Graphs in Print

Jeff Zacks[1], Ellen Levy[1], Barbara Tversky[1], and Diane Schiano[2]

Stanford University[1] and Interval Research Corporation[2]

Please address communications to:

Jeff Zacks

Washington University

Psychology Department

St. Louis, MO 63130-4899

314-935-8454

jzacks@artsci.wustl.edu

Abstract

Diagrams for presenting quantitative data are an important component of print communication. Their rate of use is high and rising. This reflects in part the recent development of software tools for generating data graphics. These programs allow a wide range of choices for data visualization—some of which may be ugly or ineffective. How has graph usage evolved during this period? A survey of graph using in academic journals, magazines, and newspapers during the years 1985-1994 revealed several dynamic trends in the characteristics of data graphics, as well as robust differences between media. However, graph features that have been singled out by experts as poor choices, such as "3-D" rendering, do not seem to be on the rise.

Introduction

Graphical presentations of quantitative information are a striking and ubiquitous component of print media. A well-chosen graphic can effectively communicate a large amount of information efficiently (Larkin & Simon, 1987), and can make that information perceptually salient and memorable. Many popular news publications, and most scientific journals, make extensive use of data graphics. Recently, three developments in the study and technique of graphing have impacted on our understanding of how graphics function.

In the past two decades a cottage industry has developed around critiques of poor design in statistical graphics. Jibes at inefficient, misleading, or inelegant graphs come from designers (Tufte, 1983), statisticians (Cleveland, 1984; Wainer, 1984), and perceptual psychologists (Kosslyn, 1993). In all these cases, criticism is accompanied by constructive insight into how to produce graphics that are efficient, perspicuous, and elegant.

At the same time, the behavioral study of how people extract quantitative information from visual displays has begun to blossom. This tradition has its roots as far back as at least the 1930s (see Macdonald-Ross, 1977 for an excellent review), but since the early 1980s there has been a flowering of both empirical research (e.g., Carswell, 1992; Casali & Gaylin, 1988; Cleveland, Harris, & McGill, 1983; Gattis & Holyoak, 1996; Shah & Carpenter, 1995; Simkin & Hastie, 1987; Spence, 1990; Tversky & Schiano, 1989) and theoretical analysis (Bertin, 1980; Cleveland, 1985; Lohse, Walker, & Rueler, 1994; Pinker, 1990).

As critical analysis and behavioral research have expanded, the technologies employed by the producers of graphs have metamorphosed. At the time Bertin's[1] (1983), Cleveland's (1985), and Tufte's (1983) major works appeared, computer technologies for creating quantitative graphics were the exclusive province of major media design shops and well-funded scientists, many in private industry. Since then, the explosion of computing power and the development of the graphical user interface have put sophisticated tools for producing statistical graphics in the hands of everyday producers and consumers of data.

These three developments have contributed substantially to what we know about how data graphics work, and have provided a wealth of new abilities to produce ever more sophisticated figures. They also raise a question: How are graphs actually being used in contemporary communications? This question is relevant to the critical enterprise, because there is little interest in critiquing that which does not occur; also, there is reason to worry that the new technologies have made it not simply easier to produce graphs, but particularly easier to produce bad ones. It is relevant to the behavioral study of graphs, because empirical and theoretical understanding is most valuable for those graphical formats that are frequently used. Finally, it is important for producers of tools for creating visualizations, because it is of interest to know what kinds of graphs are of interest to different communities, and it is worthwhile to identify areas that are poorly served by current technologies.

The research described here focuses on the use of graphs in print media. This is a question that has received limited attention in previous work.

---

[1] Bertin's book was published in English in 1983. It originally appeared in French, in 1967

Cleveland (1984) surveyed the use of graphs in scientific publications. He classified graphics as depicting one, two, or three variables. Most graphs (83%) showed two variables. There was marked variability in the amount of space per page devoted to graphs, ranging from 31% to 0%, depending on the publication. Cleveland also examined graphs from a sample of Science magazine in detail, and was dismayed to find that 30% had errors of construction, reproduction quality, explanation, or discriminability of items. Tufte (1983) surveyed 15 news magazines and newspapers from around the world, and coded their graphs as relational (linking two or more variables) or not. Relational graphs were taken to be more graphically sophisticated than non-relational ones. No publication in Tufte's sample contained more than 10% relational graphs, and no publication from the United States even reached 1%.

The purposes of the current project were threefold: first, to characterize the types of graphs typically used in current print publications; second, to compare different media in their use of data graphics; third, to examine dynamic trends in the features of graphs—especially those that have been facilitated by new computing technologies and / or those that have been targeted by critics of graphical practices as bad form. These issues were addressed by sampling a range of print media during a ten-year period, from 1985 to 1994, and coding a large number of graphs on a large number of features.

Method

(Bertin, 1967).

We began by selecting 17 publications as the database to be studied. To allow comparisons between different publication media, we selected 7 academic journals, 6 magazines, and 4 daily newspapers. The journals were the Journal of the American Statistical Association, Educational Review, the Journal of Experimental Psychology: Learning, Memory and Cognition, the Journal of Political Economy, Nature, and Science. The magazines were Byte, Fortune, Newsweek, Popular Mechanics, the Economist, Time, and US News and World Report. The newspapers were the Los Angeles Times, the New York Times, the Wall Street Journal, and the Washington Post. In each category, publications were selected to cover a range of interests and to be representative of the medium. Area experts were consulted in the selection of the scientific journals. For each publication, the first issue with a January date and the first issue with a July date from each of the years 1985-1994 were selected for coding, for a total of 3400 issues coded (17 publications x 2 issues / year x 10 years).)

Each issue to be coded was examined by one of two trained coders (a small number of graphs that were missed by the coders were coded by the first author). Coders examined each publication for graphs. A graph was defined to be a figure on the page that used spatial configuration to represent numerical information. Each graph in the publication was identified by page number and recorded. The coder then rated each figure on 39 features, described in Table 1. Questions about coding were resolved by discussion between the two coders and the first two authors.

Insert Table 1 about here

As can be seen in Table 1, the features coded sorted into eight groups. The first group (Type) captured the general type of graph. Often a given graph would be of more than one type (e.g., a figure that is both a bar and a line graph, as in Figure 1). The second group (Dimensionality) captured the apparent dimensionality of the figure. The third group (General) captured some general features of the figure. The fourth group (Orientation) captured whether the figure contained vertically-oriented or horizontally-oriented elements (or both, or neither). The fifth, sixth, and seventh groups captured some specific features of the X, Y, and Z variables, respectively (if each existed). The final group (Legend) captured properties of the legend. Figure 1 and

Figure 2 give examples of graphs coded according to the coding system.

Insert Figure 1 and Figure 2 about here

In addition to the experimental data, each of the two coders also coded one issue of Scientific American from January 1995. Inter-coder reliability was good. For the 11 graphs in this issue, the coders agreed on 417 (97%) of the 429 features rated. The 12 differences in coding can be summarized as: one disagreement on graph type (giving rise to two differences in coding), one disagreement on color vs. black-and-white, one disagreement as to whether 2 graphs were repeated, one disagreement on whether a graph was a time series, three disagreements on whether data points were labeled, one disagreement on whether there was a y-axis grid, one disagreement about whether a figure contained a legend, and one disagreement on the number of levels in a legend.

Data were entered into a database and tabulated by year and issue. Means were calculated by year, publication, and medium.

Results

Number of graphs per issue

The sample consisted of 8159 graphs (5250 from academic journals, 1365 from magazines, and 1544 from newspapers). The mean number of graphs per issue was 24.3. The number of graphs per issue during the period sampled rose sharply for academic journals (from 34.7 in 1985 to 61.2 in 1994) and in newspapers (from 10.1 in 1985 to 24.5 in 1994). For magazines, the number of graphs per issue held roughly to its mean of 9.75 graphs per issue. (See Figure 3.) These data support the notion that graphs are an important means of communication across the range of print media, and their importance has increased in recent years.

Insert Figure 3 about here

Types of graphs

As can be seen in the prior discussion of the coding scheme employed here, there are an impressive variety of types of data graphics available to authors and designers. The different media here appear to utilize these different graph types in systematically different ways. In the sample, journals employed line graphs (72.5%) and point graphs (46.6%) almost exclusively, while magazines and newspapers used mainly line graphs (49.3% for magazines, 50.1% for newspapers) and bar graphs (44.1% for magazines, 27.9% for newspapers). The latter two media also occasionally used pie graphs, and newspapers also occasionally employed cartographs (primarily weather maps) and other graph types (see Figure 4). Note that for each medium, the graph types sum to more than 100%, as a given graph can be an instance of more than one type (e.g., a scatterplot with a curve fit would be both a point and a line graph).

Insert Figure 4 about here

Apparent dimensionality and “3-D” effects

Graphs can be rendered in a number of different styles. One way to vary the rendering style of a figure is to manipulate the apparent dimensionality of the data elements (see Method sections above). Graphic designers (Tufte, 1983), perceptual theorists (Kosslyn, 1993), and experimentalists have all spilled ink on the relative merits of different rendering styles. Tufte’s argument for rendering data using minimalist data elements (“simple” dimensionality, in our terminology) has been influential. The experimental data suggest that in fact viewers of data graphics are more accurate in making magnitude judgments from simple or area graphs rather than volume graphs, but that this effect is relatively small and not particularly robust (Zacks, Levy, Tversky, & Schiano, 1998). Given the range and strength of opinion regarding this issue, it is of interest to see how graphs of varying dimensionality are used across different media.

In the publications sampled, "3-D" graphs were quite rare. Volume graphs made up only 1.6% of our sample, and surface graphs only 1.5%. Academic journals chose primarily simple graphs (85.3%), while magazines and newspapers employed a mix of simple and area graphs (for magazines 47.7% simple and 50.0% area; for newspapers, 49.3% simple and 36.1% area). Furthermore, as Figure 5 shows, the proportion of volume and surface graphs did not increase markedly over the period surveyed. It does, however, seem that for magazines the relative proportion of simple graphs increased, at the expense of area graphs (see Figure 6).

Insert Figure 5 and Figure 6 about here

There were systematic relationships between graph type and apparent dimensionality. Figure 7 indicates that point and line graphs tended to be rendered with simple dimensionality (97.1% of point and 92.9% of line graphs were simple), while bar and pie graphs were generally rendered with filled areas (85.5% of bar and 64.0% of pie graphs were area). Volume and surface rendering was used more heavily in conjunction with the more "esoteric" graph types (pie graphs, pictographs, and cartographs).

Insert Figure 7 about here

"Eye candy:" color and background picture

A number of design features were associated with attention-grabbing, high-impact graphics. Two of these are the use of color (as opposed to black-and-white) printing and of background pictures. In our sample, both of these features varied strongly by medium. Color graphs were dominant in magazines, constituting 69.9% of the graphs appearing there, while they almost never appeared in academic journals (1.0%) or newspapers (0.0%). There was some indication that color graphs increased in prevalence in magazines (see Figure 8). Background pictures were rare over all: Only 1.4% of the figures sampled contained them. Magazines were more likely to contain background pictures (7.7%) than journals (0.1%) or newspapers (0.6%).

Insert Figure 8 about here

Features that support quantitative precision: error bars, grids, labeled data points and color maps

Graphs can be used to convey a rough general message or detailed data patterns. A number of features facilitate the latter usage. Error bars permit the viewer to make inferences about the stability of reported data values. Such information is especially valued in the sciences, so it is not surprising that they were used more in journals (10.7%) than in magazines or newspapers, where they never occurred in our sample. In journals, the use of error bars showed no increasing or decreasing trends.

The addition of a grid helps viewers make fine judgments about data values (Kosslyn, 1993). On the other hand, a grid adds clutter, disdained by Tufte

(1983) and other minimalists. Based on this considerations one might expect more frequent use of grids in scientific publications, in which accuracy is valued and design experience is rarer. Thus, it is somewhat surprising that grids (associated with either the X, Y, or Z axis) were least popular in academic journals (2.0%), and most popular in newspapers (71.6%), with magazines showing an intermediate disposition to use grids (38.8%); no medium showed any trends over time.

Similarly, labeling data points permits exact judgments of numerical values, but labeled data points (either X, Y, or Z values) were also rare in academic journals (1.7%), and more prevalent in magazines (26.0%) and newspapers (14.6%). There was some indication that labeled data points are becoming less common in magazines (see Figure 9).

Color maps, which permit the display of quantitative information about an additional variable, were rare in all three media (0.7% for journals, 0.1% for magazines, and 0.0% for newspapers).

Insert Figure 9 about here

Number and type of variables

As Descartes originally demonstrated, a figure on a page can be an exquisite medium for depicting the relationship between two variables. The habit of assigning of one variable to the horizontal axis and another to the vertical axis is deeply ingrained in western visual literacy. Indeed, horizontal and vertical lines have a privileged status in perception (Howard, 1982). Not surprisingly, the vast majority (94.9%) of graphs in this sample depicted relationships between two variables. This figure agrees reasonably well with

Cleveland's (1984) figure of 83% for scientific publications. However, 11.6% of graphs in magazines depicted only one variable. These were typically figures that picked out a set of values along a single continuum, as in a time-line. Such figures were very rare (< 1%) in journals and newspapers.

One basic function of a data graphic is the use of space to represent one or more quantitative variables (Bertin, 1980). It is therefore not surprising that most graphs in our sample (97.6%) included at least one quantitative variable. Quantitative variables can be plotted against other quantitative variables or against qualitative (categorical) variables. In our sample, figures in magazines were most likely to contain at least one qualitative variable in the X, Y or Z axes or in the legend (38.9%). Newspapers and journals also made some use of figures with qualitative variables (16.3% and 11.9%, respectively).

Repetition

Plotting several data sets on the same or similar sets of axes can be a powerful technique for making visual comparisons (Cleveland, 1985). All the media studied made substantial use of this technique (35.9% of graphs in journals , 30.3% in magazines, and 46.6% in newspapers). During the period studied, the prevalence of repeated graphs in newspapers increased, and their prevalence in magazines decreased (see Figure 10).

Insert Figure 10 about here

Orientation

Graphical elements vary in their orientation. As described in the Method section, a graphical element can depict a data value by its horizontal or vertical extent or location. Most of the figures in the sample (88.6%) used vertically-oriented elements; a few (6.4%) contained horizontally-oriented elements. As can be seen in Figure 11, horizontally-oriented elements appeared mostly in magazines, and seem to be on the decline. (A given figure could contain vertical elements, horizontal elements, neither, or both.)

What might go into the choice of vertical or horizontal orientation? Kosslyn (1993, p. 38) suggests three principles. First, let reality decide between vertical and horizontal elements. Second, use horizontally-oriented elements if the labels are too long to fit under a vertical display. Third, when in doubt use vertically oriented elements because that is the convention. This last principle may be related to the very general association of "up" with "more" across cognitive domains (Clark & Clark, 1977; Lakoff & Johnson, 1980; Tversky, 1995). Given that some of the data being depicted surely have a natural horizontal orientation, and given that long labels indicate for horizontal orientation, the dominance of vertical orientation suggests that designers' choice of orientation is being driven by social convention. It would be interesting to examine the genealogy of this convention by historical study.

Insert Figure 11 about here

Discussion

The most straightforward result of this study is its confirmation that data graphics are an important component of communication in print media. The average publication in our sample had more than 24 graphs per issue. Clearly, as a culture we rely heavily on data graphics to communicate about quantitative information. Data graphics took on a wide variety of forms and styles, indicating a fertile ground for the study of their use.

Moreover, we stand at a particularly interesting time in the history of data graphics. Recent increases in the speed and power of personal computers have powerful design tools available to the “manufacturers” of data, eroding the boundaries between author and designer. For example, in 1999 it takes about 10 minutes to download from the Internet and install a program that allows one to create 3-D rendered charts using texture-mapping, modeled objects and photorealistic background images (Adrenaline Software, 1993). Has this increase in the availability and sophistication of graphing technologies had profound affects on common usage?

One way graphing software may affect graphical communication is by simply making it easier for producers of print articles to generate data figures to accompany them. One would expect such effects to be most pronounced in scientific publications, in which the author(s) of an article also typically generate any figures that accompany it. During the period surveyed, the number of graphs per scientific journal issue nearly doubled, indicating that researchers are indeed taking advantage of the new tools.

But what kinds of graphs are being produced? In fact, the features identified with high-powered computing systems for data graphics seem to be the same ones that are abhorred by critics of graphical practice. Tufte’s (1983) well-know imperative to minimize the “data-ink ratio” provides a concise formulation of a minimalist ethic in design. Features such as added 3-D appearance, background pictures, and pictographs all violate this maxim. These are just the features one is likely to see on the packaging for an up-and-coming graphics software package.

Given that more powerful systems make it easier to produce graphs with precisely the features the experts decry, one might suspect (as we did) that graphs with these features are becoming more common. This intuition was not supported by these data. On the contrary, features such as 3-dimensional appearance and background pictures were rare and do not seem to be increasing in popularity. Pictographs were seldom used. The modal graph in our sample was a simple line graph. Perhaps the impression of increasing use of graphics that are more glitzy than communicative has come from a few blatant examples.

Just as graphics tended to be conservative in style, the data they represented tended to be plain in structure. Rather than depicting complex high-dimensional interactions, most figures in the sample showed relationships between two variables.

The choices by authors and designers as to how to render data are probably governed both by features in the data and by conventions particular to their community. Some design choices are constrained by the data to be depicted. Categorical variables are typically not used with point graphs, probably because interpreting them would require difficult visual discriminations. A large number of levels of a categorical variable makes drawing a pie graph unwieldy, and long category names encourage the use of horizontally-oriented figures, to allow room for the labels. Other design choices may reflect the influence of one’s cultural community or of the genre of communication. For example, in our sample scientific journals used almost exclusively simple rendering styles, while magazines and newspapers tended to use a mix of simple and area rendering. (Such effects may also reflect differences between the media in the kinds of data being graphed.) For a discussion of how conventions can emerge in graphic communication see (Zacks & Tversky, in press).

In sum, worries about negative recent developments in the design of data graphics were not supported by this survey. The graphics coded were by and large simple and elegant in exactly the ways advocated by designers (Tufte, 1983; Tufte, 1990) and psychologists (Kosslyn, 1993). To the extent that developments in computing hardware and software have influenced data graphics, these influence don’t seem to be obviously negative.

The placement of sophisticated tools in the hands producers of graphs offers a unique opportunity for experts in graphical design and perceptual psychology to influence the production of data graphics. Systems for data graphics differ in the features they afford. By embedding principles of graphical perception in software packages, one might hope to guide authors’ and designers’ choices in felicitous directions. One example of such a tight integration between graphical perception and software design is the Trellis graphics library for S-Plus (MathSoft, 1995). Designed for exploratory data analysis, this system reflects the influence of a large body of research on the part of its designers (Cleveland, 1985). The application of a similar approach to packages for publication graphics for scientific authors and media designers would no doubt be of great value.

In advocating that designers and perceptual psychologists build their understanding of data graphics into computer systems, one major caveat is in order. Our experience has been that most variables of interest in graphical design interact with several other variables. This means that simply designing a system with a good default value for each feature is likely to be insufficient.

Rather, good solutions are likely to resemble case-based reasoning systems [Kerpedjiev, 1998 #892; see also Fish & McCartney, Kerpedjiev this volume], in which exemplars of situations similar to the one at hand are identified and elegant solutions to those examplar situations (as identified by research or design) are recommended.

References

Adrenaline Software, I. (1993) [Computer Software]. Charts Pro 1.04. Québec, Canada: Author.

Bertin, J. (1967). Sémiologie graphique. Les diagrammes, les réseaux, les cartes. Paris: La Haye Mouton.

Bertin, J. (1980). The basic test of the graph: a matrix theory of graph construction and cartography. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), Processing of visible language 2 (pp. 585-604). New York: Plenum Press.

Bertin, J. (1983). Semiology of graphics (William J Berg, Trans.). Madison, Wis.: University of Wisconsin Press.

Carswell, C. M. (1992). Choosing specifiers: An evaluation of the basic tasks model of graphical perception. Special Issue: Visual displays. Human Factors, 34, 535-554.

Casali, J. G., & Gaylin, K. B. (1988). Selected graph design variables in four interpretation tasks: A microcomputer-based pilot study. Behaviour & Information Technology, 7, 31-49.

Clark, H. H., & Clark, E. V. (1977). Psychology and language : an introduction to psycholinguistics. New York: Harcourt Brace Jovanovich.

Cleveland, W. S. (1984). Graphs in scientific publications. The American Statistician, 38, 261-269.

Cleveland, W. S. (1985). The elements of graphing data. Monterey, CA: Wadsworth Advanced Books and Software.

Cleveland, W. S., Harris, C. S., & McGill, R. (1983). Experiments on quantitative judgments of graphs and maps. The Bell System Technical Journal, 62, 1659-1674.

Gattis, M., & Holyoak, K. J. (1996). Mapping conceptual to spatial relations in visual reasoning. Journal of Experimental Psychology. Learning, Memory, and Cognition, 22, 231-239.

Howard, I. P. (1982). Human visual orientation. New York: John Wiley & Sons.

Kosslyn, S. (1993). Elements of graph design. New York: W.H. Freeman and Company.

Lakoff, G., & Johnson, M. (1980). Metaphors we live by. Chicago: University of Chicago Press.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. Cognitive Science, 11, 65-100.

Lohse, G. L., Walker, N., & Rueler, H. H. (1994). A classification of visual representations. Communications of the ACM, 37, 36-49.

Macdonald-Ross, M. (1977). How numbers are shown. AV Communications Review, 25, 359-409.

MathSoft. (1995) [Computer Software]. S-Plus 3.3. Cambridge, MA: Author.

Pinker, S. (1990). A Theory of graph comprehension. In R. Freedle (Ed.), Artificial Intelligence and the future of testing (pp. 73-126). Hillsdale, NJ: Lawrence Erlbaum Associates.

Shah, P., & Carpenter, P. A. (1995). Conceptual limitations in comprehending line graphs. Journal of Experimental Psychology: General, 124, 43-61.

Simkin, D., & Hastie, R. (1987). An information-processing analysis of graph perception. Journal of the American Statistical Association, 82, 454-465.

Spence, I. (1990). Visual psychophysics of simple graphical elements. Journal of Experimental Psychology: Human Perception & Performance, 16, 683-692.

Tufte, E. R. (1983). The Visual display of quantitative information. Cheshire, Connecticut: Graphics Press.

Tufte, E. R. (1990). Envisioning information. Cheshire, Conn.: Graphics Press.

Tukey, J. W. (1977). Exploratory data analysis. Reading, Mass.: Addison-Wesley Pub. Co.

Tversky, B. (1995). Cognitive origins of graphic conventions. In F. T. Marchese (Ed.), Understanding images (pp. 29-53). New York: Springer-Verlag.

Tversky, B., & Schiano, D. J. (1989). Perceptual and conceptual factors in distortions in memory for graphs and maps. Journal of Experimental Psychology: General, 118, 387-398.

Wainer, H. (1984). How to display data badly. The American Statistician, 38, 137-147.

Zacks, J., Levy, E., Tversky, B., & Schiano, D. J. (1998). Reading bar graphs: Effects of extraneous depth cues and graphical context. Journal of Experimental Psychology: Applied, 4, 119-138.

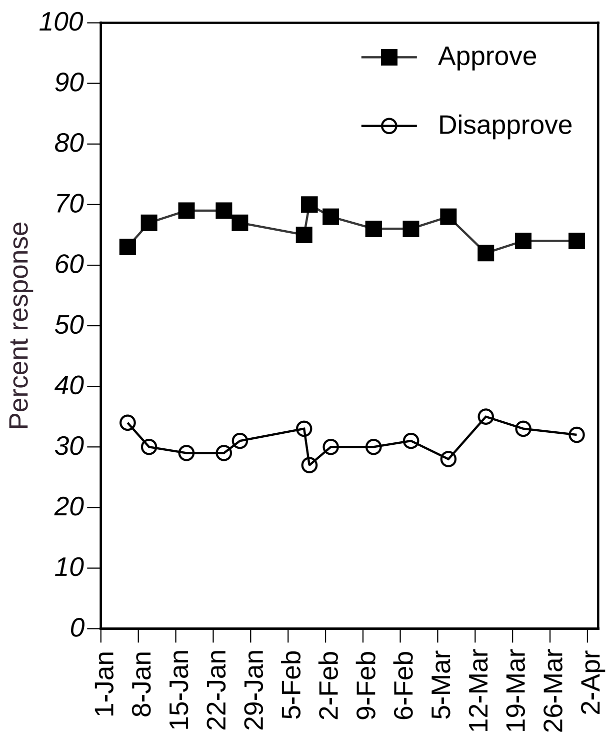
Zacks, J., & Tversky, B. (in press). Bars and lines: a study of graphic communication. Memory and Cognition.

Table 1: Features of graphs coded in this study. Groups of features are indicated in the left column, individual features in the middle column, and their definitions in the right column.

	Feature	Description
Type	point	True if the figure contained data elements that were points whose height and width were not informative.
	line	True if the figure contained data elements that were lines whose width was uninformative.
	bar	True if the figure contained data elements that were rectangular areas of consistent height or width.
	pie	True if the figure contained a data element that was a circle or ellipse whose area was subdivided from the center.
	pictograph	True if the figure used the scale of a picture, or the number of a set of repeated pictures, to convey a numerical value.
	cartograph	True if the figure contained a map onto which numerical data were projected (by any technique). If the cartograph was a weather map, this was noted.
	stem and leaf	True if the graph was a stem-and-leaf plot, as described by Tukey (1977) or any variant.
	box and whiskers	True if the graph was a box-and-whiskers plot, as described by Tukey (1977) or any variant.
	other	True if the graph was of a type other than those listed above.
Dimen-sionality	simple	True if the figure contained data elements with negligible enclosed area (e.g. points, lines, curves).
	area	True if the figure contained data elements that formed enclosed areas (shaded or not) and did not give the appearance of depth (e.g. rectangles, irregular polyhedra).
	volume	True if the figure contained data elements that gave the appearance of solid volumes.
	surface	True if the figure contained data elements that gave the appearance to surfaces with extent in depth, but negligible volume.
General	color vs b/w	Coded as “color” if the figure contained multiple hues, else “black and white.”
	repeated	If the figure was a member of a group that used a similar frame that was repeated separated by space, coded as true. If the figure contained multiple data sets drawn on the same connected axes, the number of data sets depicted was drawn.
	background picture	True if the figure contained a background picture.
	error bars	True if the figure contained error bars.
Orientation	horizontal	True if the figure contained elements that depicted numerical values using horizontal extent.
	vertical	True if the figure contained elements that depicted numerical values using vertical extent.

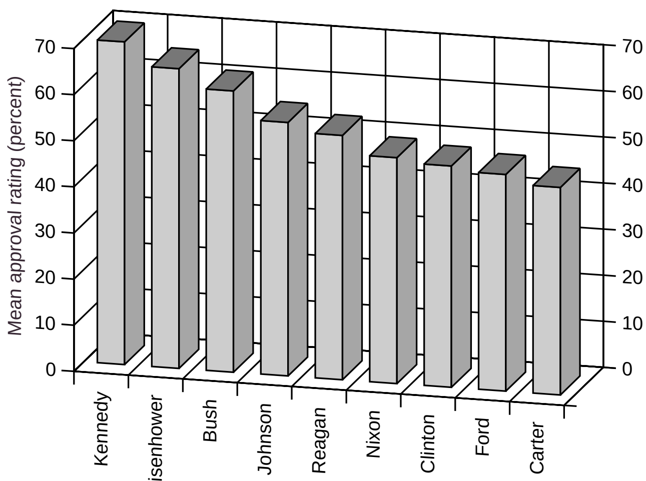
	Feature	Description
	stacked	True if the figure contained elements that depicted numerical values using the distance between adjacent elements.
X Variable	qualitative	True if the dimension represented a nominal variable.
	quantitative	True if the dimension represented an ordinal, interval, or ratio variable. For discrete quantitative variables, the number of levels was noted.
	time series	True if the quantity described was a time or date.
	labeled data points	True if any the data points were accompanied by text labels.
	grid	True if the figure contained a horizontal and/or vertical grid.
Y Variable		(Same features as "X Variable" above.)
Z Variable		(Same features as "X Variable" and "Y Variable" above.)
Legend	qualitative	True if the dimension represented a nominal variable. For legends, the number of levels in the legend was recorded, and the dimension(s) to which the legend referred (x, y or z) was also noted.
	quantitative	True if the dimension represented an ordinal, interval, or ratio variable. For discrete quantitative variables, the number of levels was noted. For legends, the dimension(s) to which the legend referred (x, y, or z) was also noted.
	time series	See above.
	color map	True if the legend was a colormap.

Figure 1: Example graph to demonstrate the coding scheme. (Data are presidential approval ratings for January-March 1999.)



Type:	point and line
Dimensionality:	simple
General:	color, background picture
Orientation:	vertical
X Variable:	quantitative, time-series
Y Variable:	quantitative
Z Variable:	none
Legend:	qualitative

Figure 2: Second example graph to demonstrate the coding scheme. (Data are mean presidential approval ratings for the last 9 presidents)



Type:	bar
Dimensionality:	volume
General:	black & white
Orientation:	vertical
X Variable:	qualitative
Y Variable:	quantitative, grid
Z Variable:	none
Legend:	none

Figure 3: Number of graphs per issue by medium and year.

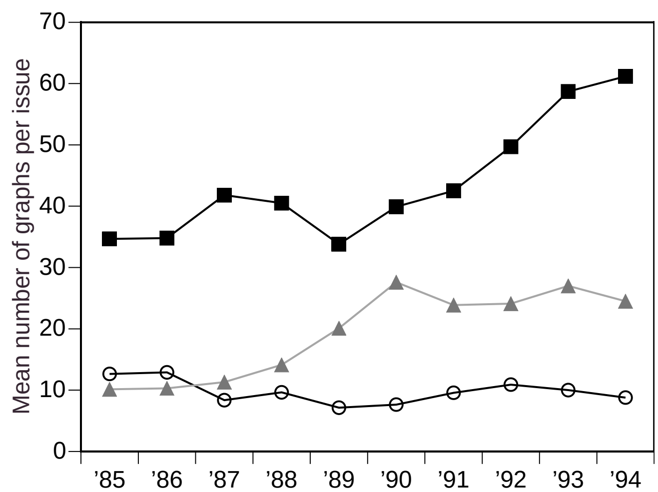


Figure 4: Distribution of graph types by medium. Each pane shows the proportion of graphs of each type for one medium. (Pictographs, cartographs, stem-and-leaf plots, and box plots are collapsed into the "Other" category because they occurred infrequently.)

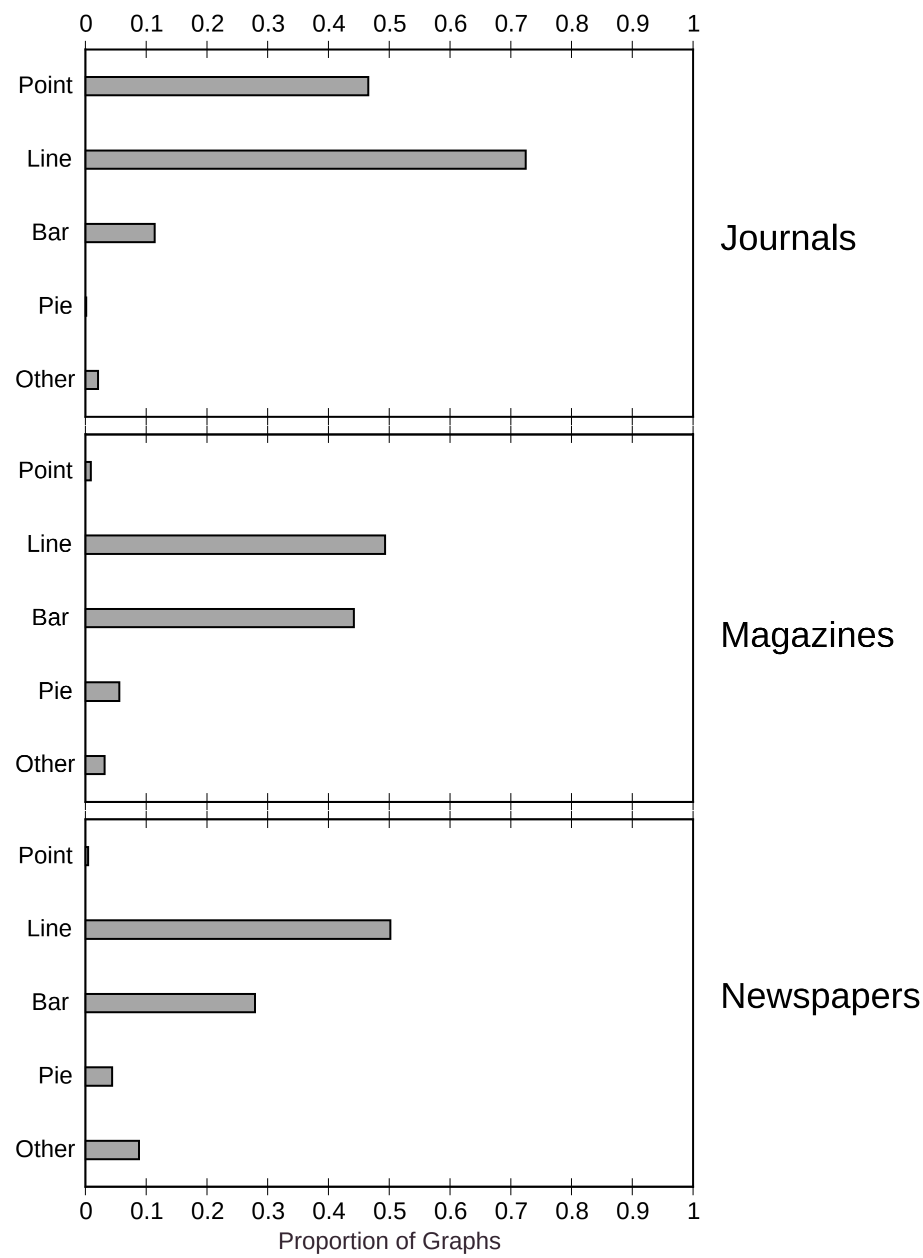


Figure 5: Volume and surface ("3-D") graphs are rare and do not seem to be on the rise.

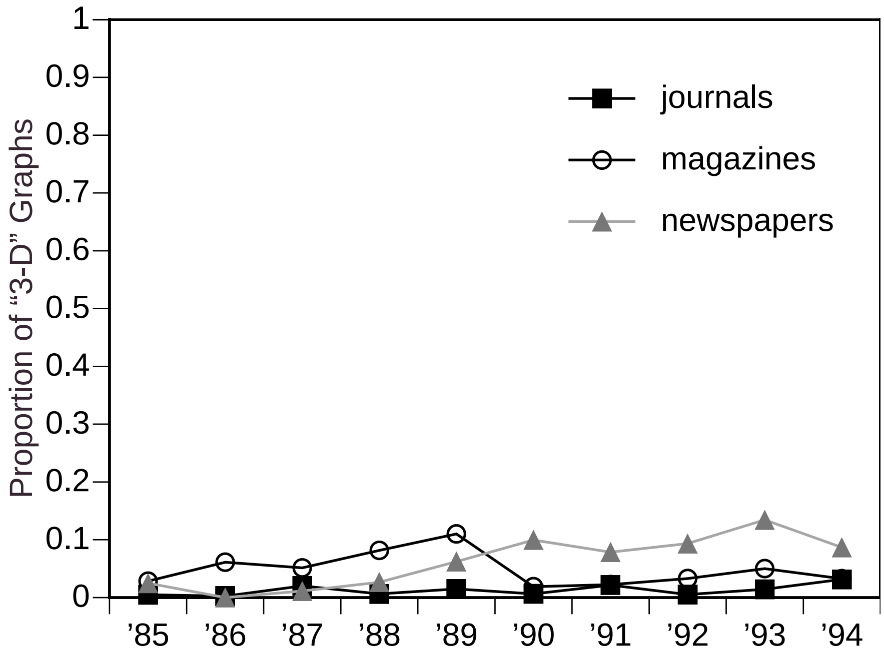


Figure 6: The prevalence of simple graphs in magazines and newspapers showed a modest increase.

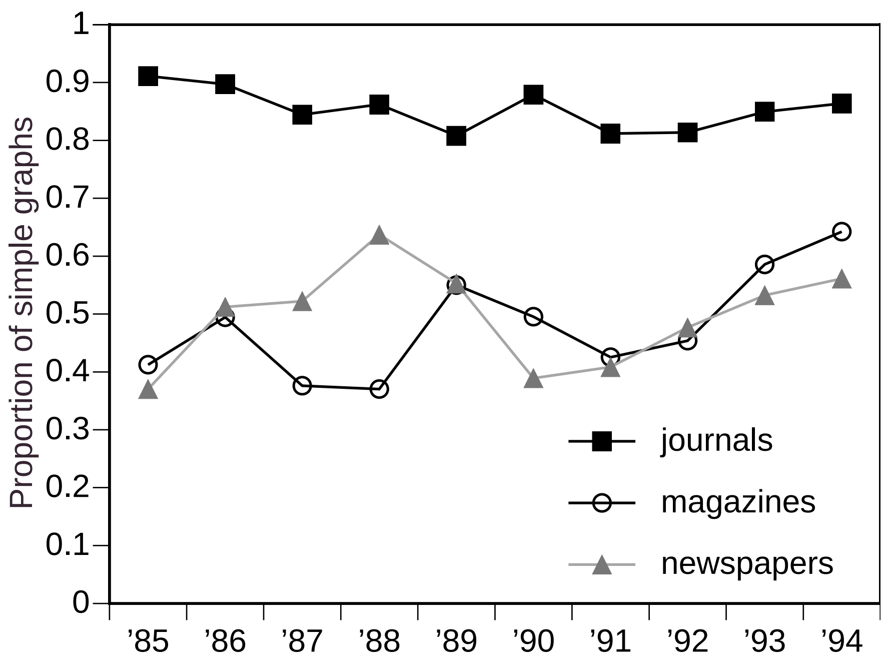


Figure 7: Distribution of apparent dimensionality by graph type. Each pane shows the proportion of graphs of one type that appeared in each apparent dimensionality. (Stem-and-leaf plots, box plots, and those classified as "other" are omitted because they occurred infrequently, as are graphs whose dimensionality could not be classified.)

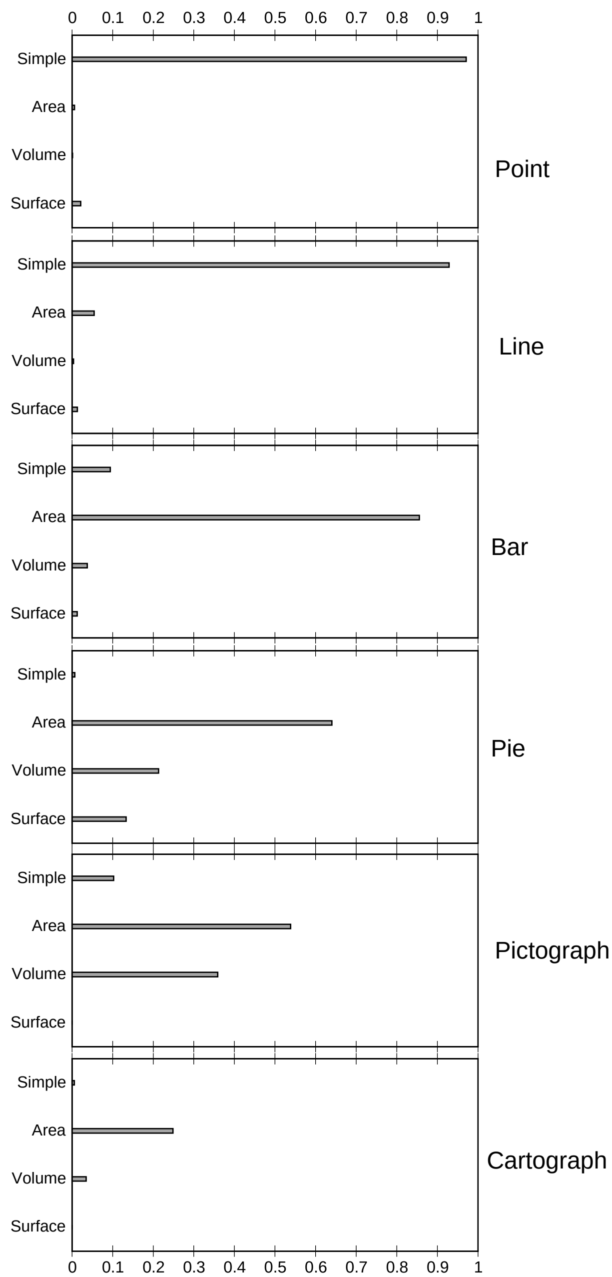


Figure 8: Color graphs appeared predominantly in magazines, where they may be on the rise.

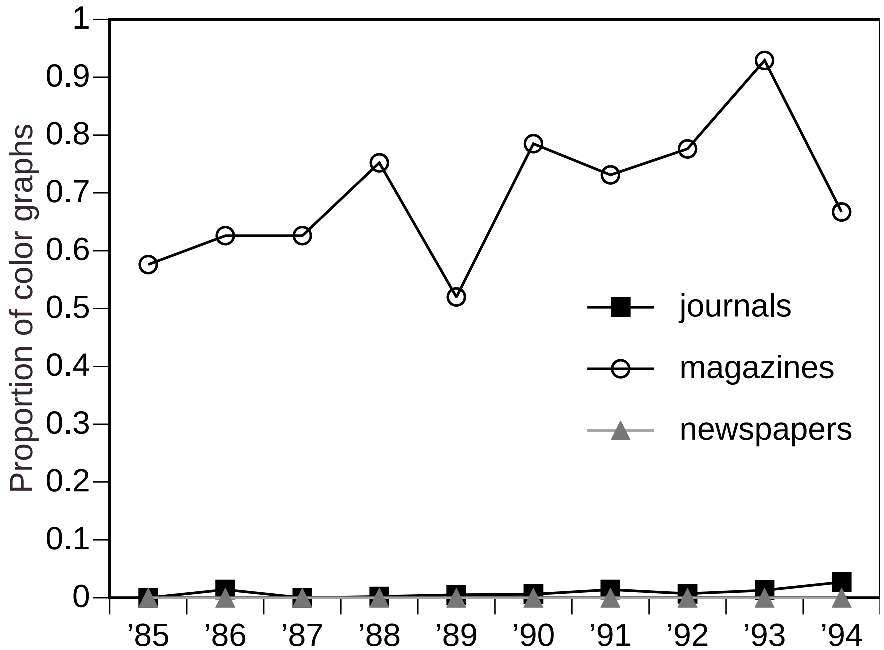


Figure 9: Data points with either X or Y values labeled were more prevalent in magazines and newspapers than academic journals, and seem to be on the decrease in magazines.

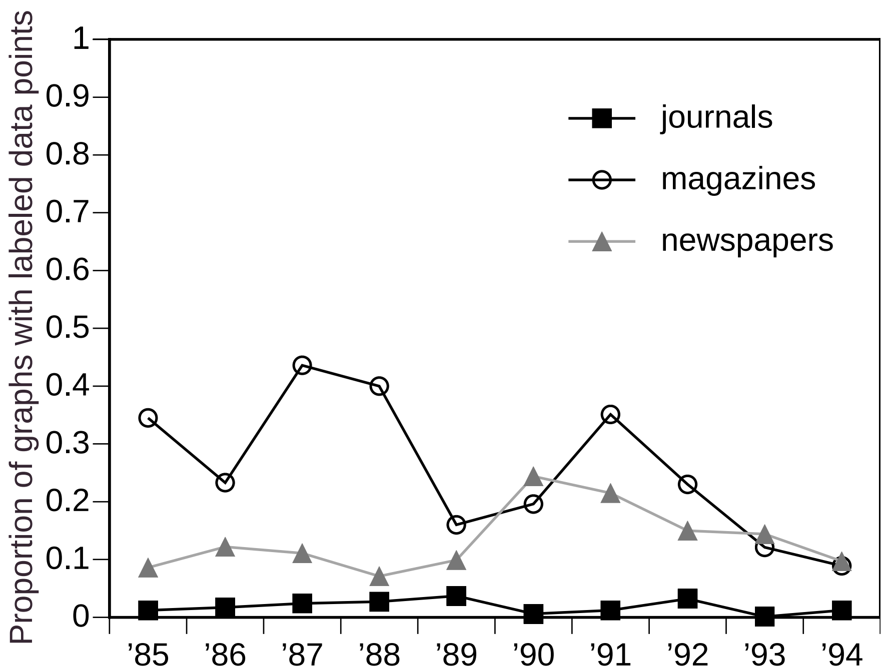


Figure 10: Repeated graphs became more prevalent in newspapers and less so in magazines during the time period sampled.

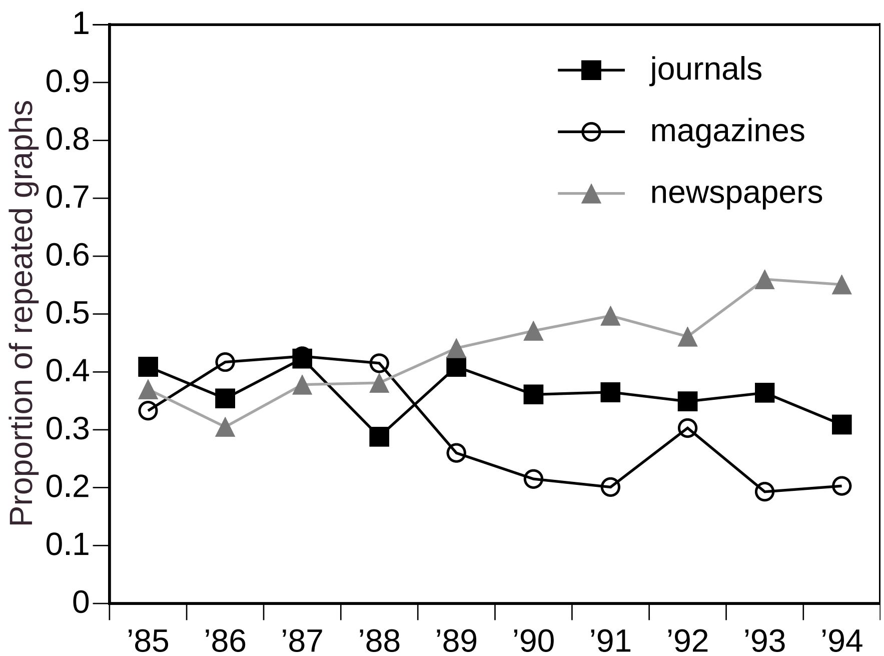


Figure 11: Graphs with horizontally-oriented elements appeared mostly in magazines, and showed a decreasing trend.