
Seminar review: *Presenting data and information* by Edward Tufte, November 9, 2016, Arlington, VA

Boris Veytsman

Edward Tufte, a prominent advocate of beauty and clarity in data presentation, is rather well known to the T_EX community. There are several packages on CTAN inspired by his ideas, for example, *tufte-latex* and *sparklines*. Tufte’s essay about the cognitive style of PowerPoint (2003) was quoted by many T_EXnicians at that time (I must say, however, that we now have T_EX tools quite capable of reproducing PowerPoint style). His web site (edwardtufte.com) and Twitter (@EdwardTufte) have been sources of many interesting ideas and beautiful pictures.

For quite some time, I have wanted to attend his seminar (edwardtufte.com/tufte/courses), and this November my employer (Harris Corp.) kindly agreed to pay for it. (Disclaimer: this review reflects only my personal opinion, and does not necessarily reflect opinions of my employer, other Harris employees or officers. This review cannot be construed as an endorsement or recommendation by Harris or any other corporate entity.)

Despite a rather steep price (\$420.00, which includes four books by Tufte in paperback editions, coffee & tea, but does not include parking or lunch), the big ballroom of the Hyatt at DCA airport was packed. This is a good sign: there *is* an interest in high quality data presentation by the public.

The presentation itself, as a matter of course, followed Tufte’s recommendations. It started with an hour of silent reading, it emphatically did not use slides with bullets, and it did not assume that the attention span of the audience was close to that of a two year old. Many of the topics covered would be familiar to those who have read Tufte’s books: from the famous map by Charles Joseph Minard to the beautiful examples of the “small multiples” graphs. There were also many new topics, including a very interesting analysis of the styles of Web sites. Despite my previous knowledge of Tufte’s ideas, the presentation turned out useful for me: it is one thing to read about the general principles, but quite another thing is to see them in action.

Edward Tufte did not limit himself to typography or graphic design. Rather, he showed a holistic approach to the presentation of data and information. He talked about such things as preparation before the presentation, ways to deal with anxiety and technology malfunction, and many other topics that are important for a successful practitioner, but often overlooked.

Boris Veytsman

Tufte discussed both sides of a presentation: the point of view of a presenter, and the point of view of an audience. There was plenty of advice on both how to show information and how to understand what was shown — and recognize when the presenter is not forthright with the truth.

This leads to another dimension of the seminar, one which was rarely mentioned explicitly, but definitely was present: the moral dimension. Tufte explained how to show the story behind the data, always assuming this was a true story. It was not only that he did not teach how to mislead the audience and conceal the truth; rather, his methods are not compatible with such subterfuge (more on this below). Data presentation as practiced by Tufte is not a neutral trade, which can be used for any goal, being it telling the truth or telling lies. Honesty is hardwired in the approach.

The seminar used projections, videos and music. The current jargon suggests the word “multimedia”. However, this word now often means some haphazardly selected fragments joined together in a more or less arbitrary fashion. Unlike those “multimedia presentations”, this one embedded the non-speech elements with tact and taste, creating an organic whole. Perhaps the closest analogy would be a theater play set by a good director, rehearsed to perfection and skillfully executed. Indeed, Tufte’s presentations are one person shows. It is strange that among the arts discussed by the author (typography, painting, music, literature) the thespian art was not mentioned — it would seem to be very relevant to the ideas espoused in the seminar.

While it is impossible to write here about all the topics raised by Tufte, I would like to mention several of them, which, in my opinion, would be of a special interest to the readership of *TUGboat* and the T_EX community.

Tufte spent much time advocating high resolution of information. One of his villains, mentioned several times during the course, was the famous “rule of seven \pm two”, which prescribes no more than about seven chunky sentences on a presentation page. Tufte thinks that audience can grasp complex ideas and flowing texts. He says that the best example of a high resolution information medium is a classic printed page. This strikes a chord with us T_EXnicians, brought up in the traditions of classical typography as embodied by T_EX design.

Tufte brought with him a couple of rare books to demonstrate these traditions and inherent innovation: one by Galileo (*Istoria e dimostrazioni intorno alle macchie solari*, Rome, 1613) with small images embedded in the text, and a second by Oliver Byrne

(*The Elements of Euclid*, London, 1847) with 3D cutouts of stereometric bodies. For some bibliophiles the opportunity to see these books alone might well justify the cost of admission.

Another important topic of the seminar was providing the audience of a presentation multiple paths through the data. The common low resolution presentation assumes a single path followed by the audience, bullets being stepping stones across the sea of information. The high resolution presentation espoused by Tufte assumes the audience members may choose different paths through the data and reach their own conclusions, perhaps unforeseen by the presenter. The recommended way of starting the presentation with a period of silent reading actively encourages this approach.

There are two aspects to this approach. First, it seems to be surprisingly close to the Free Software philosophy. One of the fundamental purposes of Free Software is to give anyone the ability to modify and combine programs in new ways, sometimes not foreseen by the original authors. Tufte applies this principle to data. Like Free Software users, his audience is expected to be active participants rather than passive consumers fed with pre-digested pieces of information. It is not coincidental that Tufte made highly complimentary remarks about the reproducible research movement, which is an analog of Free Software in the data world.

Second, as mentioned above, this approach is inherently incompatible with misleading and manipulation. A manipulator wants the audience to reach the pre-defined conclusions, and carefully constructs the path to them, avoiding anything which might accidentally reveal the truth. The last thing a manipulator wants is to let the audience engage in an independent analysis. While reproducible research and high resolution presentation were not specifically designed to prevent data manipulation, it is an inherent feature of these approaches. Similarly, while Free Software was not specifically designed to prevent backdoors in software, backdoors are not compatible with it.

It should be said that Free Software was prominent in the recommendations spelled out in the course handouts. Besides T_EX, Tufte mentioned *R*, which is used by many T_EXnicians. By the way, there are several *R* packages inspired by Tufte: for example, *ggthemes* includes a Tufte theme and Tufte plots such as range frames are part of the *ggplot2* framework. (There is a detailed review of these packages by Lukasz Piwek at motioninsocial.com/tufte.)

Still, there was a topic of the seminar which left a mixed aftertaste for me. Namely, Tufte several times exhorted the audience to produce high quality presentations. This was done in a way that seemed to imply that the common low quality presentations are caused by the laziness of the practitioners. This is definitely not the complete case. The high quality display Tufte espouses requires a huge amount of thought. While they are immensely satisfying for the producer and very useful to the audience, they require a most valuable resource: the work of an extremely skilled professional. Therefore they are insanely expensive to make. In many corporate environments, where an incessant stream of (necessarily low quality) presentations is expected, Tufte's ideas are a losing proposition. In a sense, Tufte's exhortations in the Hyatt ballroom could be taken as similar to a motivation speech by a renowned *haute cuisine* chef to an audience of fast food cooks. Of course these cooks would be happy to work on something better than greasy burgers. However, they are not in the position to decide: the menu is determined by their managers, and ultimately by the customers. If we want to change the bleak landscape of corporate presentations, we need to change corporate culture, including the expectations of upper management and individual choices of the information producers and consumers. It seems this would also imply rather fundamental changes in our society. At any rate, shaming the overworked practitioners, in my opinion, would not help.

Nonetheless, there is a role for us T_EXnicians, especially those who are in the business of writing packages. Obviously our tools cannot do the creative thinking for the presentation producer. However, there is much mundane work, which we can and should make much easier, leaving the practitioner free to be creative. Many common tools today subtly (or not so subtly) nudge the users in the direction of making poor presentations. It is possible (albeit difficult) to make tools nudging the users to make better ones.

A seminar like this one could be an inspiration and a source of ideas for this effort. Thus I would recommend it for TUG members.

◇ Boris Veytsman
Systems Biology School and
Computational Materials
Science Center, MS 6A2,
George Mason University,
Fairfax, VA 22030
`borisv (at) lk dot net`
<http://borisv.lk.net>