that Dr. Frees and Dr. Miller (Section 1.1 and Section 2) make quite convincingly. Consequently, if Figure 1 is used to try to illustrate the *concept* that Factor B lags so far behind Factors A and C that some action is required, then the figure might be less helpful because it was designed and presented to exemplify one or more *facts*, without a satisfactory structure for *concept* presentations.

Figure 1

The Units of Factors A, B and C for January



We differ a bit with Dr. Frees and Dr. Miller's assertion of ". . . the complexities of coding and decoding . . ." color effectively (Section 4.1). Indeed, while research into specific visuals or visual skills associated with the different levels of achievement has been limited (Dwyer 1994, p. 386; Kleinman 1997, pp. 59–61), studies have confirmed that color can enhance achievement primarily at the concept level of learning (Dwyer 1987, p. 77; Kleinman 1997, pp. 113–14).

For example, if each of the factors in Figure 1 were in color, especially if factors A and B were in different colors, then the discrepancies among the factors would be exacerbated. If the color red were identified as a cautionary hue and if factor B were colored red, then color as an instructional variable could be used efficiently and effectively to call attention to the less satisfactory performance for factor B. This would be a conceptual event in that two or more elements are linked by a relationship.

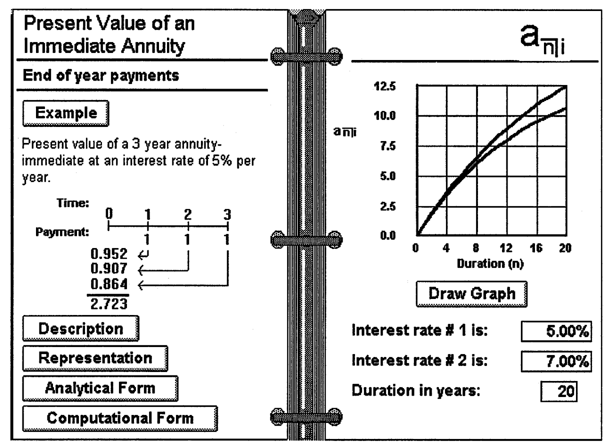
At procedural and problem-solving levels, visuals should become more complex in design. Consider Figure 2, taken from a computer program that was designed specifically for participants who are interested in the *Mathematics of Finance* (Shapiro 1994). In this case, there are provisions for interactivity between the student and the exercises on that page.

As indicated, the student can provide different values for variables and thus generate different solutions for the equations that are being learned. This process exemplifies learning at the higher levels of rule and problem-solving. Since those levels are more complex by necessity than the levels of factual and conceptual knowledge acquisition, the visuals that accompany them also are more complex.

Consistency and congruency imply that if specific content is presented at the fact level, for example, then the visuals that accompany the fact content should reflect only the fact or facts that are presented. They should not depict concepts, procedures, or problem-solving techniques.

Figure 2

Mathematics of Finance Interactivity



If consistency and congruency are observed at each level of learning in the construction and utilization of visuals, then understanding and knowledge acquisition can be facilitated by those visuals. The discrepancies mentioned by Dr. Frees and Dr. Miller might then be resolved so that the presenter and the participant are better able to communicate.

REFERENCES

DRISCOLL, M., AND GAGNE, R. 1988. *Essentials of Learning for Instruction*. Englewood Cliffs, N.J.: Prentice Hall, Inc.

DWYER, F. M. 1978. *Strategies for Improving Visual Learning*. State College, Penn.: Learning Services.

DWYER, F. M., ed. 1987. *Enhancing Visualized Instruction: Recommendations for Practitioners*. State College, Penn.: Learning Services.

DWYER, F. M. 1994. "One Dimension of Visual Research: A Paradigm and Its Implementation," in *Visual Literacy: A Spectrum of Visual Learning*, edited by D. M. Moore and F. M. Dwyer. Englewood Cliffs, N.J.: Educational Technology Publications, pp. 383–401.

GAGNE, E., YEKOVICH, C., AND YEKOVICH, F. 1993. *Cognitive Psychology of School Learning*, 2nd. ed. New York, N.Y.: Harper Collins College Publications.

GAGNE, R. 1985. *The Conditions of Learning and Theory of Instruction*, 4th. ed. New York, N.Y.: Holt, Rhinehart and Winston.

KLEINMAN, E. 1997. "An Investigation of the Relationships among Selected Visual Skills and Academic Achievement at Four Different Levels of Learning," Doctoral dissertation, The Pennsylvania State University.

KLEINMAN, E., MOORE, M., AND DWYER, F. 1997. "Effects of the Use of Color in Technology," Working paper, The Pennsylvania State University.

SHAPIRO, A. 1994. *AFS Mathematics of Finance*. State College, Penn.: Arnold F. Shapiro & Associates.

AUTHORS' REPLY

EDWARD W. FREES AND ROBERT B. MILLER

In contrast with many academic articles, this paper makes bold statements to challenge readers to consider their own graphical practices. We hoped that these bold statements would encourage a lively discussion of the paper. Discussions serve to position a paper; they allow readers to appreciate what a paper does and does not do. We were not disappointed; the discussions enhance the value of the paper, and we thank the discussants for their excellent contributions. In lieu of a point-by-point response, we summarize the main points of the six discussions.

Two of the discussants, Cutlip and Lange, emphasize the role of graphs that appear in business outlets as compared to graphs that appear in scientific publications. Although graphical design guidelines are common to both media, the principle "Know your audience" comes to the forefront in business outlets. The importance of this guideline was underscored by Eckley, who considered it the most important of our eight design guidelines. Shapiro/Kleinman also reinforce this point and provide many references from the educational literature on communication targeted at specified learning levels.

Lange and Mailander reinforce the message that through more effective graphical practices, actuaries can become better consumers and producers of graphs. Lange cites the seductiveness of graphical software. Further, with improved communication skills, we should be able to present more information effectively. In particular, Mailander suggests that actuaries present more basic raw data, which are input to an analysis, in addition to conclusions that arise from an analysis.

The discussions by Cutlip and Shapiro/Kleinman emphasize the distinction between graphs that appear in a static medium, such as a professional or scientific publication, and those that appear in a dynamic medium, such as through a live presentation or interactive software package. Many of the guidelines are common to the two forums, yet the relative emphasis on the importance of the guidelines differs. Our paper focuses on the static medium, and the discussions help to clarify this distinction.

MacNeil emphasizes the difficulties encountered when interpreting logarithmic scales, a favorite way to "lie" with statistics. Like MacNeil, in our own teaching we find that an appreciation for logarithmic scaling in plots must be cultivated; it is easy to confuse or deceive the uninitiated. Nonetheless, multiplicative relationships are common in actuarial practice, such as present values in interest theory or population growth in mortality studies. Logarithmic scaling is an important device for communicating these relationships graphically.

Both Mailander and Eckley comment that this, or a follow-up, paper could include tabular presentation of data. As we note in the paper, individual actuaries can improve their skills by accessing references such as Ehrenberg (1977) and Tufte (1983). We find that the rules that guide professional editors, such as those described in the *Chicago Manual of Style* (1993), and the rules for designing effective tables are seldom in conflict. Only occasionally are there conflicts between these two sets of rules. On the whole, therefore, we believe that there are greater opportunities for improving graphical, rather than tabular, practices.

We were pleased to see the Shapiro/Kleinman comments on the use of color in graphs. We do not view color as something to avoid in graphs; colors can be used in exciting and creative ways to display important patterns in data. With evolving technologies, we anticipate greater use of color in static media than in the past. Despite the many opportunities for excellence in graphical displays using color, we still see many opportunities for deception using this feature. To keep our article focused, we chose not to present guidelines on the use of color. The remarks by Shapiro/Kleinman are welcome additions to the paper.

Again, we thank the discussants for their remarks. Communicating quantitative information graphically is important to actuaries; we are pleased that the discussants share this viewpoint. We hope that this paper encourages readers to adopt a critical eye when they view the "graphics of Ross Perot and *USA Today*." More importantly, we hope this paper will help actuaries design and use effective graphs.

Additional discussions on this paper can be submitted until October 1, 1998. The author reserves the right to reply to any discussion. Please refer to the Submission Guidelines for Authors for instructions on the submission of discussions.