Geospatial Visualization Software
Impediments To Creativity and Progress

Daniel B. Carr
Department of Applied and Engineering Statistics
George Mason University, email: dcarr@gmu.edu

1. Introduction

The context of the paper is the desire to promote creative progress in geospatial visualization software. I suspect that most of us who attempt to create new methodology want to make a difference that we can see in our lifetime. Thus the notion of progress here is the degree to which the new methodology is moving toward being used by the targeted audience. The desire to see new methodology put into practice motivates noting impediments to progress and thinking about strategies to deal with the impediments. This paper notes impediments in the hope of promoting some discussion about strategies and solutions. I apologize in advance for the informal writing style, the typos to which I am prone and for writing so much about my own experience rather than providing a scholarly work in which impediments have been scientifically established to be widespread. I hope the impediment are neither to mundane nor to hopeless and that useful discussion will follow.

The paper is also motivated by The National Research Council's recent publication, *Beyond Productivity Information Technology, Innovation and Creativity* (BP for short). BP says that "Individuals and groups involved with information technology and creative practices (ITCP) benefit from participating in venues that support, motivate and display this type of work."

The publication goes on to describe three classes of modern studio-laboratories. A time diagram lists major 20th century centers. I only recognize a few of these stand alone centers, such as Bell Labs, MIT media lab and PARC-Pair. I strongly remember my 1984 visit to Bell Labs. One of my graphs sparked a discussion between John Tukey and Colin Mallows that I could not follow. This visit and subsequent visits provided inspiration. I always left with thinking that even I could do wonderful things if guided and encouraged by the brilliant statistics staff. Inspiration is important

The Bell Labs of today is far different from what I remember. BP addresses longevity issues for centers embracing technology and creativity.

BP also covers hybrid networks that cross institutions and virtual-space-based strategies. This immediately brings to mind the groups to which I am a member and those that are drawing me from awareness toward involvement. How might I link to new groups and how might the links among my favorite groups be strengthened? Can virtual collaborative research environments really work? Can their scope extend to strategically address a variety of impediments?

2. Impediments

I work to develop new visualization methodology and delight in interacting with researchers of vision. They provide the context and direction for the development of ideas. Thus I have a reasonable strategy for developing ideas centered on needs that have been conceptualized by key people. However I do not have the expertise and energy to adequately promote the transitions from ideas to

- software development
- software evaluation and refinement;
- papers,
- recognition as being important for deployment,
- community understanding, and
- distribution and support.

I manage to develop prototype software, increasingly with the help of students. My coauthors help me to get a fair number of papers published. That seem to be where things most things end. I think I once read that the average readership for a journal article is 0.8. From this I infer that even author(s) do not read their published articles. Sadly, my most cited paper seems to be reference more to indicate awareness than actual use. Whether the 0.8 statistic is correct or not, having a paper published is only a tiny step toward progress. The following is about the bulleted items other than papers .

2.1. Software development, recognition, distribution and support

My thoughts are drawn to friends and software development teams in statistical software companies such as Insightful, SPSS, and SAS. I have seen enough in professional software **design** at one of these companies to wonder how in the world a professor working in isolation can compete in developing visualization software. Then there are the company coding teams, testing teams, professional marketers, established sales and distribution channels, and software support teams.

With recognition, it is possible to get new methodology in commercial products. Professional recognition can sometimes get a foot in the door. The key recognition to achieve is that selling the new methodology makes commercial sense. This can take some serious salesmanship even when well justified. The fact the new methodology has taken hold in a high reputation collaborative research environment could be a big step toward recognition.

Open source software offers another venue. At the moment two groups draw my attention: the R group and the GeoVista Studio group at Penn State. These groups would seem to provide foundations for virtual collaborative environments that could be extended to include more capabilities.

In terms of software development tools, the absence of really convenient widget sets for GUI's is a major impediment. It is not that hard to link visualization software to databases or to a statistics engine like R. Rendering software like OpenGL is straight forward to use. Developing good interfaces poses many hard challenges. Of course part of the problem is how the widgets should best work together to serving the goals.

2.2 Software evaluation and refinement

.

Wood (1992) says "all graphics have an agenda." The possible agenda for quantitative graphics include software sales, deception, distraction, ritualistic participation, entertainment, description, summarization, discovery, story telling, and decision making. When the design focus is constructive (knowledge oriented) scientific visualization it is desirable to scientifically assess the designs in terms of meeting conjectured objectives.

Two common impediments to the development and promotion of software tools are 1) the lack of access to usability assessment resources and 2) the lack of access to the desired subjects.

I was fortunate to work at NCI and help in the development of the state cancer profiles web site (www.statecancerprofiles.cancer.gov). The direct encounter with usability assessment was a positive experience. It seemed so right to have usability groups within the federal agencies the disseminate statistics (often geospatial) to a variety of audiences. Unfortunately the NCI group is shrinking and my access professional usability assessment is going away.

As a creative graphics designer I often want to go against convention. This, of course, is in violation of the famous graphics design principle, "follow convention". Ware (2000) put his finger on my dilemma when he discussed the distinction between sensory and arbitrary symbols. He defines sensory symbols as deriving their expressive power from their ability for using the perceptual processing power of the brain without learning. He defines arbitrary symbols as being arbitrarily derived, learned and predominantly culture based. While specific cases may not be purely sensory or arbitrary, the distinction is useful. Graphic design should help to exploit the power of our eye-brain systems. It seems desirable to favor sensory symbols, layouts and interfaces whenever there are no well-established cultural conventions. I want to go further and attack some moderately well-established conventions hoping that education and time will change the conventions that are holding us back. However, before I can do this I need usability assessment/cognitive tests to establish that there are merits to my presumed sensory symbols.

Consider an example. There is a dot plot with perceptual groups of five. I want to connect the dots in groups of five to help guide the visual tracking from dot to dot. I think that connecting the dots leads to similar eye traversal in repeated viewing, to faster memorization of the dot pattern, and to creating a group-of-five symbol that can be more easily compare to other group-of-five symbols.

People inform me that connecting the dots violates a convention. The most common rational for the convention seems to be that some people (not those explaining this to me) might interpret the connection as implying interpolation. I imagine a dot plot with just three connected dots indicating mortality rates for California, Montana, and South Carolina. I have trouble imagining a person that would seriously interpolate the values. In my mental model of the world people that are smart enough to interpolate in their heads would know better. The only reasons I can think for not connect dots is to reduce increase the data ink to total ink ratio Tufte (1983) and to reduce clutter. Contrarily I think judicious addition of lines often simplifies appearance. I need evidence to support my conjectures before braving a cultural battle.

There can be huge differences in audiences. I suspect that lot could be learn by usability studies specifically designed to benefit from the input of sophisticated researchers. Access to such subjects is limited.

2.3 Understanding and education

I heard that when box plots were introduced at the food and drug administration, researchers meet them with considerable resistance. Now box plots are taught in grade school. The propagation of new paradigms in institutions meets with resistance. Substantial change involves substantial education.

I am hoping to promote change by making software available called conditioned choropleth maps (downloadable from www.galaxy.gmu.edu/~dcarr/ccmaps.) While a simplified version may be very useful in grade school my first targeted audience is older. Education may not be so easy.

There many things about the software that the targeted audience may not discover on its own. One, for example, involves figuring out how to study estimate uncertainty for values symbolized on a map. To address this in CCmaps one can attach either standard errors or confidence interval differences (upper bound – lower bound) to a 3 class slider and dynamically change the thresholds for partitioning cases such as counties in three classes. The three classes are highlighted separately in three juxtaposed maps. It is simple look for the spatial patterns of estimate uncertainty while keeping the estimates in view.

I see multimedia components as good way to provide tutorials about software features, exploration strategies and interpretation. The research on the merits of utilizing both auditory and visual memory seems clear (Meyer 2001).

I wish for better tools for implementing educational components. However the big impediment is my lack of talent and skill for preparing well focus diagrams, well written scripts and cogent dynamic tutorials. A collaborative research consortium that provides access to people with these skills would really draw my attention.

3. Closing remarks

A list of impediments can go on and on. Having interesting data in hand helps me to think about new designs and to promote the software showing it. There are many impediments to obtaining data not the least of which is the failure to collect data. Currently I am frustrated by the lack of good data on cigarette smoking rates (percents in different packs per day classes) at the county level. Other impediments include confidentiality rules that often prohibit different federal agencies from sharing data. Even when data is available there are often web access impediments. One common impediment involves data access designed to provide small isolated elements rather than substantial data sets. A second common impediment is the failure to format data in convenient form for input to analysis software. Moving on to the topic of metadata leads to more thought about impediments.

I have chosen to be part of collaborative research communities that provide both guidance and data. These communities have been crucial to my continued learning, productivity and creativity. These are virtual communities in the sense that the locations are distant and interaction depends substantially on electronic communication. They are not pure virtual because they include many friends who find ways to meet together. Overlapping research interests, respect for talent and skill, and compatible personalities make it work. Reading BP leads me to imaging virtual collaborative research communities that provide an even wider range of professional talent and tools. Perhaps the workshop will consider such communities as a topic.

References.

Committee on Information Technology and Creativity, 2003. W.J Mitchel, A.S. Inouye and M.S Blumenthal Eds. *Beyond Productivity, Information Technology, Innovation, and Creativity*. The National Academies Press, Washington D. C.

Meyer, Richard 2001. *Multimedia Learning*. Cambridge University Press, New York,NY

Tufte, Edward, 1983, *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, Connecticut.

Ware, Colin. 2000. *Information Visualization Perception for Design*. Academic Press, San Diego CA.

Wood, Denis. *1992. The Power of Maps*. Gilford Press New York, NY