Telling the Story in a Compelling Way:  
Improving Data Visualization by Using Metaphors

Timotheos Frey*

February 2021

Abstract

Data visualization is a powerful tool in information transmission. However, its potential is often partly fulfilled because the design is heavily influenced by the tools rather than a conceptual idea. This article shows the importance of consciously designing data visualization and the power of metaphors to improve data visualization.

1 Visualization Matters

The Purpose of Data Visualization

Graphical representations of data are fast and powerful. Fast, because they are the most efficient way to convey complex information. And powerful, because images stick in people’s minds. Numbers get forgotten, images don’t. Many scholars – from Tufte (1983, 1990, 1997) to Cairo (2013, 2019) – have emphasized the importance of data visualizations in information dissemination as well as the importance of deliberate visualization design.

There are of course researchers and journalists who deliberately seek to mislead the audience, as evidenced by the famous *Daily Express* in the 1950s (see Murray, 2015, for a discussion). But this is another story and not the purpose of data visualization as we are discussing in this article.

This paper intends to show, in a nutshell, how to tell a story clearly, honestly and quickly with improved data visualization – it requires a little effort, but it is worth it.

Good Design Pays Off

With the development of powerful computing systems during the last decades, data visualization became commonplace (see Friendly, 2006 for an overview on the historical development). They are not only ubiquitous, they can be created and distributed increasingly rapidly – in our daily life as well as in science. There are good reasons for this: Kastellec and Leoni (2007), for example, argue for an increased use of graphics instead of tables in the presentation of research results in political science. They put forward two arguments. Compared to tables, data visualization does not suggest a level of precision that

*mailto:timotheos.frey@unifr.ch, http://timfrey.wordpress.com
it cannot give. Second, it constrains researchers to be precise, because with visualization
the overall picture becomes clear to the reader in a quick and comprehensive way.

Unfortunately many data visualization designs are heavily influenced by existing tools
and their templates (see Nussbaumer Knaflic 2015: 2ff.), rather than by a conceptual
idea. As Healy (2019) points out, there are better and worse visualizations. With this
paper I would like to show how data visualization can be better designed. “Better” refers
to a design that minimizes the room for misinterpretation, speeds up the transmission
of information, and makes the information better accessible by presenting it in a form that –
compared to other methods – is easy to understand.

To achieve this goal design matters. First, the viewer needs an immediate overview
and orientation. Second she should be able to grasp the story rapidly without the need
for further explanation. By translating the core message into graphic patterns based on
signs, concepts and shapes already known to the viewer, the transmission of information
can be accelerated and the likelihood of correct interpretation increased.

A very simple example of this is the use of symbols instead of points in a scatterplot
(see fig. 1). The figure shows the political space based on elections exit polls in the United
Kingdom in 2001. While parties are depicted as dots, stars represent the educational
level of respondents, crystals the religious groups, and crossed hammers social classes.
The choice of the crystal was inspired by the religion neutral protective symbol of the
International Red Cross and Red Crescent Movement, and the crossed hammers represent
a historical symbol for mining in particular, and (hard) work in general. The deliberate
choice of distinct symbols symbols are just a detail, but it helps to read the graph.

Xin, Ai and Ai (2018) take a step further. They show that a map-like visualization in
the analysis of abstract non-spatial data provide vivid and concrete features. The map-like
visualization process helps users to make an intuitive comparison and analysis of the data.
However, while there is a broad consensus among designers about what constitutes good
and accessible design, it remains difficult to scientifically show the effect of good design.

Ziemkiewicz and Kosara (2008) suggest that compatibility of visual metaphor and ver-
bal metaphor can speed the processing of visual information. They also provide evidence
that visual metaphor influences the representation of information in the reader’s mind. In

Figure 1: Symbols instead of dots: Representation of the political demand side as a two-
dimensional space in the UK in 2001 (Kriesi et al., 2008, 199)

Xin, Ai and Ai (2018) take a step further. They show that a map-like visualization in
the analysis of abstract non-spatial data provide vivid and concrete features. The map-like
visualization process helps users to make an intuitive comparison and analysis of the data.
However, while there is a broad consensus among designers about what constitutes good
and accessible design, it remains difficult to scientifically show the effect of good design.

Ziemkiewicz and Kosara (2008) suggest that compatibility of visual metaphor and ver-
bal metaphor can speed the processing of visual information. They also provide evidence
that visual metaphor influences the representation of information in the reader’s mind. In
addition, they find signs suggesting that visual metaphors help for a better understanding of visualization.

In practice, this means that conscious design and the use of visual metaphors in data visualization will most likely help to better convey relevant information. And that is why we should do it.

2 Tell Your Story in a Compelling Way

Respect the Data and the Rule of Logic

First the data should match with the visualization project (see Healey, 2019; Ryan, 2016, 139 ff.). Is the data fit for purpose? The aggregation level and the measurement level, the handling of missing data and other features of the data should be carefully examined to choose an appropriate form of visualization.

Second, hierarchical orders and causalities must always be respected and shown. A timeline is a directed vector, a bigger item should be depicted bigger than a smaller one. An important element should be highlighted in order to provide orientation to the viewer. Causal effects have a direction, which must also be defined visually. In short: show the world as it is, otherwise the best design is useless.

Use Metaphors as Shortcuts

Metaphors help us understanding our environment, they are thus important for poets as well as for designers. The shape and appearance of an object often also conveys its purpose. This works because our brain associates similar shapes and derives the function from them. This way, many everyday objects become almost self-explanatory (Hekkert and Cila, 2015).

One of the best-known example is the analog radio: nowadays rarely encountered, but essential design elements have remained with us to this day as metaphors. The reason for this can be found in their immediate logic. The so-called radio buttons have been transferred to areas of application in which they no longer represent a solution to an immediate technical need, but rather serve as a familiar metaphor for a selector switch.

Design is a Conscious Process

It is very worthwhile to try out different design options. Often, the first idea is not the best. In many cases, the optimal solution does not result from the concept, but from the entire process.

A few years ago, my colleague Simon Bornschier and I developed a series of design options to place voters and parties on political scales. The starting point was the question of how to contrast political supply with demand, i.e., how to graphically display and compare party positions with voter preferences on a left/right axis. Based on individual voter survey data and media data on political supply, two continuous variables that can be compared were formed. Figure 2 on page 5 shows six design variants. All are based on the same data and the same concept, mapping the left/right axis horizontally, with a zero point at the center. Left parties and voters are placed on the left, right parties on
the right. Using the same data the individual graphs are very different, both in terms of design and message.

All these variants show basically the same: the political space – defined by the classical left-right scale – of France during the 1995 parliamentary elections. Some map the parties’ program – the political supply side – on the axis, while others also map the voters’ position – the political demand side. The data used for the positioning of the parties are, as mentioned above, aggregated position statements of party exponents in the media. The positioning of the voters is based on survey data (see Kriesi et al., 2008 for more information on the data and the methodology).

Graphs (a) and (f) depict only the policy supply, all others show both the supply and the demand side. The graphs showing only the policy supply display not only the mean but also the dispersion of the observations, for (a) as a transparent box and for (f) as a fine line. Both plots make overlaps among the individual parties visible.

The design (c) was one of the first attempts. As with (d), the scale metaphor is clearly recognizable, making the immediate measurability of the data palpable. Only one scale is shown, which might create an issue: this design can lead to the erroneous assumption that same data is used for both the positioning of parties and the positioning of voters. This is not the case. Variant (b) would be a possible answer to this potential confusion, but it becomes – due to the higher number of elements – more complex and difficult to read.

The (e) variant is the author’s personal favorite. It shows discrepancies between the party base and the party leadership. The connection between the party and its electorate is drawn as a fine, dashed line - which is indeed a facet of reality that can also be shown graphically. The scale values were deliberately omitted because direct comparability is limited due to the different measurement methods used to measure supply and demand. On the other hand, this illustration clearly shows that the parties at the poles of the spectrum orient themselves to their extreme positions in the election campaign and thus distinguish themselves from their opponents. It is particularly striking that the voters of the Communist Party (PCF) are on average more moderate than the voters of the Socialists (PSF) – the “converted positions” immediately comes to mind thanks to the crossing lines.

In the end, the variant (f) was used (Bornschier, 2010), whereby the diagram was drawn twice, once for the supply side and once for the demand side.

This example highlights three points: First, there is not only one valid solution. Second, a real design process is useful. Third, quite different stories can be told with the same data.

3 Transforming a Concept Into a Design

Getting a Clear Picture of the Story

Depending on the purpose, some forms of visualization are more suitable than others. The size ratios of three or four groups can be well shown using a pie chart. But if there are more groups or if additional information needs to be included, a bar chart or a bubble chart is probably more suitable. If the groups are parliamentary groups, then a form
Figure 2: Mapping party positions and voter preferences: Six design variants for left/right axis compared

Voters and parties: demand and supply side compared
France 1995: cultural dimension
of presentation that is oriented to the seating arrangement in parliament is preferable because it awakens associations in the viewer’s mind. She immediately knows what it is about without the need of additional information.

As this example shows, good data visualization provides contextual information graphically. We must therefore always ask ourselves three questions at the beginning of a project: First, what story do we want to tell? Second, how can the basic concepts behind the story be shown graphically? And third, whether the data, the story and the graphic are coherent and consistent.

Making the Underlying Concept Visible

Using the example of commercial media analysis – arguably currently one of the most common data visualizations as every major company and organization closely monitors its media presence – it is easy to show how visualizations can be improved with little effort and a metaphor.

First, the story: Organizations strive for visibility and brand awareness as well as a good reputation. Taking into account the mechanisms behind perception building, we can assume that structural effects play a strong role, while cyclical effects are rather weak, i.e. extraordinary successes or isolated pieces of bad news have little impact on the overall perception. On the other hand, rumors and false stories can cause serious reputational damage if they are spread over a long period of time.

Second, the basic concept that should be shown graphically: Usually the results of such kind of media analysis are measured as frequencies over time, sometimes weighted by channel reach. They tend to be presented as bar charts or line charts, where the time is mapped on the X-axis. Unfortunately, peaks, as they occur in extraordinary situations, always catch the eye first – independently of their real impact. Considering the presumed structural effects, we should not be interested in peaks, but rather in the publicly perceivable information density. Therefore, the crucial question is: How can density be represented?

Density can be represented, for example, with a bubble chart, where the bubbles are semi-transparent and the bubble size represents the aggregated reach per day. This way a single large bubble appears less dense than a series of overlapping bubbles, even if the latter are smaller. Figure 3 — showing a comparison between the neutral and positive media coverage and the negative coverage of a particular event in a crisis situation — provides a very good example.

Third, the examination of coherence and consistence: The available data can tell this story and allows for an easy to create meaningful chart. While the selected chart is not as commonplace as a bar chart, its complexity is no greater than that of comparable chart types. The strength of this type of diagram is that it shows the mechanism behind the formation of public opinion – the publicly perceivable information density. At the same time, short-term peaks are discriminated, which is consistent with the underlying theory. This form of representation thus helps the viewer to better understand the story.
Figure 3: Example of a graphic implementation of the concept of the *publicly perceivable information density*. This graphic was created as part of a commercial project and is therefore anonymized.

4 Discussion

Data visualization is a powerful tool. The use of metaphors can improve data visualization because they help the viewer to better understand the chart. There is evidence that data visualizations, as well as the use of metaphors, are effective and efficient means to accelerate the transmission of information. However, there is still a need for more research looking at the cognitive processes involved.

Too often the design of data visualization is left to the tools. This is unfortunate. The ultimate objective of authors is to tell their story in meaningful way for their audience. But the way the story is told matters. The examples provided in this paper show that a lot can be achieved with simple means.
References


