A Review of Edward R. Tufte's "The Visual Display of Quantitative Information"

John W. Eshleman
Atlanta

Book Reviewed: Tufte, E.R. (1983). *The visual display of quantitative information.* Chesire, Connecticut: Graphics Press. Hardbound

Many methods exist for you to present information to others. You can render verbal descriptions of the information. You can quantify the information and give to others the numbers. You can sort and put the numeric information into tables. And you can illustrate the information by way of picture or graphic. Each method has its strengths and weaknesses, depending on how you apply the method. Many of us quantify information, and find it expedient as well as useful to illustrate the information. Accordingly, we resort to using charts and graphs, or other types of diagrams or figures. The basic idea is to convey information to others so that we communicate effectively to them.

Charts, graphs, and even tables of numbers as well as verbal descriptions all serve as discriminative stimuli. The concept of "discriminative stimulus" comes from behavior analysis. Basically, it refers to some object or event that precedes a behavior in time. A discriminative stimulus also, in a loose sense, represents a sign or signal to the person to take some action or to make some decision. Charts, graphs, and other figures can help our reasoning when we need to interpret information. They can assist us in making decisions based on the information. They can instruct others about what we discovered. But charts, graphs, and other figures can also distort information, conceal information, or be used for persuasion instead of honest interpretation. Charts constructed for persuasive purpose may deceive people. They may make us commit errors in our reasoning. They may lead us to make poor decisions.

What forms the essence of science? To a large degree, science consists of how you collect information and then put it together so that it may be interpreted productively. You need to collect the right information. The information will likely include measures and quantities, which will allow you to make comparisons and decisions. You need to make sure that when you quantify information that you do so in the most clear, honest way possible. This means selecting measures that tell us something useful about what it is you are measuring. Are you measuring trends? Are you comparing values? Do you want to see the effect of changing the values of an independent variable? Do you want to see how people, behaviors, or events distribute themselves? If so, then you need the right units of measure.

Having the right measures helps, but it helps further to view these measures so you can reason well about what they indicate, and to make decisions based on them. While a series of numbers may indicate a trend, a trend may be better depicted as an angled line. Hence, a chart of those numbers works well to let us view the trend, and does a better job than the numbers. You can see at a glance whether a line is angled upward, flat, or downward. Putting numeric measures onto a chart or graph represents, then, the visual display of quantitative information.

Nearly 20 years ago Edward R. Tufte published a book dealing with the visual display of quantitative information. He analyzed what makes for a good display, and singled out practices that lead away from good displays. What makes a chart useful and helpful to the reader? What leads the reader astray? What are the do's and don'ts of charting? These questions form the topic of Tufte's book, one that skilled Precision Teachers and researchers will find highly valuable. For the person interested in deep issues concerning graphical display of quantitative information, Tufte's book provides invaluable background information. It can also serve as a guide for us to check up on our own practices. Is the Standard Celeration Chart, by itself, good at helping the visual display of quantitative information? Can we misuse the chart? Can we lower its effectiveness? Can we use it to persuade people about something we advocate? Can we conceal what happened? Armed with Tufte's book, we can address these issues and concerns.

Tufte divided the book into two main parts: Part 1 deals with "Graphical Practice," and covers the topics of graphical excellence, graphical integrity, and sources of graphical integrity and sophistication. Part 2 deals with the "Theory of Data Graphics," and covers the topics of data-ink and graphical redesign, chartjunk: vibrations, grids, and ducks, data-ink maximization and graphical design, multifunctioning graphical elements, data density and small multiples, and aesthetics and techniques in data graphical design.

Excellence in Graphics

Tufte begins the chapter on Graphical Excellence by commenting: "Excellence in statistical graphics consists of complex ideas communicated with clarity, precision, and efficiency." This forms a useful working definition of excellence. It indicates, perhaps, a goal we should strive for when we use charts. Tufte next identifies nine criteria of excellence in graphics used to display quantitative information. He states that "Graphical displays should:

- show the data

- induce the viewer to think about substance rather than about methodology, graphic design, the technology of graphic production, or something else

- avoid distorting what the data have to day

- present many numbers in a small space

- make large data sets coherent

- encourage the eye to compare different pieces of data

- reveal the data at several levels of detail, from a broad overview to the fine structure

- serve a reasonably clear purpose: description, exploration, tabulation, or decoration

- be closely integrated with the statistical and verbal descriptions of a data set." (page 13).

Without exception, the Standard Celeration Chart (Pennypacker, Koenig, and Lindsley, 1972) has the potential to meet all of Tufte's criteria for graphical displays. The cumulative record as developed by B.F. Skinner, an early staple of behavior analysis, meets those same criteria as well. Both of those charts are standard charts, having standard dimensions, whereon the "meaning" of dots and lines always runs with respect to particular standards. For instance, on the Standard Celeration Chart a line drawn from bottom left corner to top right corner has an angle of 34 degrees, which always represents a celeration of X2 (read as "times two.") Any line parallel to that standard also has a celeration value of X2. A grid of "standard celerations" can be drawn for reference celerations; typically this grid includes referencecelerations for X4, X2, X1.5, X1.25, X1.1, and X 1.0. Similarly, the cumulative record, which Lindsley has described on several occasions as a "standard frequency chart," also has standard reference grids. These grids used to be part of the visual display in older publications. A certain angle would be a frequency or rate of 1 per minute. An angle somewhat steeper would reference a rate of 2 per minute. A yet steeper line, about the same angular distance as from 1 to 2 would reference 4 per minute, and so on. In addition, both charts show data of measures that are at once universal, standard, and absolute measures.

Both the Standard Celeration Chart and the Standard Frequency Chart (cumulative record):

(a) show the data

(b) induce viewers to think about the substance of the graphic rather than the methodology of graphic production or design

(c) clearly avoid distorting what the data say

(d) present many numbers in a relatively small space

(e) make large data sets coherent

(f) encourage the eye to compare different pieces of data (indeed, compel such comparisons)

(g) reveal the data at several different levels of detail (e.g., the Standard Celeration Chart simultaneously illustrates frequencies, but also celerations, and variability, cycles and various trends; the Standard Frequency Chart similarly illustrates local rates and overall rates.)

(h) serve a very clear purpose, which is to show precisely and unambiguously a clear record of behavior and behavior change, thereby permitting not only precise and reliable description but also allowing for exploration. In the case of the Standard Celeration Chart straight line celeration predictions can be made.

and finally,

(i) both charts are very closely integrated with the statistical and verbal descriptions of the data sets. In the case of the Standard Frequency Chart (cumulative record) in particular, the angled lines and tick marks often ARE the data set. Typically, precision teachers do not make statistical transformation or conversion from numbers to graphics. Instead, they record the data directly and continuously as dots on the Standard Celeration Chart.

In contrast, most of the time the various other charts and graphs used in behavior analysis violate Tufte's suggested standards of excellence. Or if they do not violate these standards they may fail to meet them. Often, many behavioral researchers ignore several of Tufte's standards at the same time. Within the same publication the same data may be put on charts with different axes, but where the grids have the same physical dimensions. Axes may be stretched or compressed on either or both dimensions, thereby changing how the plotted data will look. Small effects may appear magnified. Large differences may appear muted. Some results such as trends or variations in the data may not show up at all. Sometimes the variations in design of behavior analytic charts unnecessarily induces viewers to spend time figuring out what the chart attempts to show.

Researchers may ignore Tufte's suggested standards of excellence for many reasons . These may have to do in part to a lack of awareness with respect to those standards. Researchers or scholars may not have been educated in the issues and problems related to graphic display. It may also reflect the operation of various macrocontingencies operating at the cultural or institutional level. Common and accepted practices for particular professions, including applied behavior analysis, may select against and thus discourage any effort to meet any of Tufte's standards. This may be especially true if there are political or economic contingencies operating that may affect how data are displayed graphically. This may even be true on a subtle level. For instance, a chart that magnified a small effect may help toward publication of an article. To what extent such macrocontingencies exist and play any role can only be speculative at this point, and is beyond the scope of this mini-review -- though it may prove an interesting investigation.

Integrity

Tufte's second chapter, on Graphical Integrity, seems more interesting than chapter one, in my opinion. Tufte begins chapter two by observing, "For many people the first word that comes to mind when they think about statistical charts is "lie."" (page 53). This forms a key issue. Tufte adds that "a graphic does not distort if the visual representation of the data is consistent with the numerical representation." (page 55). From these observations Tufte introduces the "Lie Factor" -- a valuable concept curiously overlooked in Iver Iverson's (1988) review of Tufte's book in the Journal of the Experimental Analysis of Behavior. Tufte's formula for the Lie Factor appears relatively simple:

$$\text{Lie Factor} = \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}}$$

In his chapter Tufte presents several classic examples of severe Lie Factors. For instance, he presents a graph published in the New York Times (August 9, 1978) pertaining to fuel economy standards based on Congressional and Department of Transportation data. These were standards that car manufacturers would have to meet. The data set began with 18 miles per gallon in 1978, and incremented to 19 in 1979, 20 in 1980, 22 in 1981, 24 in 1982, 26 in 1983, 27 in 1984, and 27.5 in 1985. The chart of these data in the Times, however, was drawn as perspective of a highway extending to a vanishing point in the distance. Imagine a perspective drawing of a highway, with the angles of the edges of the road converging and meeting in the distance. The drawing even included shading and the typical dashed lane marker lines characteristic of roads. On this chart the fuel economy standards were drawn as horizontal lines from one side of the road to the other. The line for 1978 was drawn in the distance, near the vanishing point. The lines for subsequent years were drawn down the chart, so that the line for 1985 was near the bottom of the chart. The long and the short of it, so to speak, was that the line for 1978 was 0.6 inches long, whereas the line for 1985 was 5.3 inches long.

The increase in the raw data is computed as 27.5 - 18.0 / 18.0 X 100, which equals 53%. Over that span of time fuel standards would increase 53%. However, as Tufte notes, "the magnitude of the change from 1978 to 1985 is shown in the graph by the relative lengths of the two lines:"

5.3 inches - 0.6 inches / 0.6 inches X 100 = 783%.

Visually, then, the graph compels an interpretation of a 783% increase in fuel standards, while the actual data denoted a 53% increase. Tufte then computed the Lie Factor for this chart:

$$\text{Lie Factor} = \frac{783}{53} = 14.8$$

In other words, the New York Times' graph lied about the data by a factor of 14.8 times.

Since a picture is worth a thousand words, the point becomes much more clear if you can see the graph that Tufte uses for an example, rather than to try to visualize it from the words written above. Tufte includes the charts in his book, for which he obtained permission to use them.

An interesting research task would be to ascertain how widespread lie factors are found in behavioral research graphical data presentations. With standard charts you are unlikely to find lie factors, of course. But with the tendency for behavioral researchers to draft their own charts and graphs the possibility for lie factors certainly presents themselves.

Tufte adds that the ideal Lie Factor has a value of 1,which means the size of the effect shown in the graphic matches the size of the effect in the data. He's not absolutely strict, however, and suggests that the range of Lie Factors from .95 to 1.05 are probably acceptable, and include minor inaccuracies in plotting. Tufte suggests that "The logarithm of the Lie Factor can be taken in order to compare overstating (log LF > 0) with understating (log LF < 0) errors" (page 57). He adds that most such distortions are those of overstating; that Lie Factors as big as 5 are relatively common.

When I first read about Tufte's concept of the Lie Factor I immediately thought of Og Lindsley's frequent commentary about "fill the frame" charts, also known as "stretch to fill" charts. Og corrected me by saying that "stretch to fill charts" are not indicative of Lie Factors per se. Og came up with the terms "stretch to fill" and "fill the frame" in response to a general tendency of behavioral researchers to draw graphs where they "stretch out" or compress the chart axes in order to fill up the resulting frame with their data. The researcher stretches or compresses one or both of the axes of a chart such that a particular effect is illustrated. By stretching the vertical axis a small effect can be made to appear larger. By compressing an vertical axis a large effect can be made to look smaller. Likewise, stretching a horizontal axis can "attenuate" a trend, giving it a lower angle.

Stretching an axis and filling the frame can produce differences to interpretation of results. Porter (1985), for instance, discovered that about one-third of the effects published in the Journal of Applied Behavior Analysis were counter-turns. A counter-turn occurs either when a frequency jump up is followed by a celeration turn-down, or a frequency jump-down is followed by a celeration turn-up. Either way, the behavior heads back towards its baseline level. The effect is only transitory, in other words. Porter discovered these counter-turns when recharting behavioral data onto Standard Celeration Charts. The effects were either invisible or at best barely noticeable in their original charts. Whether the researchers created charts to conceal the counter-turns remains unanswerable, but the outcome is the same. By the way the chart is drawn the reader may not notice the counter-turn and may therefore not conclude that the procedure investigated has an effect potentially at variance to what is claimed.

Related to stretching axes is the practice of truncation, where only a segment of an axis is shown. This segment, of course, can then be stretched out to fill the vertical or horizontal space on paper. Stretching and truncation are issues that Tufte deals with in a section of chapter two titled "Context is Essential for Graphical Integrity." What this means is that charts that have had axes manipulated to show certain results in a certain way are at the very least deceptive to the extent that vital contextual information is removed or distorted.

One of the key features of the Standard Celeration Chart is that it avoids both stretching of axes as well as truncation. Thus, context is always preserved. Now, often data on the Standard Celeration Chart will not cross all six logarithmic cycles. The Chart captures frequencies from .001 per minute (about once per day) in the lowest cycle up six cycles to frequencies of 1000 per minute (a X 1,000,000 differential). Because few behaviors ever traverse all six cycles there may be a tendency to regard the unaffected cycles as superfluous, and hence worthy perhaps of being dispensed with. Under such a tendency there would be a further tendency to stretch out the remaining axes to fill the space left by the excised axes. That, of course, would defeat the purpose of having a Standard chart, as well as violate Tufte's admonition against removing context. In any event, even the blank cycles provide useful context in that they help locate the frequencies of behavior on a standard frequency scale. To help make the point, think of the location of the visible light portion of the electromagnetic spectrum and the value of including the whole spectrum for contextual purposes. If all you showed was that portion of the spectrum pertaining to visible light the viewer would not be able to tell where it stood with respect to other types of EM radiation. Context is valuable.

Tufte's Theory of Data Graphics

The second part of Tufte's book covers his theory of data graphics. Here, he discusses vital issues such as "data ink," "chartjunk," multifunctioning graphical elements, data density, and finally aesthetics and technique in graphical design. Only some of these issues concern standard celeration charting. Furthermore, many of these issues will pertain primarily to published charts, not necessarily to those used on an everyday basis as working charts.

Data-Ink

In the chapter on Data-Ink Tufte presents another valuable ratio: The amount of ink used to make a graphic that presents data or information. Ink that does not present data or information is not considered to be data-ink. Note that the word ink already suggests something that has been published or made ready for archival purposes.

The data-ink ratio equals the amount of data-ink divided by the total ink used to print the graphic. As he suggests, one way of viewing this ratio comes from how much of a graphic you can erase without loss of data or information. Because precision teachers use commercially manufactured charts, the data ink issue may be of marginal concern: People use the same charts and have no need to draw their own. Moreover, working charts build up a data-set over time. Thus, a chart with a lot of unused lines and unused space may seem, at first, to have a poor data-ink ratio. But as days proceed and frequencies charted, this ratio would surely change.

The data-ink issue may have more relevance when precision teachers present or publish their results. At chart shares at ABA conventions speakers sometimes use the transparency templates. Some do not, however. Some people simply show the data without any grid, which exemplifies the point that Tufte makes regarding data-ink. Likewise, in our publications our published charts rarely show the internal up and down day lines or the horizontal frequency lines. Indeed, removing this mass of data-ink actually helps viewing the data.

Chart Junk

Chart junk, as best I can tell from reading the chapter, concerns various decorations or additions made to charts. Such charts become cluttered. They may produce unintended optical effects, too, such as Moiré patterns. Again, Tufte approaches the issue from the point of view of people creating their own charts.

In precision teaching a blank Standard Celeration Chart will contain no chart junk. Is it possible to add chart junk to our charts? Possibly so. Some people add keys, bulleted information, window boxes of explanations, etc., inside the grid area of Standard Charts. This makes use of the empty space that may well be likely on many standard charts. For convention presentation purposes this may be fine, if it let's one combine talking points with visually-displayed quantitative information. However, for publication purposes, and indeed for certain key points when presenting charts, our best practices may be to reduce or eliminate the amount of this potential chart junk.

Data Density and Aesthetics

The final chapters of Tufte's book cover how much data are shown on a graphic and to what extent one should take artistic sensibilities into account when developing a chart. For data density Tufte offers yet another one of his many ratios. The data density equals the number of entries in a matrix divided by the area of the graphic. The aesthetic issues concern the design of the graphic, the selection of letters and numbers used, the balance and proportion of scale, the professionalism of the drawing and the avoidance of chartjunk decorations. Again, in both cases I regard Tufte's suggestions as strategic considerations for properly communicating information without exaggerating or minimizing claimed effects. For precision teachers who use manufactured Standard Celeration Charts, what Tufte suggests in these chapters may prove only of marginal significance. Precision teachers do not have to draw their own charts.

Regarding Standard Celeration Charting

While Tufte's book may have mixed significance overall for precision teaching, the general suggestions he brings up are instructive. The whole point of having a chart is to communicate information. Moreover, it is to communicate effectively; to induce the reader to make certain conclusions. Toward that end, Tufte's suggestions can assist people who share charts.

For precision teaching the general guidelines covered in Tufte's book can help us clarify what makes for a good chart. Following Tufte's guidelines, a properly developed Standard Celeration Chart will:

1. have dates filled out in the blanks across the top;

2. have the labels filled in across the bottom of the chart;

3. have a pinpoint specified in its appropriate blank;

4. have frequency dots that are dark enough, large enough, and otherwise discriminable;

5. have one or more celeration lines drawn through a series of frequency dots;

6. have a straight-line projection of the celeration;

7. have a celeration line that is also dark and thick enough to be easily seen;

8. not have a whole lot of extraneous text, bulleted points, or other items or markings within the chart frame; and

9. not have either of its axes truncated or laminated in order to "fill the frame."

There are exceptions to some of these suggestions, of course. A chart may have a collection of celerations, and thus in that instance adding in the frequency dots will produce confusion. Likewise, a chart may display a series of conditions, starting with a baseline, and thus there should be both vertical lines separating the different conditions, along with appropriate labels.

Some people may want to make "pinprick" dots, under the belief perhaps that this adds to precision. However, very tiny dots are difficult to see. Furthermore, the Standard Celeration Chart has a multiply divide scale, so having a large dot will not detract from precision. Just center it on its day line and frequency line. Then, recognize that the important difference in frequencies is multiplicative. It matters little if a large dot standing for 12 per minute correct also covers up where 11, 10, and 13, and 14 go.

Another exception concerns an important issue regarding "chartjunk." Not all extraneous marks on a chart should be considered as junk. There are cases where it may be best to reprint an actual working chart, rather than "clean it up" to make it presentable. A student's chart may have doodles, comments, and so on drawn onto it. Publishing it "as is" helps preserve the record of what actually happened and may help illustrate how working charts may look. The real world is not always neat and tidy. A working chart containing such junk had that junk on it while it was being used for making decisions. Therefore, publishing one or more working charts as a record of what one actually did also preserves some honesty about monitoring behavior.

Conclusion

Tufte's book presents a number of issues of strategic importance to the display of quantified information. For our purposes it provides some additional arguments that buttress the importance of using standard charts. The book provides useful counter-arguments to the questionable charting practices that exist in behavior analysis and in education. Moreover, it casts these issues in a colorful, visually interesting volume, one that replicates many unusual charts, some of which are of historical importance.

However, the book is not for everyone. Beginning charters and people new to precision teaching may find the issues either too complex or too irrelevant to their needs. Trying to teach them Tufte's arguments and suggestions may add confusion. Advanced charters and established professionals, however, may find greater value in Tufte's book, especially when faced with countering periodic attacks leveled against the Standard Celeration Chart. Or, they may find value when trying to explain some subtle charting issues to colleagues and to fellow professionals.

Tufte's book primarily concerns published charts. It does not concern itself with what we in precision teaching call working charts. Most of what Tufte deals with has to do with charts that will end up in books, journal articles, or in other archival documents. For those purposes it has many valuable considerations. However, for everyday working charts, which teachers and learners use to make decisions, the points made by Tufte will seem largely superfluous – unless such individuals are tempted to go off and create their own charts. There is a certain tendency of people to want to do this, and if they do, then what Tufte says comes into play at that point.

Tufte's book does not explain how to draw charts and graphs. For that purpose a general source such as Robertson (1988) may prove more valuable. Moreover, Tufte's book does not contain a taxonomy of different chart types, nor coverage of the unique properties of each one of the basic types of charts. For such purposes a book such as Schmid's (1954) may have much more use. Finally, Tufte does not address the specific issue of standard charts, and the concept of standard charts appears not to be a consideration. Thus, the rationale for using standard charts, and how to use them, will require other sources, such as Pennypacker, Koenig, & Lindsley's (1972) Handbook of the Standard Behavior Chart. However, bearing these deficits in mind, all in all Tufte's book would make a valuable addition as a resource for established professionals and for those who have a specific interest in measurement and charting issues and topics.

REFERENCES

Iverson, I. (1988). Tactics of graphic design: A review of Tufte's "The visual display of quantitative information." *Journal of the Experimental Analysis of Behavior, 49*, 171-189.

Pennypacker, H.S., Koenig, C., & Lindsley, O. R. (1972). *Handbook of the standard behavior chart*. Kansas City, Kansas: Precision Media.

Porter, K. (1985). *Standard Celeration Chart Summary of Individual-Percentage-Interval Recording Studies from the Journal of Applied Behavior Analysis 1968-1984*. Doctoral dissertation, University of Kansas.

Robertson, B. (1988). *How to draw charts and diagrams*. Cincinnati, OH: North Light Books.

Schmid. C.F. (1954). *Handbook of graphic presentations*. New York: The Ronald Press Co.

Tufte, E.R. (1983). *The visual display of quantitative information*. Chesire, Connecticut: Graphics Press.