Commentary on 'Improving Graphic Displays by Controlling Creativity'

Linda Williams Pickle

Linda Williams Pickle for the past 30 years has been focused on developing better statistical methods and data visualization tools for analyzing and presenting health-related data. Her research has led to improved techniques of modeling incidence rates at the county level, taking into account hierarchies and spatial relationships. Simple comparisons or regression analyses that do not take into account these hierarchies yield incorrect results. This method has been adopted by the American Cancer Society for their *Cancer Facts and Figures* annual publication.

Before our work, the National Center for Health Statistics (NCHS) processed and published death certificate data collected by all states, but it never published maps based on these data. Manning Feinleib, NCHS director at the time, decided the center should produce its own atlas and asked me to lead the project.

The challenge at NCHS was not to control rampant creativity, but to encourage some degree of it. NCHS has a long history of producing reports of health statistics for the United States. Staff statisticians take their role of guarding the integrity of the data very seriously. They, therefore, preferred to publish tables from which questions could be accurately answered, rather than graphics from which only an approximation of the true value could be visually extracted. Similarly, they preferred to keep the formats of the tables constant, year after year, so that, as Wainer points out, the user need not learn a new format each year. Thus, the challenge for the *Atlas* team was to overcome the complaint "… But we've never done it that way before."

Most agreed that the atlases we had previously published at the National Cancer Institute (NCI) demonstrated the value of looking for geographic patterns in mortality rates. They were hesitant, however, to try any design deviating from the design that had been used before (i.e., a red-blue color scheme, rates categorized by a combination of rank and statistical significance, and most areas with nonsignificant rates blanked out).

We asked five visualization experts how we could improve upon the NCI design and got five opinions, so we began a research program at NCHS to test the effectiveness of each map design element. The rigorous experimental design of these studies helped us move from a design based on personal preference to one more scientifically grounded. Our efforts paid off because they resulted in approval of the unique overall design by upper-level management and an award for best illustrated government book in 1997.

NCI Cancer Trends Progress Report

The design for the NCI Cancer Trends Progress Report was developed in a different way. For that report, an advisory team consisting of NCI clinicians, epidemiologists, social scientists, and statisticians debated and proposed design elements for the entire report. An external advisory committee reviewed the initial proposal and recommended changes. Time limits for publication and the desired high degree of automation of graph production limited the customization allowed for the many graphs. As the graphical design advisor, I needed to make compromises for scale, placement of tick marks and grid lines, etc., that would make sense for all the graphs but would be consistent throughout the report. All advisors agreed consistency was important for readability.

Unlike the *Atlas* maps of mortality rates, this report included graphs for various types of variables. As an example of the issues that arose when considering design options, we had to consider questions such as the following: Should the graph for the average daily number of ounces of red meat consumed be a standard square format with vertical axis beginning at zero, or should its range be truncated and/or its aspect ratio modified to better highlight the slight downward time trend of consumption? The resulting "standard" design for the report was obviously not optimal for all graphs, as Wainer points out. His reformatted graph is certainly an improvement over the original. However, because this report would be updated frequently on the web as new data became available, the graphics production system needed to be as automated as possible, precluding optimization of every graph.

Focus Influences Graphic Style

I agree with Wainer that one can improve graphic displays by controlling creativity, but wish to point out that this works both ways. Groups that are more focused on the data than the presentation will often resist change and need to be convinced that the judicious use of creative graphics can help get an accurate message across to the reader. On the other hand, those who focus more on the presentation than the actual data may be overly creative, interfering with the reader's ability to clearly see patterns in the data (Tufte's concept of "chart junk" comes to mind). Government agencies, with their typical multiple layers of approval required for publication, tend to incorporate too little creativity into their graphical designs, whereas academics, more interested in methods than data someone else collected, tend to be overly creative. The best graphics, in my opinion, are those that find a middle ground and push the envelope of design while retaining those elements found to work well by cognitive studies or experience.

Further Reading

Pickle, L.W., and Herrmann, D.J. (eds.) (1995). "Cognitive Aspects of Statistical Mapping." NCHS Working Paper Series Report, No. 18.

Pickle, L.W., and Herrmann, D.J. (1999). "Cognitive Research for the Design of Statistical Rate Maps." ASA Proceedings, Section on Survey Research Methods, 186–191. Alexandria, Virginia: American Statistical Association. *www.amstat.org/sections/SRMS/proceedings/papers/1999_029.pdf*.