Realism and Perceptions of Data Quality in Computer-Displayed Maps

Tony Boughman and Sara Irina Fabrikant

Department of Geography

University of California, Santa Barbara

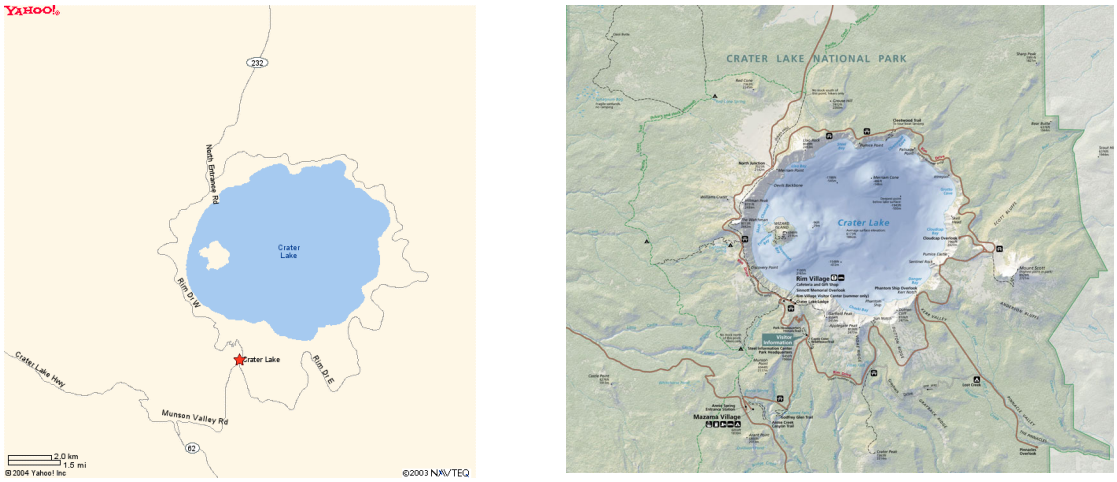
Santa Barbara, CA 93106

boughman@geog.ucsb.edu

Issues in data quality are of particular interest as we continue to accumulate huge quantities of digital information. Creators of geographic information are familiar with the inherent shortcomings in translating real-world observations into digital representations, but do users understand the limitations of data? Cartographers have long been aware of the necessity of presenting an abstract view of the infinitely complex world. Improvements in computer technology allow increased realism in digital images and there seems to be a general movement toward more realistic display (Patterson, 2002). This movement is expressed in plan-view, orthographic maps as well as 3-D visualizations, video games, and virtual environments. More realistic visual rendering does not mean that the underlying data are closer to reality, or more accurate and precise (Longley et al., 2001). Researchers are currently looking into the idea of communicating the degree of uncertainty in digital data to consumers by somehow presenting or visualizing the uncertainty (Buttenfield, 1993; McGranaghan, 1993; Goodchild &

Clarke, 2002). Is the trend toward more realism doing the opposite and conveying, perhaps inadvertently, the impression of more certainty in the data?

This paper presents an experimental design to examine if people infer a higher level of data quality, accuracy, and precision from a more detailed, realistic-looking map compared to a more generalized and abstract map. The two map types used for comparison have roads as the primary theme and are designed for locational reference or navigation (Figure 1). Both maps will have graphical scalebars. The more abstract map is a vector graphic with generalized features, simple colors, and little added contextual information. It is modeled on the typical route-finding maps available on the World Wide Web (Figure 1a).



a) abstract, generalized map[1] b) detailed, realistic-looking map[2]

Figure 1: Two map styles used as inspiration for the design of abstract and realistic-looking test stimuli

[1] http://maps.yahoo.com

[2] http://www.shadedrelief.com/realism/crater.jpg

The more realistic-looking map combines both raster and vector graphics and has additional contextual information (water features, trails, railroads, buildings), topographical additions (natural tone coloring, hill shading, land cover texture), and greater geometric detail with less generalization (Figure 1b). Generalization is a method of data compression commonly used in cartography when reducing scale (McMaster and Shea, 1992). A reduction in map scale, as when zooming out with a computer display, has a concomitant increase in area covered by the map. To keep the volume of data constant the map is generalized. When generalization is used without a change of scale the map appears simpler and less detailed. The stimuli in this study are designed to examine if additions of context and topography along with more detailed geometry lead map viewers to have an impression of higher data quality in terms of locational accuracy and precision.

With the ascendance of the World Wide Web, the internet has surpassed printing as the leading medium of map delivery, therefore this study uses as its stimuli digital maps displayed on computer monitors (Petersen, 1997). Detail is limited by computer monitor resolution, typically 72 pixels per inch, while normal human visual acuity is much greater than this (Clark, 2002). When viewing maps on computer displays or websites with interactive capabilities more detail is often enabled by zooming in to a larger scale but this study only looks at statically displayed maps at a fixed scale.

This study is aimed at non-expert map users who most likely have no training in digital map creation. Participants drawn from the research pool of undergraduate students at the University of California, Santa Barbara are asked to make accuracy and precision judgments while performing common map use tasks. The maps used here are typical of those used for locational reference or navigation. Participants are randomly assigned to one of two maps and asked to estimate distances between objects and to estimate travel distances along the road network. They are asked to place confidence intervals around their estimates. The confidence intervals expressed as plus or minus units are used as a basis of comparison between maps. Participants are also asked to compare distances between pairs of objects on the map. They are asked, for example, "Is the distance greater from point A to point B or from point A to point C?" and given the option of "I don't know". This is another method used to establish how precisely distances are estimated. The differences in distance among pairs of objects will vary, with some pairs having an obvious difference (A to B is much farther than A to C). Other pairs will have less difference in distance, all the way down to A to B being equal in distance as A to C. In this way a threshold of where the "I don't know" response begins can be found. Similarly, questions regarding the road network are something like "Would it take longer to travel by car from A to B or from B to C?" Again, they are given the option of "I don't know". The variation in distance estimations along with the magnitude of plus or minus confidence intervals are compared between maps. The point-to-point

distance comparisons are analyzed for the level of difference in distances at which uncertainty in the form of “I don’t know” responses appears.

The results of this study provide insight into the effects of cartographic realism and detail on people’s inferences of the underlying quality of the data. A relationship found here could then be employed in visualizing geographic data with regard to the conveyance of uncertainty. It may be feasible to use realism or level of generalization as visual variables in situations that involve communicating the degree of data quality.

Acknowledgements

This work is supported by the National Science Foundation’s Integrative Graduate Training Program in Interactive Digital Media (IGERT), Award Number: DGE-0221713

References

Buttenfield, B.P. (1993). Representing Data Quality. *Cartographica*, Vol. 30, No. 2-3.

Clark, R.N. (2003). Experiments with Pixels Per Inch (PPI) on Printed Image Sharpness. Clarkvision.com, http://www.clarkvision.com/imagedetail/printer-ppi

Goodchild, M.F., and Clarke, K.C. (2002). Data Quality in Massive Data Sets. In Abello, J. et al. (eds.), *Handbook of Massive Data Sets*, 643—659. Netherlands, Kluwer Academic Publishers.

Longley, P. A., Goodchild, M.F., Maguire, D. J., and Rhind, D. W. (2001). Geographic Information Systems and Science, Sussex, England, John Wiley & Sons Ltd.

McGranaghan, M. (1993). A Cartographic View of Spatial Data Quality. *Cartographica*, Vol. 30, No. 2-3.

McMaster, R.B. and Shea, S. (1992). *Generalization in Digital Cartography*. Washington, D.C., Association of American Geographers

Patterson, T. (2002). Getting Real: Reflecting on the National Park Service Maps. Presented at: International Cartographic Association (ICA), Mountain Cartography Workshop, Timberline Lodge, Mt. Hood, Oregon http://www.shadedrelief.com/realism/

Petersen, M.P. (1997). Trends in Internet Map Use. Proceedings of the 18th International Cartographic Conference 1997, pp. 1635-1642.