Tversky, B., Morrison, J. B. and Bétrancourt M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies*, 57, 247–262.
Winn, W. (1993). Perception principles. In M. Fleming and W. H. Levie (Eds.), *Instructional message design: Principles from the behavioural and cognitive sciences* (pp. 55–126). New Jersey: Educational Technology.
Wouters, P., Paas, F. and van Merriënboer, J. G. (2008). How to optimise learning from animated models: A review of guidelines based on cognitive load. *Review of Educational Research*, 78, 645–675.

35 Toward Empirically Verified Cartographic Displays

Clare Davies, Sara Irina Fabrikant, and Mary Hegarty

Introduction

Maps, like many other kinds of purposefully constructed graphic displays, are so familiar in everyday life that it is easy to overlook the complex decisions made by the cartographers who design them. Although cartography has a long history as both art and science, there has been very little empirical work on the scientific evaluation of map designs. This chapter summarizes current knowledge and research on perceptual and attentional aspects of cartographic design, and empirical methods that can be used to evaluate the design of both static maps and more flexible, and dynamic, computer-based geographic visualizations. We focus on the use of "salience maps" to characterize the map stimuli, and the use of eye fixations to measure visual attention. We also summarize some alternative empirical methods.

All of these methods can be used to evaluate not only the more common static map displays such as road maps and topographic maps, but also the more realistic-looking geographic visualizations used in remote sensing and photogrammetry (the extraction of measurements, physical topography, and other information from aerial photographs or satellite imagery). Furthermore, the evaluation methods can also be used with more abstract cartographic products, such as statistical maps and multivariate spatiotemporal displays, including map animations and three-dimensional-globe viewers. As such, our discussion pertains not only to conventional cartographic designers but also to other visual design disciplines.

Cartographic Design

Like many areas of visual design, cartography – the making of maps – was traditionally viewed as a skilled craft but after the Enlightenment gradually became an academic discipline, the study and creation of "products of art clarified by science" (Eckert 1908/1977). One might imagine, then, that scientific evaluation of map designs has been a critical part of the cartographic process. However, this is not the case. One of the best known cartographic writers of the twentieth century, Arthur Robinson (1952), bemoaned the general lack of empirical evaluation by cartographers of the time, pointing out that mere self-testing by the cartographer was inadequate since "the quality of parental pride is not always objective" (Robinson, 1952, p. 10).

There are many structured design decisions that a cartographer must make when constructing a map. Principles of the cartographic design process can be found in many of the widely used cartography textbooks (see, for example, Dent, 1999; Kraak and Ormeling, 2010; Slocum et al., 2010). The design process typically starts with a *conceptual design* phase in which the fundamental operation of thematic or object generalization is applied. The cartographer selects, classifies, and restructures the spatial data to be displayed on the basis of the intended theme and purpose of the map, also considering user requirements as well as usage and technical constraints.

The subsequent *cartographic generalization and design* phase typically entails graphic symbolization, where the cartographic designer systematically assigns graphic marks to the restructured data, again always keeping the map's theme, purpose, and target audience in mind. In addition, cartographic generalization often includes simplifying clutter, for example, by reducing the number of bends shown in a very twisty road, replacing multiple buildings in a small-scale map with a uniformly colored polygon implying an urban area, or using a specific symbol to indicate multiple closely located instances of a phenomenon.

Some degree of subjectivity is always implicit in maps designed by human cartographers. Although this is often unnoticed by map users, it is occasionally highlighted in popular books and exhibition catalogs (e.g., Monmonier, 1991; Barber and Harper, 2010). It occurs especially with respect to the selection (or omission) of features based, for example, on human legibility limitations, or simply technical and graphic reproduction constraints. It can also include aesthetic and even political sensitivities, especially where these are relevant to the map's purpose. Yet Eckert's phrase from a century ago suggests that scientific knowledge of human perception can also have a large part to play in determining good design, within the specific goals and constraints of the cartographer and the intended audience.

Eckert's thoughts on the role of "logic" (i.e., consistent and scientifically based design) in maps occurred several decades before maps started appearing on computer screens. In the digital age, we might expect more objective methods to be applied to map design and expect these to be partly automated rather than manually applied (e.g., Regnauld, 2001; García Balboa and Ariza López, 2008). Many of the basic design issues and user tasks remain constant in the digital age, while others have been added. Most map images still typically remain static at any one moment of viewing, but they can be altered or replaced instantly by the user through zooming, panning, hiding, or showing "layers" of symbols that represent geographic phenomena or real world processes. Some map visualizations are dynamic, changing either to reflect real-time movement (as in satellite navigation systems) or to simulate the passing of time at longer scales (as in historical reconstructions and weather maps).

Other geographic visualizations, such as remotely sensed images and aerial photographs, are barely cartographically designed "maps" at all, in that they are recordings (i.e., snapshots) of the Earth's surface by means of a camera or other sensor instrument (e.g., infrared or radiometers). They sometimes involve little purposeful graphic symbolization or cartographic generalization, beyond some color coding of different spectral values (or of noted changes between two snapshots from different times). However, more and more often we now see mergers of map *and* imagery, with topographic and/or thematic maps overlaid on a photographic or remotely sensed backdrop (e.g., in route-directing Web sites and in weather maps). Viewers can then see some of the physical landscape context surrounding each symbol and text label. In many professional map use contexts (and increasingly for the general public), the maps and images may also be "zoomed in" on a computer to a very large cartographic scale, with only a few buildings, fields, or streets visible within a typical screen display.

Meanwhile, as automated methods improve to capture, model, and display three-dimensional (3-D) topographic data, cartographic design may increasingly need to be applied to interactive 3-D visualization. Even (and perhaps especially) in such complex, multiple-viewpoint visualizations, key task-relevant information will still, as always, have to be identified and distinguished (e.g., Evans, 2009; Jobst, Kyprianidis, and Döllner, 2008). In all cases, we can still hope to fulfill Eckert's plea to "clarify" each visualization design scientifically, in order to optimize users' comprehension and problem solving as they view it.

We now turn to the core principles that can be applied to achieve this clarification, and their verification and application to actual map displays.

Cartographic Principles and Scientific Validation

One of the best-known attempts to establish the principles of good visualization design is Bertin's *Sémiologie Graphique* (1967/1983). Bertin identified the different "visual variables" that can be used by a graphic designer to convey information. He argued that specific visual variables – color, line orientation, shape, and pattern granularity – would be more effective than others in symbolizing and distinguishing the quantitative and qualitative concepts represented in a map or diagram. Bertin proposed principles for deciding which visual variables (i.e., *representing* variables) should be matched to which underlying *represented* variables. These principles include matching the dimensions of the visual variables with the underlying variables that they represent, in terms of scales of measurement (nominal, ordinal, etc.). He also suggested that the visual variables should be ordered such that the most thematically relevant data are the most perceptually salient in the image. Thematic relevance, in turn, depends on the task and user. Although Bertin's analytic framework is at a level of specificity to allow empirical testing, he did not himself attempt any scientific tests of his hypotheses. However, more recent empirical work has validated and clarified many of his ideas.

Applying Salience Models

Perhaps the most basic principle of cartography in general, and Bertin's (1967/1983) scheme in particular, is the principle that visual variables (e.g., symbol size, color shades, line orientation) that represent different variables in the data (for example, air temperature and wind direction in the case of a weather map) should be applied in such a way that important thematic information is visually salient. In order to test this principle we need methods of objectively evaluating the relative salience of locations (or meaningful graphical elements) in maps. Work in vision and attention in the past decade or so has provided such methods. Notably, Itti, Koch, and their colleagues (e.g., Itti, Koch, and Niebur, 1998; Itti and Koch, 2001) computationally modeled the role of bottom-up visual salience (or "'saliency") in directing users' initial attention within a visual image. Their model considers three component visual features of a display: color hue, color intensity (or color lightness/value), and orientation. Although Itti et al. based their choice of variables on the psychophysics literature, two of these visual features had also been among Bertin's key visual variables, and the third – color intensity – is also sometimes seen as simply another dimension of what Bertin labeled "color value." Values for each of these features are computed independently, and a "feature map" is calculated comparing both local and global center-surround differences. The feature maps for the different features are then combined to produce a single "salience map" for the whole image, essentially an objective measure of the relative salience of different regions of the display.

The validity of salience models is typically established by their ability to predict the locations of eye fixations on a display. The models assume that viewers will at least initially fixate the most salient regions of a display, which in this model are defined as regions that differ maximally from their neighboring regions on the analyzed visual features. Eye movement studies have shown mixed results in validating salience models. The models work best in predicting people's fixations within highly controlled images used in typical visual attention and psychophysics experiments, such as those studied within the "pop-out" and conjunctive search work of Treisman and Gelade (1980). However, they do not predict human eye fixations well when viewing more naturalistic images (Henderson, 2003). Presumably this is because the models measure bottom-up effects of *visual* salience on attention, whereas perception of naturalistic images is also significantly affected by top-down influences of the viewer's knowledge-based expectations and goals. This leads us to ask how far top-down influences apply to the perception of map displays.

In our laboratories we have used salience models both to design maps and to predict eye fixations on maps and aerial images. Fabrikant and Goldsberry (2005) used the Itti and Koch (2001) model in conjunction with an informal task analysis to redesign weather maps to make task-relevant information salient. A traditional weather map (showing pressure and temperature isofields, shown in Figure 35.1 (left panel) was downloaded from the World Wide Web. Assuming a task of inferring wind direction from pressure by nonexperts, Fabrikant and Goldsberry used cartographic principles (Bertin, 1967/1983) to make the task-relevant pressure information more salient and the task-irrelevant temperature information less salient. This essentially involved muting the colors showing temperature and darkening and thickening the lines showing pressure. The salience model was applied to the resulting maps, and the redesign and test cycle was repeated until the pressure systems were identified by the model as the most salient display regions.

The lower panel in Figure 35.1 depicts a grayscale reproduction of a colored weather map designed using the Itti salience model specifically for the task of inferring wind from pressure (Fabrikant, Hespanha, and Hegarty, 2010). The result of the salience model, presented in Figure 35.1c, depicts the weather map in the background, overlaid by the predicted sequence (arrows) of the first three eye fixations (circles). The salience map in the right panel of Figure 35.1c also uses a spotlight metaphor: More visible and lighter locations in this map indicate more visually salient locations in the map shown in the center panel of Figure 35.1b. The white spot in the right panel of Figure 35.1c shows that the low pressure cell indicated by the letter "L" is the most salient location in the map. This is probably due to the color contrast of the letter L (which is dark blue in the original) with its surrounding area (which is mainly pale pink), and the orientation contrast of the letter with the narrowly spaced isobars (i.e., lines of equal air pressure) that surround it.

In empirical testing, college freshmen were shown either traditional maps or redesigned maps. Their task was to judge whether an arrow in one region of the map indicated the direction that the wind would be blowing in that region, and thus required inferring wind direction from pressure. After training in meteorological principles that allowed them to infer wind from pressure, the students performed the wind inference task more quickly (Fabrikant et al., 2010) and more accurately (Hegarty, Canham, and Fabrikant, 2010) with the redesigned maps. However, this neat matching of salience with performance does not tell the whole story.

Limitations of Salience Models

The Itti salience model appears to predict eye fixations quite well for maps that closely follow the generally accepted cartographic design principles,

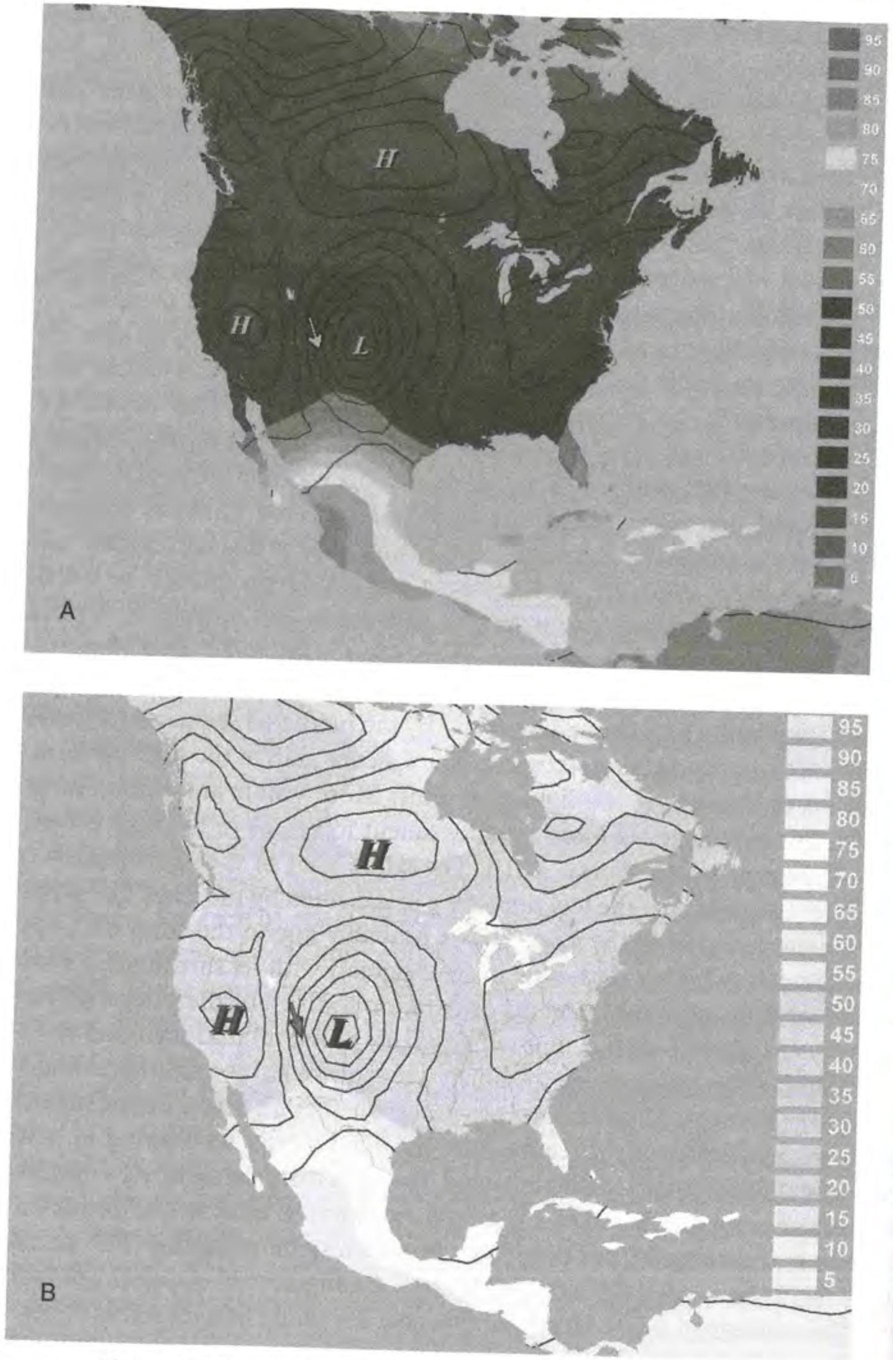

Figure 35.1. *(a) Original weather map (see Plate 10), (b) redesigned weather map (see Plate 11), and (c) predicted eye fixations based on a salience model (right). (Adapted from Fabrikant et al., 2010.)*

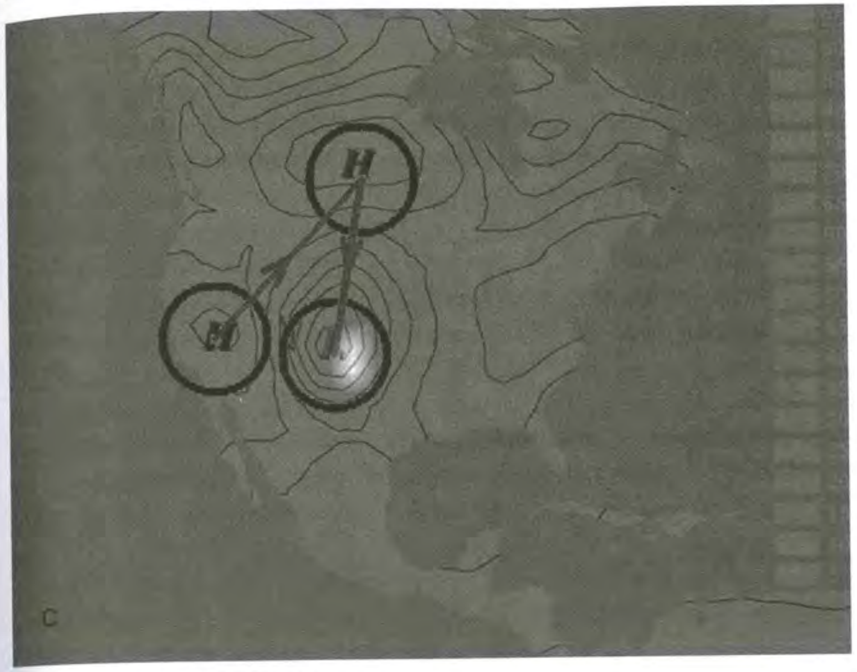

Figure 35.1. *(cont.)*

and thus optimize contrast between task-relevant and task-irrelevant items (Fabrikant and Goldsberry, 2005; Fabrikant et al., 2010). It also works well for predicting eye fixations on remotely sensed images created without any task specifically in mind (Swienty, Kurz, and Reichenbacher, 2007). However, when people are given a specific task to accomplish with a meaningful display, salience has more nuanced effects on eye fixations. For example, Hegarty et al. (2010) found that highly salient regions of a display did not attract eye fixations if they were not task relevant, and making task-relevant variables salient sometimes had a large influence on task performance even when there were only small effects on eye fixations.

The prediction of eye fixations from salience models is also more complex for experts than for novices. In one series of experiments (Davies et al., 2006; Lansdale, Underwood, and Davies, 2010) professional photogrammetrists, who worked for a national mapping agency, viewed aerial imagery with a view to spotting changes and/or supporting later recognition. Both tasks were common in their usual work, although the experimental task context was different. The Itti visual salience model predicted the initial gaze locations for both experts and novices. Within a few seconds after the image appeared on the screen, however, experts' gaze direction was no longer especially drawn to visually salient areas – unlike the novices'. It is possible that experts were instead choosing to view more *cognitively* salient areas, that is, objects in the image that were most likely to change, according to the experts' experience.

Subsequent change detection tasks (i.e., identifying where any change occurred) and memory tests showed a clear advantage for experts over novices. However, change detection and recognition performance were stronger in both participant groups for cues located at the most salient areas of the image, even though experts spent less or no time looking at those salient areas. A greater number of salient features increased accurate performance in both tasks. It thus appeared that experts could encode and recall even the parts of the image upon which they had not explicitly fixated, with visual salience contributing to their awareness of those areas. These results are consistent with some previous evidence concerning the likely effect of expertise on map interpretation (Lowe, 2003).

This research also suggests that experts can be driven by task differences to follow different scanning strategies and to utilize noncentral vision, so that visual salience plays a role beyond simply predicting gaze fixations. As predicted by Hoffman (1990), the concept of *perceptual learning* appears to apply in areas such as remote sensing, as experienced viewers adapt their attention and perception to specific tasks. The research also suggested that foveal gaze direction is inadequate for predicting experts' superior cognitive performance in visual display tasks (see also Davies, Cannon, and Gould, 2007). Furthermore, these results match similar expertise findings in such varied domains as chess (Charness, Reingold, Pomplun, and Stampe, 2001) and soccer (Ward and Williams, 2003).

Clearly, then, "top-down" influences on attention come into play more in situations involving specific tasks and the potential for expertise development. Although alternative salience models (e.g., focusing on clutter, as in Rosenholtz, Li, and Nakano, 2007) have sometimes shown more promise in such situations (Fabrikant et al., 2010), overall the work reminds us that visual variables by themselves will always play only a limited role in directing a map user's attention (see also Brunyé and Taylor, 2008; Noudoost et al., 2010).

Indeed, sometimes a specific goal may even render the visual variables far less of a driving factor than other design considerations. One example is a scenario in which people are presented with a visual scene (either from a simplified 3-D virtual world or from a real world outdoor scene or photograph), and with a map on which they must indicate the location and orientation from which they would see the scene. In this situation, most people seem to rely more on salient characteristics of the scene than of the map – even when this causes them to make errors because the map cannot reflect the 3-D geometry of the scene (e.g., Pick et al., 1995; Gunzelmann et al., 2004; Davies and Peebles, 2010). Specifically, people select a highly salient landmark or other unique aspect of the scene, and then attempt to match that to the 2-D map, rather than trying to abstract and match the broader 2-D geometry of their surroundings. The proposal by some (e.g., Hermer and Spelke, 1994) that geometry may be more fundamental to people's orientation and spatial cognition than landmarks seems not to hold in these somewhat more realistic scenarios, at least for most adults. Thus, in this very common map use scenario, it may be more important and helpful to select and symbolize landmarks on a map that can be easily matched to the most salient features of the real world environment.

It should not, of course, surprise us that the way we view a map will depend very much on the task we are attempting to perform with it. In the past this has been viewed as a problem for deducing any generic knowledge from cartographic design research, after researchers had initially fallen into the trap of assuming that a user would have only one type of goal in viewing a given map (see Petchenik, 1983; Montello, 2002). This need to prevent overprescription suggests that we need more cognitive and perceptual research and systematic empirical evaluation within cartographic design. As with the design of any artifact, designers and researchers first need to identify an appropriate range of intended tasks and user characteristics, and then use these to direct their research studies or design evaluations.

In summary, the preceding body of research has shown that following Bertin's analysis of visual variables and designing maps according to his principles can and does make a difference. By redesigning maps according to Bertin's principles we can to some extent "offload cognition onto perception" (Scaife and Rogers, 1996) by making salient perceptual attributes draw our eyes to the task-relevant stimuli, or at least speed up a top-down search.

This research has also established the importance of objectively measuring visual salience and evaluating visualization designs individually within the relevant task context. Finally it suggests that designers may be able to optimize the visual salience of key task-critical map features and areas by calculating salience maps of the tested designs, based on the model of Itti et al. (1998). At the time of writing, this model can be downloaded from Laurent Itti's Web site at the University of Southern California (ilab.usc.edu/bu). Increasingly sophisticated models are likely to evolve from this and similar work.

Assessments of Performance with Different Cartographic Designs

We now turn to a more general review of how we can use methods from experimental psychology to evaluate the effectiveness of cartographic designs. We start by reviewing how simple measures of performance and reaction time can be used for this purpose, but focus primarily on eye tracking and change-detection paradigms, which have recently been adapted to the evaluation of map designs and which can yield more in-depth information about the online cognitive processes of viewers as they work with map displays.

Measures of Accuracy and Response Time

Since many map use tasks involve searching for information, one obvious way to test the effectiveness of map designs is to compare the speed and accuracy with which participants locate a specified item within two or more contrasting map designs. Following standard experimental design procedures, a series of tasks is required to minimize the effect of random response variability, and ideally the tasks are randomly ordered to prevent experimental bias.

In a map interpretation task, participants may be required to respond in various ways, from the traditional method of hitting a specified key or button to respond to a yes-no question, to locating the object by physically pointing at a touch screen. This kind of procedure, while most strongly associated with psychophysics research, can also be very useful with cartographic map designs (e.g., Lloyd, 1997). However, such methods on their own provide little insight regarding the specific cognitive processes used to accomplish different tasks, what map features attract users' attention, and so on, although debriefing participants can help with this. The methods also require experimental controls that often mean that the studied tasks are relatively artificial, so they sometimes provide little insight regarding how quickly and efficiently users might perform a realistic task with the map in an actual use context.

Eye Movement Tracking

Tracking eye movements (eye tracking) has been used in cartographic research since the early 1970s, but early studies led to few advances in design knowledge (MacEachren, 1995; Montello, 2002). Eye tracking research was expensive and difficult, given the technologies available at the time. In addition, the complexities of the relationship between eye fixations and visual attention were not fully appreciated in the early 1970s. For example, we now know that while eye fixations are correlated with visual attention, attention can also be dissociated from eye fixations (Posner, 1980), so that there is no strict guarantee that a visually fixated object is what the viewer is attending to at the time. A third reason is that eye movement research at that time was concerned primarily with reading, so that methods for the analysis of eye fixations on visual-spatial displays had not been fully developed (i.e., for analysis of fixations over fields of view versus text).

In recent years eye tracking technologies have become more available, useful, and usable, and we know more about the relationship between eye fixations and visual attention (Henderson and Ferreira, 2004). We also have better statistical methods for analyzing eye fixations on visual displays, including "area-of-interest" analyses in which we examine the number of fixations or total gaze duration on different (and sometimes prespecified) regions of a display.

As a result, researchers are now in a better position to apply eye tracking methods to evaluate map designs and principles of cartography. Eye tracking can help to evaluate the design of maps, especially in situations where users need to search the display visually as a critical part of their tasks. This method can augment other usability measures, and typical measures of task performance such as accuracy and speed, by indicating aspects of the user's perception and attention during the task. Eye fixations can reveal both how the users went about the task and how they may have been distracted from performing the task optimally (Fuhrmann et al., 2005; Çöltekin, Fabrikant, and Lacayo, 2010; Çöltekin, Heil, Garlandini, and Fabrikant, 2009).

For navigation and other tasks that involve immersion within the geographic environment itself, less intrusive helmet- or spectacle-mounted mobile eye trackers now allow for field studies with far more precise and useful data recording, which can also take into account the use of other sources of visual information (such as the surrounding scene) as well as the map itself. This may even be taking us to the point where maps can become secondary or omitted for many tasks that are immersed in the spatial environment itself. Instead, feedback from an eye tracking system can allow for geographic information to be directly superimposed upon the scene within the viewer's vision using "augmented reality" techniques (e.g., Baldauf, Fröhlich, and Hutter, 2010). Augmented reality might also be used deliberately to direct the user's attention in a task-relevant direction (e.g., Biocca, Owen, Tang, and Bohil, 2007). Of course, these partial and dynamic visualizations are still subject to visual design issues in themselves, and require evaluation techniques like those that we advocate for static maps.

The advent of advanced methods that allow the tracking of eye fixations during movement in the world, and the application of eye fixations to moving stimuli in dynamic visualizations, have raised new challenges for analyzing the data from eye tracking studies. In addition, because visual attention is a complex phenomenon, eye tracking and other techniques need to do more than simply identify initial fixations or relative proportions of time spent gazing at different areas of a static visual display.

Sequential analysis approaches can help us to understand the time course of users' interactions with a visualization, which is a first step to detailing the reasoning strategies involved. It is especially relevant where changes in a physical environment over time are a key aspect of the user's task. Changes over time have been studied by examining how people view static visualizations that represent dynamic phenomena via multiple small images ("small multiples") (Fabrikant et al., 2008). However, as analysis tools improve, eye tracking techniques are becoming particularly attractive for studying performance with dynamically changing computer-based visualizations, so that the impact of display changes on the viewer's sequence of fixations can be better understood. In a neat twist, researchers have harnessed the spatial analysis techniques built into the same geographic information systems (GIS) that produce maps and other spatial visualizations in the first place, to analyze and interpret the patterns of users' eye movements upon those visualizations (Çöltekin, Fabrikant, and Lacayo, 2010).

While eye tracking can provide important insights regarding how users interact with maps, a challenge in moving forward is that some of the time-saving features of modern eye tracking software can sometimes be difficult to apply to geographic visualization displays. This is especially true when analyzing ecologically valid tasks rather than artificial "view-then-answer-questions" lab experiments. For example, "areas of interest" analyses (e.g., comparing the relative amount of time or number of fixations spent viewing specific map regions) can prove difficult or even impossible to apply meaningfully in highly detailed displays (Riley, 2006). An added complexity of maps is that they often overlay variables, so that the same area of interest in a map can present information about different variables (e.g., pressure and temperature in a weather map; Hegarty et al., 2010). In a digital display, layers of information can also be shown or hidden, panned and zoomed, so that it is crucial to match fixation data to the content of the screen at its location *at that moment* in time.

It is also important to distinguish time spent attending to the user interface of the geographic information system or other mapping software from time spent studying the actual map content. For example, "interface" time might include time spent issuing commands to zoom, pan, label, query, print, save, change visual variables, or hide/show data layers. This separation can be difficult when the technology allows users to perform commands via keyboard shortcuts, or where user interface items are closely integrated with the map itself (for example, via "popup" information boxes or menus).

For both of these problems, part of the solution is to integrate the eye movement data with logging by the geographic information system itself, so that each gaze fixation is linked correctly to a known map or user interface item that was displayed at that screen location at that time. Another part of the solution might be to collect think-aloud protocols concurrently with eye fixations, to aid interpretation of the user's intentions during a given period of interaction.

Finally, some professional map use environments can impose difficulties for evaluating designs using eye tracking. These include situations in which people view maps under poor or variable lighting conditions, or in which they use stereoscopic imagery. Remote desktop eye tracking systems (the least intrusive kind for desk-based task recording) still do not work well with people who wear glasses. However, in general more is now possible at far less cost (and, with increasingly sophisticated software, far less analysis effort) than in previous decades. As a result, eye tracking is now being used routinely with digital map displays not only in psychology laboratories, but in evaluations by professional geographic information providers such as national mapping agencies (e.g., Brodersen, Andersen, and Weber, 2001; Davies et al., 2007).

In the recent enthusiasm for eye tracking as a method, it is often forgotten that this method has only become widely available and accessible to researchers in the past decade or two. Yet, of course, prior to that time, psychologists and other researchers still managed to do much useful work on visual attention and design. The methods they used, and also more recent interest in phenomena such as "change blindness," include some potentially useful alternatives or additions to eye tracking. These will now be briefly outlined.

Probe Tasks

An alternative but less precise method of identifying where a participant has just been looking on a visual display is the letter-probe or number-probe task. In this paradigm, immediately after a map image disappears from the screen, it is replaced by an array of letters or digits for around 100 milliseconds (allowing at most one saccade, i.e., a single eye movement typically taking around 50 milliseconds). This in turn is immediately followed by a visual "mask" (e.g., an array of stars or dots), displayed for 500 milliseconds. The mask disrupts the visual afterimage, so that the viewers cannot mentally "scan" around it. Participants are then asked to recall what letters (or numbers) they managed to see. (Of course, in experiments using this task, it is usually important also to include a subsequent memory or interpretation task that requires recall of the map itself, so that participants do not ignore it and just focus on the letter probe!) This method has the advantage of being inexpensive and nonintrusive, but it locates a participant's visual gaze within the initial exposure with only an approximate spatial resolution. It also introduces an artificial task into the procedure, which reduces ecological validity. Nevertheless, it can be sufficient to test relatively straightforward hypotheses about the impact of design and task variables on visual attention, especially in the immediate early stages of viewing a map display.

Davies et al. (2006) used this method to examine the interaction between photogrammetry expertise and visual salience in people's initial fixations on aerial photographs (where the only image-related task was a subsequent recognition test), as mentioned earlier. Whenever a letter from the probe display was recalled by a participant, the researchers calculated the relative visual salience of the specific area of the initial image that had been replaced by that letter, and thus compared the two participant groups on their tendency to look more at salient than nonsalient areas before the image was replaced by the grid of letters. The method also allowed the researchers in this case to compare different types of image (e.g., rural versus urban imagery) for their relative tendencies for salience to guide the viewer's attention. Thus, despite its cruder spatial resolution than eye tracking, this paradigm can still be used to compare two or more designs or visualization types, as well as different participant groups or viewing conditions.

Change Detection Paradigms

Change detection or change "blindness" paradigms have been used in research with map visualizations (Garlandini and Fabrikant, 2009; Lansdale et al., 2010), with the assumption that users will be more likely to notice a change in the map if it is either more perceptually salient or more relevant to the task. As with the number-probe paradigm, this can be done either with or without additional eye tracking. In its simplest form, a change detection task involves showing a map or other visualization for a certain period on a single screen, then following it with a different version in which at least some objects or features are altered in some way (after a briefly blank screen, to prevent the impression of movement that would draw attention to the change). Alternatively, to reduce memory load in the task (which could affect performance regardless of original attention), the images may be compared side-by-side – although visual working memory is of course still required as the viewer switches attention from one to the other. Change detection of this kind was a central task in the research described earlier on performance of expert photogrammetrists (Lansdale et al., 2010), as it is an important part of their work.

A third method, which again reduces but does not completely avoid working memory load, is to alternate the two images repeatedly every second or two (again interspersed with a brief blank image to preclude the impression of movement), until the participant responds in some way to indicate the presence/absence or location of the change. This is called the "flicker paradigm" (Rensink, O'Regan, and Clark, 1997).

Garlandini and Fabrikant (2009) give an example of using the flicker paradigm with geographic visualizations and discuss the usefulness of this as an evaluation method for map design. Their study focused on comparing the relative effects of changing different visual variables (as in Bertin's scheme). The key dependent variable was response time (length of elapsed time before participants hit a key to indicate that they spotted the change), although again eye fixations were also recorded. Participants were also asked to locate and describe the change. The results indicated that changing size, color hue, and color value (effectively altering contrast) was far more noticeable to participants than changing the orientation of a map feature. Change blindness tasks can thus allow the researcher either to identify and/or to eliminate the use of visual variables that appeared to mislead participants on error-prone trials, or to compare accuracy of performance between different visual variables, participant types, or map designs. Again eye tracking is not essential to this method, although it can greatly aid interpretation of results.

Assessing User Preferences

What role do user preferences play in evaluating map designs? If people have good intuitions about the effectiveness of visualization, cartographers and geographic systems engineers who are working under time pressure may wonder whether objective empirical evaluations as described earlier can be avoided, in favor of subjective user evaluations. In addition, since modern mapping software allows users of displayed graphics to alter their visual appearance at will, one might ask whether "design" can now be effectively left in the hands of end users. Examination of this issue has long suggested otherwise (Green 1993; McGuinness, 1994; Smallman, Cook, Manes, and Cowen, 2007; Hegarty, Smallman, Stull, and Canham, 2009). Studies have shown repeatedly that when allowed to choose (or given a chance to create) more realistic or more iconographic (cartographically symbolized) visualizations, inexperienced viewers prefer to see or create more realistic and complex terrain maps or imagery with much extraneous (task-irrelevant) detail. However, the same participants perform decision-making tasks in various domains significantly less effectively with those preferred visualizations than with cartographically principled and iconographic maps that present only task-relevant information. The importance of this simplification to match task requirements follows the same principle as for other types of information display (Tufte, 2001; Kosslyn, 2006; Hegarty, 2011).

This is an important point for applied domains in which traditional "market" research is likely to clash with empirical validation of cartographic design, particularly for less experienced users. In a commercial setting a compromise may have to be reached, but in professional or safety-critical environments, where optimally fast and/or accurate performance can be essential, the preceding findings suggest strongly that map design should not be left to personal preference or customization, especially by untrained users. Geographic visualization designers and researchers can thus demonstrate the cost-effectiveness of skilled and task-focused design and empirical evaluation, in terms of increased efficiency and productivity.

Conclusions

Despite the traditional focus on cartography as a craft, recent studies have shown a surprising degree of congruence between conventional cartographic (and general graphic design) principles and measures of people's performance such as eye movements, response times, accuracy, recall, and recognition of visual elements in the display. Eye tracking, although revealing, is not always essential in evaluating a geographic visualization. However, multiple measures and a systematic empirical approach are critical, given that user preferences in this area do not always match with perceptual and cognitive reality. Many questions remain, particularly regarding the extent to which the principles and methods discussed here may still be applicable to novel technologies such as augmented reality or stereoscopic 3D displays. Nevertheless, future researchers should be able to address these via the same general approach – learning from theories and models in psychophysics, but bearing in mind the broader cognitive task and context of the real world map user, and the continuing value of simplifying visualization displays to reflect, and draw the user's eyes to, the most task-relevant information.

References

Baldauf, M., Fröhlich, P. and Hutter, S. (2010). KIBITZER: A wearable system for eye-gaze-based mobile urban exploration. In H. Saito et al., AH'10: Proceedings of the 1st Augmented Human International Conference. (Article No. 9) New York: ACM.

Barber, P. and Harper, T. (2010). *Magnificent maps: Power, propaganda and art.* London: The British Library.

Bertin, J. (1967/1983) *Sémiologie Graphique: les diagrammes – les réseaux – les cartes (Semiology of Graphics).* Madison: University of Wisconsin Press.

Biocca, F., Owen, C., Tang, A. and Bohil, C. (2007). Attention issues in spatial information systems: Directing mobile users' visual attention using augmented reality. *Journal of Information Systems*, 23(4), 163–184.

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In: Jordan, P. W., Thomas, B., Weerdmeester, B. A. and A. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189–194). London: Taylor and Francis.

Brodersen, L., Andersen, H. K. and Weber, S. (2001). *Applying eye-movement tracking for the study of map perception and map design.* Copenhagen: Kort and Matrikelstyrelsen (National Survey and Cadastre), Publications Series 4. Available from KMS, Rentemestervej 8, DK – 2400 Copenhagen NV, Denmark. Downloaded from www.kms.dk.

Brunyé, T. T. and Taylor, H. A. (2008) When goals constrain: Eye movements and memory for goal-oriented map study. *Applied Cognitive Psychology*, 22, 1–16.

Canham, M. and Hegarty, M. (2010). Effects of knowledge and display design on comprehension of complex graphics. *Learning and Instruction*, 20, 155–166.

Charness, N., Reingold, E. M., Pomplun, M. and Stampe, D. M. (2001). The perceptual aspects of skilled performance in chess: Evidence from eye movements. *Memory and Cognition*, 29(8), 1146–1152.

Çöltekin, A., Fabrikant, S. I. and Lacayo, M. (2010). Exploring the efficiency of users' visual analytics strategies based on sequence analysis of eye movement recordings. *International Journal of Geographical Information Science*, 24(10), 1559–1575.

Çöltekin, A., Heil, B., Garlandini, S. and Fabrikant, S. I. (2009). Evaluating the effectiveness of interactive map interface designs: A case study integrating usability metrics with eye-movement analysis. *Cartography and Geographic Information Science*, 36(1), 5–17.

Davies, C., Cannon, C. and Gould, J. (2007). The eyes have it: Human change detection. In G. Hart et al., *Ordnance Survey Research Annual Review 2006–07.* Southampton, UK: Ordnance Survey, pp. 10–11. Available from: Ordnance Survey, Adanac Drive, Southampton SO16 0AS, UK. Downloaded from www.ordnancesurvey.co.uk.

Davies, C. and Peebles, D. (2010). Spaces or scenes: Map-based orientation in urban environments. *Spatial Cognition and Computation* 10 (2–3), 135–156.

Davies, C., Tompkinson, W., Donnelly, N., Gordon, L. and Cave, K. (2006). Visual saliency as an aid to updating digital maps. *Computers in Human Behavior*, 22, 672–684.

Dent, B. D. (1999). *Cartography: Thematic map design.* Dubuque, IA: Wm. C. Brown.

Eckert, M. (1908/1977). On the nature of maps and map logic (trans. W. Joerg). In L. Guelke (Ed.), *Cartographica monograph 19 (Supplement 1 to Canadian Cartographer*, Vol. 14, 1977), 1–7.

Evans, S. (2009). 3D cities and numerical weather prediction models: An overview of the methods used in the LUCID project. Working Papers Series Paper 148 – May 09. Centre for Advanced Spatial Analysis, University College London, ISSN 1467–1298. Downloaded from http://eprints.ucl.ac.uk/17404/1/17404.pdf, 22 October 2011.

Fabrikant, S. I. and Goldsberry, K. (2005). Thematic relevance and perceptual salience of dynamic geovisualization displays. Paper presented at the 22nd ICA/ACI International Cartographic Conference, A Coruña, Spain.

Fabrikant, S. I., Hespanha, S. R. and Hegarty, M. (2010). Cognitively inspired and perceptually salient graphic displays for efficient spatial inference making. *Annals of the Association of American Geographers*, 100 (1), 1–17.

Fabrikant, S. I., Rebich-Hespanha, S., Andrienko, N., Andrienko, G. and Montello, D. R. (2008). Novel method to measure inference affordance in static small-multiple map displays representing dynamic processes. *The Cartographic Journal*, 45(3), 201–215.

Fuhrmann, S., Ahonen-Rainio, P., Edsall, R. M., Fabrikant, S. I., Koua, E. L., Tobón, C., Ware, C. and Wilson, S. (2005). Making useful and useable geovisualization: Design and evaluation issues. In J. Dykes, A. M. MacEachren and M.-J. Kraak, *Exploring geovisualization* (pp. 553–566). Amsterdam: Elsevier.

García Balboa, J. L. and Ariza López, F. J. (2008). Generalization-oriented road line classification by means of an artificial neural network. *Geoinformatica*, 12(3), 289–312.

Garlandini, S. and Fabrikant, S. I. (2009). Evaluating the effectiveness and efficiency of visual variables for geographic information visualization. In K. S. Hornsby et al. (Eds.), *COSIT 2009. Lecture notes in computer science*, Vol. 5756. Berlin and Heidelberg: Springer-Verlag, 195–211.

Green, D. R. (1993). Map output from geographic information and digital image processing systems: A cartographic problem. *The Cartographic Journal*, 30, 91–96.

Gunzelmann, G., Anderson, J. R. and Douglass, S. (2004). Orientation tasks with multiple views of space: Strategies and performance. *Spatial Cognition and Computation*, 4(3), 207–253.

Hegarty, M. (2011). The cognitive science of visual-spatial displays: Implications for design. *Topics in Cognitive Science*, 3, 446–474.

Hegarty, M., Canham, M. S. and Fabrikant, S. I. (2010). Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 36(1), 37–53.

Hegarty, M., Smallman, H. S., Stull, A. T. and Canham, M. (2009). Naive cartography: How intuitions about display configuration can hurt performance. *Cartographica*, 44, 171–187.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Science*, 7, 498–504.

Henderson, J. M. and Ferreira, F. (2004) Scene perception for psycholinguists. In J. M. Henderson and F. Ferreira (Eds.), *The interface of language, vision and action*. New York: Psychology Press.

Hermer, L. and Spelke, E. S. (1994). A geometric process for spatial reorientation in young children. *Nature* 370, 57–59.

Hoffman, R. R. (1990). Remote perceiving: A step toward a unified science of remote sensing. *Geocarto International*, 2, 3–13.

Itti, L. and C. Koch. (2001). Computational modeling of visual attention. *Nature Reviews Neuroscience*, 2, 194–203.

Itti, L., Koch C. and Niebur E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 1254–1259.

Jobst, M., Kyprianidis, J. E. and Döllner, J. (2008). Mechanisms on graphical core variables in the design of cartographic 3D city presentations. In A. Moore and I. Drecki (Eds.), *Geospatial vision: New dimensions in cartography*. Selected Papers from the 4th National Cartographic Conference GeoCart'2008, New Zealand. Lecture Notes in Geoinformation and Cartography. Berlin: Springer, pp. 45–59.

Kosslyn, S. (2006). *Graph Design for the Eye and Mind*. New York: Oxford University Press.

Kraak, M.-J. and Ormeling, F. (2010). *Cartography: Visualization of spatial data* (3rd ed.). Harlow, UK: Pearson.

Lansdale, M., Underwood, G. and Davies, C. (2010). Something overlooked? How experts in change detection use visual saliency. *Applied Cognitive Psychology*, 24(2), 213–225.

Lloyd, R. (1997). Visual search processes used in map reading. *Cartographica*, 34 (1), 11–32.

Lowe, R. K. (2003). Animation and learning: Selective processing of information in dynamic graphics. *Learning and Instruction*, 13, 157–176.

MacEachren, A. M. (1995). *How Maps Work: Representation, Visualization and Design*. New York and London: Guilford Press.

McGuinness, C. (1994). Expert/novice use of visualization tools. In A. M. MacEachren and D. R. F. Taylor (Eds.), *Visualization in modern cartography* (pp. 85–199). Oxford: Pergamon (Elsevier).

Monmonier, M. (1991). *How to lie with maps*. Chicago: University of Chicago Press.

Montello, D. R. (2002). Cognitive map-design research in the twentieth century: Theoretical and empirical approaches. *Cartography and Geographic Information Science*, 29(3), 283–304.

Noudoost, B., Chang, M. H., Steinmetz, N. A. and Moore, T. (2010). Top-down control of visual attention. *Current Opinion in Neurobiology*, 20(2), 183–190.

Petchenik, B. B. (1983). A mapmaker's perspective on map design research 1950–1980. In D. R. F. Taylor (Ed.), *Graphic communication and design in contemporary cartography* (pp. 37–68). Chichester, UK: John Wiley and Sons.

Pick, H. L., Heinrichs, M. R., Montello, D. R., Smith, K., Sullivan, C. N. and Thompson, W. B. (1995). Topographic map reading. In P. A. Hancock, J. M. Flach, J. Caird and K. J. Vicente (Eds.), *Local applications of the ecological approach to human-machine systems*, Vol. 2 (pp. 255–284). Hillsdale, NJ: Lawrence Erlbaum.

Posner, M. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology* 32, 3–25.

Regnauld, N. (2001). Contextual building typification in automated map generalization. *Algorithmica*, 30(2), 312–333.

Rensink, R. A., O'Regan, J. K. and Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8, 368–373.

Riley, C. (2006). Eye tracking as a tool in cartographic research: Assessing map scale for suitability of data overlay. M.Sc. (Ergonomics) thesis, Loughborough University, Loughborough, UK.

Robinson, A. H. (1952). *The look of maps*. Madison: University of Wisconsin Press.

Rosenholtz, R., Li, Y. and Nakano, L. (2007). Measuring visual clutter. *Journal of Vision*, 7, 1–22.

Scaife, M. and Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies, 45*, 185–213.

Slocum, T. A., McMaster, R. B., Kessler, F. C. and Howard, H. H. (2010). *Thematic cartography and geographic visualization* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Smallman, H. S., Cook, M. B., Manes, D. I. and Cowen, M. B. (2007). Naïve realism in terrain appreciation. Proceedings of the 51st Annual Meeting of the Human Factors and Ergonomics Society. Santa Monica, CA: Human Factors and Ergonomics Society, pp. 1317–1321.

Swienty, O., Kurz, F. and Reichenbacher, T. (2007). Attention guiding visualization in remote sensing IIM systems. *Photogrammetrie, Fernerkundung, Geoinformation*, 4, 239–251.

Trafton, J. G., Marshall, S., Mintz, F. and Trickett, S. B. (2002). Extracting explicit and implicit information from complex visualisations. In M. Hegarty, B. Meyer and N. H. Narayanan (Eds.), *Diagrams 2002*. Lecture Notes in Artificial Intelligence. Vol. 2317 (pp. 206–220). Berlin: Springer-Verlag.

Treisman, A. and Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, 12, 97–136.

Tufte, E. T. (2001). *The visual display of quantitative information* (2nd ed.). Cheshire, CT: Graphics Press.

Ward, P. and Williams, A. M. (2003). Perceptual and cognitive skill development in soccer: The multidimensional nature of expert performance. *Journal of Sport and Exercise Psychology*, 25, 93–111.