Star Plots
How Shape Characteristics Influence Classification Tasks[1]

Alexander Klippel[a], Frank Hardisty[a,b], Chris Weaver[c]

GeoVISTA Center, Department of Geography, Pennsylvania State University, PA, USA[a]
Dutton e-Education Institute, Pennsylvania State University, PA, USA[b]
School of Computer Science and the Center for Spatial Analysis, University of Oklahoma, OK, USA[c]
[klippel,hardisty]@psu.edu, cweaver@ou.edu

Corresponding address:
Dr. Alexander Klippel
The Pennsylvania State University
Department of Geography, GeoVISTA Center
302 Walker, University Park, PA 16802, USA
Phone: 1-814-222-2324
Fax: 1-814-863-7943

Abstract. Our research addresses the question of how to design interfaces for spatial analysis such that they support cognitive processes. In this paper we specifically target the question of map symbol design for the analysis of multivariate data, which is a common problem in cartography and related fields. We focus on *star plots* and the largely unaddressed question of how to assign variables to rays in a star plot and which consequences specific shapes—as the result of data characteristics and the assignment of variables to rays—have on interpretation and classification. We conducted an experiment with two conditions that were designed to shed light on the question: does the shape of a star plot influence the interpretation (meaning) of the data it represents in a classification task? While previous research on multivariate point symbols has addressed this question for Chernoff faces, for example, few connections have been made to the shape of a star plot and its potential influence on meaning. We found that certain salient shape characteristics induced by variations along the horizontal and vertical axis increase the classification speed. However, we also found that salient shapes, such as *has one spike*, introduce a perceptual similarity that overrides the assumed similarities in the meaning of the represented data.

Introduction

Maps are one of the most ancient forms of human communication (Harley & Woodward, 1987). While (very) early map design focused on the representation of environmental spaces, map design has evolved to depict abstract, multidimensional data spaces for which geographic location establishes the frame of reference but the thematic attributes create the spatial patterns (e.g., Slocum, McMaster, Kessler, & Howard, 2005). Modern information systems and a vast number of freely available tools allow everyone to create maps from a wide variety of data sets. While cognitive research on maps and map symbols is recent, beginning in the 20th century (Montello, 2002), design choices are for the most part still left to the user with often little guidance provided through the software (for exceptions / overviews see, Harrower & Brewer, 2003; MacEachren, 1995; Bertin, 1974; Bollmann, 1981; Wilkinson, 2005; lNelson, 2007).

The importance of behavioral studies in cartography was recognized in the early 20th century by the German cartographer Eckert (1921) and came into the focus of cartographic research with Arthur Robinson's book "The Look of Maps" (Robinson, 1952) (see Montello 2002 for an overview). Since then, the behavioral validation of map design has not kept pace with the sensational developments producing

[1] Article will appear in *Cartography and Geographic Information Science*

information graphics and tools for mapping and visual analysis. Our research on star plots is placed in the tradition of cognitive cartographers seeking to establish behavioral validation of design choices of maps and map symbols (e.g., Arnheim, 1976; Brewer & Pickle, 2002; Chang, Antes, & Lenzen, 1985; Olson, 1984; Robinson & Petchenik, 1976). One of our over-arching goals is to apply knowledge gained by cognitive cartographers to current map-making methods. Thus, the current research can be seen as an example of cognitive cartographic research methods applied to assessing user-driven, contingent map graphics. Furthermore, we aim to formally implement findings on cognitive and perceptual interaction with maps and map symbols with the goal of integrating decision support on how to optimally—from a cognitive point of view—design maps and map symbols.

The rest of this paper is structured as follows: Section 2 provides background information on the visualization of multivariate data. We focus specifically on point symbols and how they have been used in maps. Section 3 details the experimental setup of a behavioral study focusing on the different design options of star plot glyphs. We distinguish two conditions that differ in the way variables are assigned to rays in a star plot. Within this section, we detail various tools that have been implemented specifically to aid in the analysis of the behavioral data. Section 3 contains also a detailed discussion of the data analysis. Section 4 concludes by relating the results to design principles for star plots and directions for future behavioral experiments necessary to establish a metric for cognitively adequate visualizations.

Background

One of the challenges of modern information technology is the need to handle the abundance of data available on individual entities such as counties, countries, cities, cars, or plant communities (Gore, 1998; Ware, 2004; Adrienko, Adrienko, Dykes, Fabrikant, & Wachowicz, 2008). With this increase in the number of variables comes increased difficulty in sensibly displaying multivariate data sets. Visual data displays become more complex and harder to understand as both the complexity of displays and the interdependence of variables grows.

Our focus in this paper is on star plot glyphs (Chambers, Cleveland, Kleiner, & Tukey, 1983), such as shown in Figures 1 and 2. The star plot glyph is one of the earliest multivariate symbolization techniques mentioned in the literature on multivariate numerical data representation in general, and on multivariate statistical mapping in particular. One of the earliest such forms was in Florence Nightingale's *Notes on Matters Affecting the Health, Efficiency and Hospital Administration of the British Army* in 1858 (Nightingale, 1858).

Figure 1 illustrates star plots in a map generated using the GeoViz Toolkit. The map displays a combination of multivariate attribute information and geographical information by drawing a star plot glyph at the centroid of each geographical unit, in this case the 48 contiguous United States. When star plots are near each other (and exhibit similar characteristics), such as in the Plains States and the Southeast, they can become a basis for discerning possible clusters or large-scale regions.

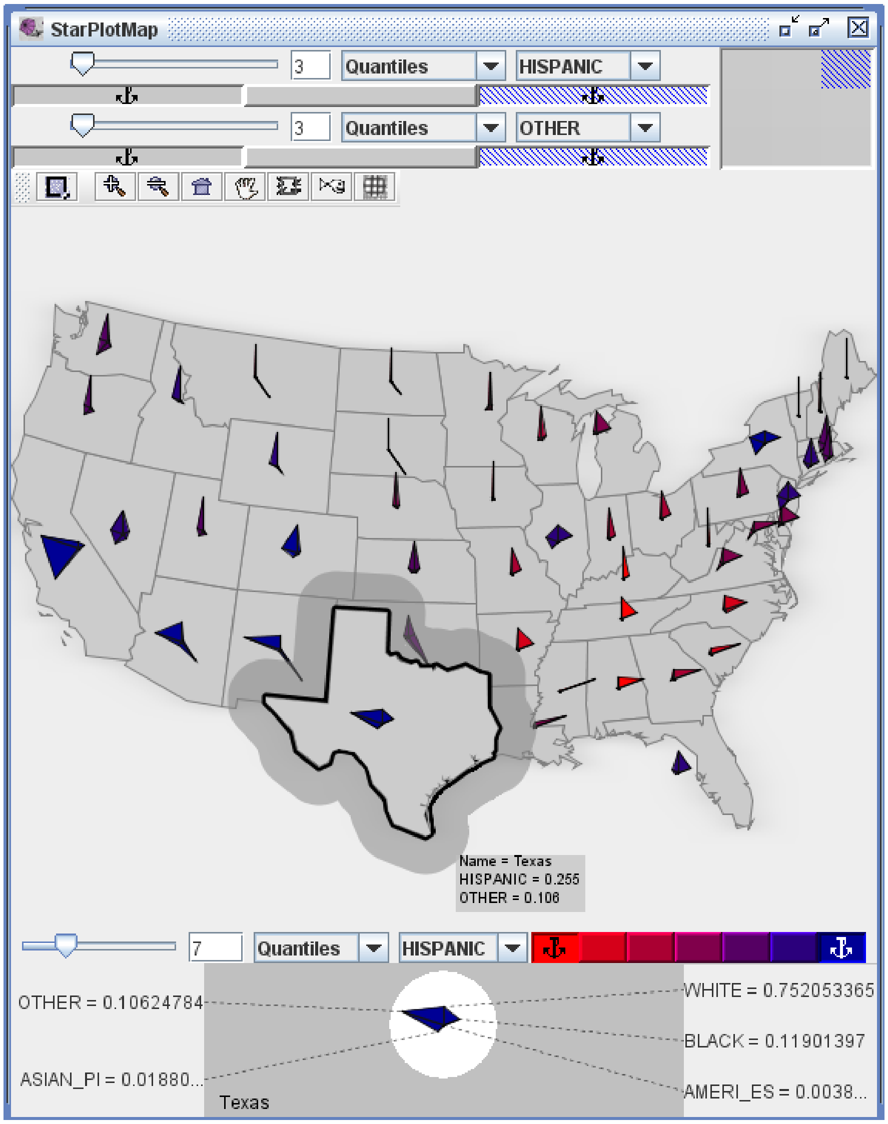


Figure 1. Star plots glyphs in map regions in an interactive visual-analytic workspace.

Figure 2 illustrates a single star plot glyph. The number of variables depicted corresponds to the number of rays emanating from the center of the star plot. The length of each ray is proportional to the value it represents, i.e. the variable. There are several variations in the actual design of star plots. The example in Figure 2 shows five variables for the state of Texas, all of them self-reported race/ethnicity categories as gathered by the U.S. Census Bureau: percentage White, percentage Black, percentage American Indian and Eskimo, percentage Asian and Pacific Islander, and percentage Other race. While in general it is possible to represent various kinds of data by star plot glyphs, some kinds of data need to be transformed before they can be reasonably represented by a star plot glyph. One such example is negative data. As one advantage of spar plot glyphs is the depiction of only minimally processed data especially for exploratory data analysis we would argue for star plot glyphs being most effective for positive data. This is, of course, a point worth discussing in greater detail although it is not the focus of this article.

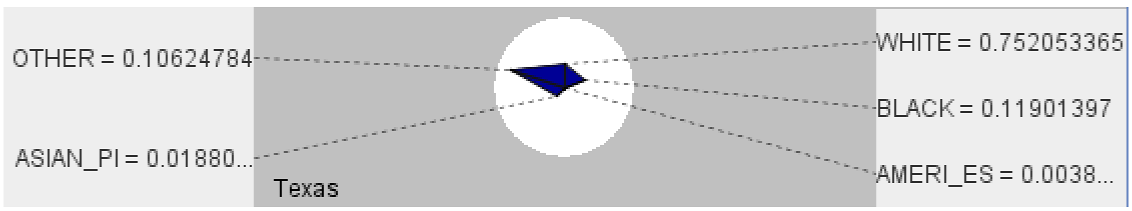


Figure 2. A star plot of three population characteristics and geographic area for the state of Texas (taken from Figure 1). Depicting the values of the five variables along the main axes results in a characteristic shape of a star plot.

Behavioral evaluation of glyphs

Several researchers have investigated the effectiveness of multivariate glyphs in general (Chambers et al., 1983; Friendly, 2008; Nelson & Gilmartin, 1996; Nelson, 2000); fewer have looked at star plots in particular (Lee, Reilly, & Butavicius, 2003; Ware, 2004; Ward, 2002). An increasing number of designed user studies are replacing the large number of evaluations based on researchers' own opinions or ad hoc evaluations (Ward, 2008; Swienty, Reichenbacher, Reppermund, & Zihl, 2008; Griffin, MacEachren, Hardisty, Steiner, & Li, 2006). Ward (2008) summarizes some of this research on star plots. We present a revised overview here.

Lee and coworkers (Lee et al., 2003) evaluated two kinds of representations of binary data: (1) representations that let the data speak for themselves, i.e. those that minimally preprocess data prior to graphical transformation; and (2) representations that preprocess the data in some nontrivial way. In Lee's et al. (2003) study, the first kind of representation includes Chernoff faces and Star plot glyphs; the second kind is called *spatial visualization* in which the *k* dimensions of a data set are reduced to two dimensions and presented in a map-like format (Skupin & Fabrikant, 2007). Lee et al.'s argument in general is that cognitively inspired preprocessing of the data fosters data interpretation and sensemaking. Their research focused on binary data, and their findings indicate superior time and accuracy of answers for spatial visualizations. However, the disadvantage of both Chernoff faces and glyphs that Lee and coworkers found can be easily attributed to the larger number of dimensions that must be compared with each other, in contrast with fewer dimensions in spatial visualizations (Larkin & Simon, 1987). Ward (2008) suggests that some of these problems can be overcome by using special layout techniques for star plots (Ward, 2002, Ward, 2008).

Being truthful to the data is an important step in the data analysis process and hence should be supported prior to application of data aggregation techniques. Our main criticism of the procedure described in Lee et al. (2003) is that it is restricted to binary data and cannot easily be extended to differently scaled data sets.

Borg and Staufenbiel (1992) discuss a variation of data preprocessing prior to graphical representation. They note, for example, that there are no established rules for assigning variables to rays in a star plot and related visual representations (see below). In their experiments, they therefore compare factorial suns with snowflakes (a kind of star plot) and simple suns. Factorial suns preprocess the data using principal component analysis (PCA). The suns are constructed over a two-dimensional PCA vector configuration.

The results of their behavioral experiments are that the grouping of the data presented by factorial suns (i.e. the similarity rating of glyphs) is superior to simple suns and snowflakes. Superiority is established by comparing grouping by participants to a grouping performed by experts. Participants performed grouping solely on the perceptual characteristics of the different glyph representations, not on the semantics of the data.

Other experiments on point symbols representing multivariate data comprise research on Chernoff faces (Chernoff & Rizvi, 1975), or, for example, Chernoff faces compared to other multivariate point symbols (Nelson & Gilmartin, 1996).

Assigning variables to rays

Behavioral research has only marginally addressed the important question of how to assign variables to rays. Quoting Wilkinson (1987), Borg and Staufenbiel (1992) state that there are no established rules for assigning variables to rays in a star plot. Moreover, different assignments can change the interpretation of a glyph. For instance, in an experiment with 96 college students, permuting the assignment of variables to the facial characteristics of Chernoff faces resulted in a 25 percent error rate in a classification task (Chernoff & Rizvi, 1975).

Ward (2008) identifies three approaches for ordering variables/rays in a star plot: correlation-driven, symmetry-driven, and data-driven. The work by Borg and Staufenbiel (1992) mentioned above follows the correlation-driven approach. Ankerst et al. (1998) propose other correlation–driven approaches, including an implemented, heuristic search algorithm that rearranges dimensions to improve overall interpretability (see also Friendly & Kwan, 2003).

The symmetry-driven approach utilizes Gestalt principles (Rock & Palmer, 1990; Ware, 2004) and shape classification to identify simple shapes. Peng et al. (2004) propose two criteria for simplifying the shape of a star plot: monotonicity and symmetry. From these they derive the characteristics of the *perfect* star plot glyph, namely neighboring rays of similar lengths, monotonically increasing or decreasing ray lengths on both sides of an axis, and symmetrically positioned rays of similar length along either the horizontal or vertical axis. The simplest shape according to these criteria would be a *teardrop*. Ward (2008) points out, however, that this approach requires a more formal evaluation as it was not behaviorally assessed. An alternative to searching for such supposedly simplest shape characteristics would be to search for characteristic shapes such as those identified and discussed by Galton (2000).

The data-driven approach (Ward, 2008) allows users to organize the data themselves by choosing a base record, i.e. one entity. For this entity, the organization of variables and rays can be determined and mirrored in the star plots for other entities.

To summarize, within the large body of research on visualization techniques, there has been relatively little behavioral research on evaluating the use of star plot glyphs to represent multivariate data. In particular, few researchers have addressed the question of how to assign variables to rays in star plots. One exception is Peng et al. (2004), who build their approach on Gestalt principles, but do not directly perform behavioral studies. They do suggest the need for an in depth evaluation of the effects of shape on the interpretation of star plots (Ward, 2008).

Behavioral approach

Shape (of objects such as a star plot glyph) is central to processing visual information (Biederman, 2007); to quote Palmer: “Of all the properties we perceive about objects, shape is probably the most important. Its significance derives from the fact that it is the most informative visible property in the following sense: Shape allows a perceiver to predict more facts about an object than any other property.” (Palmer 1999, p. 363) The questions that these observations pose on the design of star plots are threefold:

- Are specific meanings associated with particular shapes (or other perceptual) characteristics of a star plot?
- Is any such meaning independent of the data set that the star plot represents?
- Are there families of star plots distinguished by shape rather than meaning?

To explore the relationship between shape and meaning in greater detail, and specifically for star plot glyphs, we have designed the following experiments.

Experiments

Figure 3 shows three star plots consisting of eight different variables displayed at the same time, i.e. by eight rays. Imagine now the following situation: There are no design guidelines (defined in the literature or elsewhere) on how to assign variables to each ray, hence, the star plots depicted in Figure 3 can potentially be representations of exactly the same data set. Six variables have high values and two variables have low values. Of course, no one would change the assignment of variables to rays in different parts of, for example, a map. Nonetheless, the question remains whether shape influences what information is extracted from star plot maps.

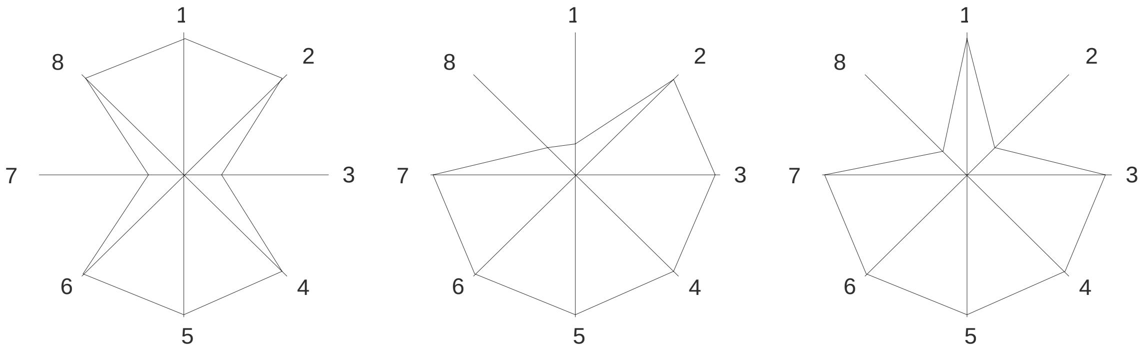


Figure 3. Three star plots that potentially represent the exact same data. Part A has low variable values on the main axes directly opposite each other (i.e. 3 and 7); Part B has low variable values on neighboring axes, i.e. a combination of main and secondary (i.e. 1 and 8). Part C has low variable values on the secondary (diagonal) axes; in the case depicted here, 90 degrees apart (i.e. 2 and 8).

Design and material

We conducted an experiment comparable to a classification task, employing a grouping paradigm to assess the similarity—and thereby the underlying conceptual knowledge structures—of data represented as star plots. Two experimental conditions used fictitious data about future cars (compare Chambers et al., 1983). The participants were advised explicitly that the cars are not current models but design studies. We ascertained that the variation of variables is plausible with respect to the possibility of interaction between variables. The two conditions were identical, except that in condition one the main variation of values was oriented along the major (horizontal and vertical) axes. These rays carry labels 1, 3, 5, 7 (see Figures 3-4). Henceforth, this condition is referred to as *1-3-5-7* for ease of interpretation; we refer to all conditions using this pattern of labels.

In the second condition, only two rays were exchanged so that the main variation occurred along rays 2, 3, 6, and 7; that is, the variables on rays 1 and 5 in condition one (*1-3-5-7*) were moved to rays 2 and 6 in condition two (*2-3-6-7*), respectively (see Table 1 and Figures 3-5). Henceforth, this condition is referred to as 2-3-6-7. The data values were identical in both conditions.

Table 1. Assignment of car data variables to rays in condition *1-3-5-7* and condition *2-3-6-7*.

Variable	Rays in Experiment *1-3-5-7*	Ray in Experiment *2-3-6-7*
Price	1	2
Safety rating	3	3
Miles per Gallon	5	6
Emissions	7	7

Variation along rays was organized as follows. Instead of raw data, variables were normalized and values were assigned in three distinct ranges *high*, *medium*, and *low*. Rays were divided into 100 units each, with 0 and 100 as the lowest and highest possible values, respectively. We assigned the three ranges as follows:

- high: 85-100
- medium: 50-65
- low: 15-30

We maintained a distance of 20 unit points between each of the three ranges to ensure perceptually distinct representations. Within each range, we randomly assigned values using the built-in *Random* function in Excel (which we consider “random enough” for present purposes). Both conditions used identical data values; only the aforementioned assignment of variables to rays changed. We considered all combinatorial possibilities ($3^4 = 81$) of distinct star plots that result from using the three value ranges (high, medium, and low) of four variables (price, safety rating, miles per gallon, and emissions).

We kept the values of the remaining four variables (weight, maximum speed, acceleration, and interior space) qualitatively constant within the *high* range (85-100). By qualitatively constant we mean that these values were randomized within this range—but held constant across conditions—just like the values of other variables (see Figure 6).

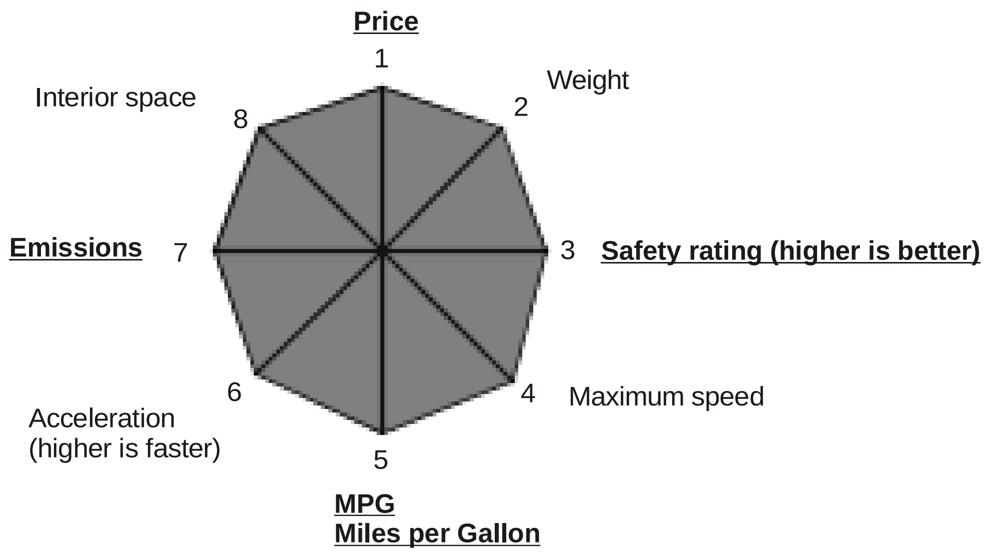


Figure 4. Assignment of variables to rays in condition *1-3-5-7*. Varied variables are located along rays 1 (Price), 3 (Safety rating), 5 (MPG), and 7 (Emissions).

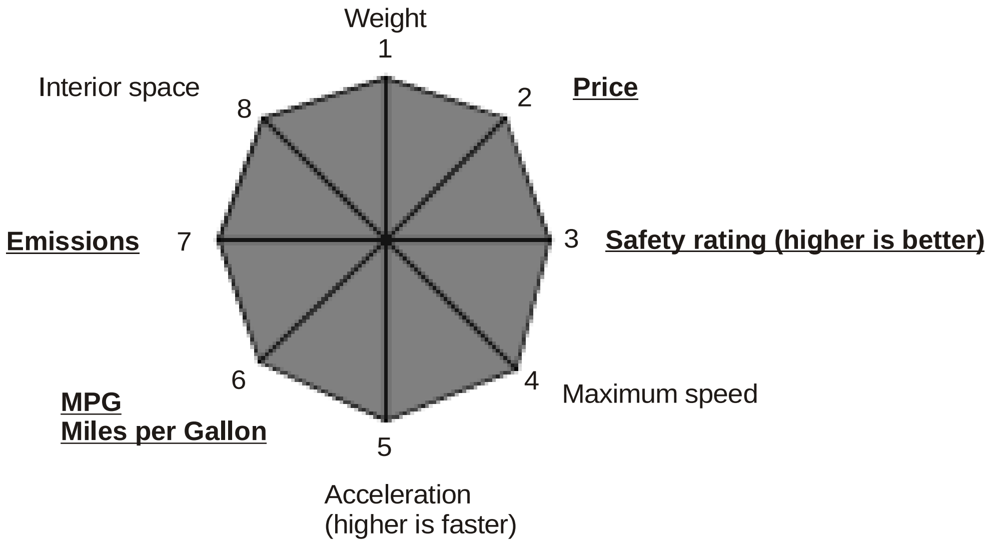


Figure 5. Assignment of variables to rays in a star plot in condition *2-3-6-7*. Varied variables are located along rays 2 (Price), 3 (Safety rating), 6 (MPG), and 7 (Emissions).

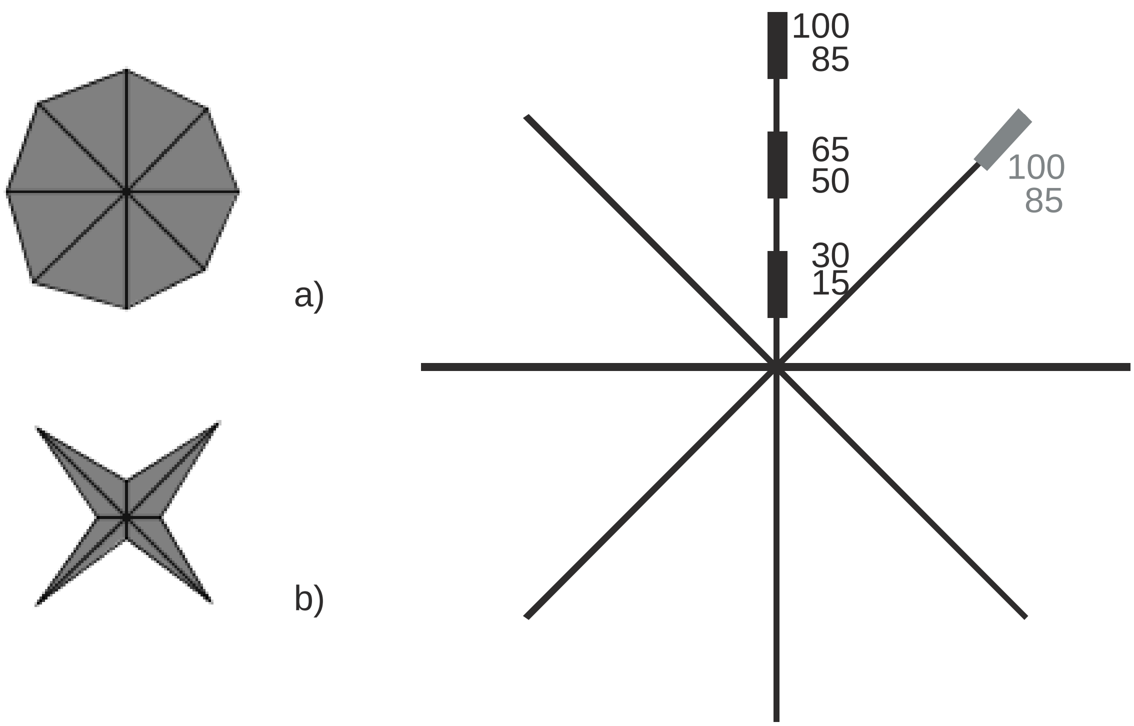


Figure 6. Design of star plots. Left: Two examples in which the values of varied variables are all high (a) and all low (b). Right: The three distinct range categories of high (100-85), medium (65-50), and low (30-15) values for varied variables; unvaried variables were randomly assigned values in the range 100-85 (light gray).

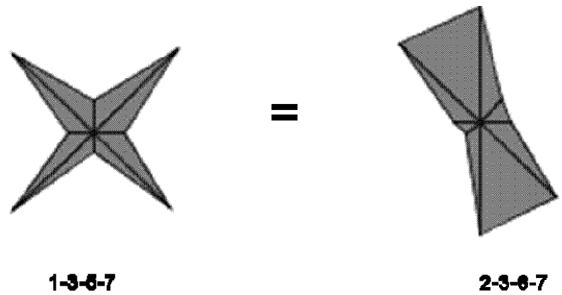


Figure 7. Comparison of the effects of assigning variables to rays. The exact same data (for one car) is represented visually very differently. This effect is based solely on changing 2 rays (1 to 2 and 5 to 6).

Figure 7 gives an example of the effect that the assignment of variables to rays can have. It shows two star plot representations of the same car data values, with the only difference that variables on rays 1 and 5 (condition *1-3-5-7*) are represented on rays 2 and 6 (condition *2-3-6-7*), respectively. The classical star shape disappears and could, instead, be termed *two wedges*.

The employment of tasks in a grouping paradigm is one of the most important methods to elicit conceptual knowledge (Cooke, 1999; Ahn & Medin, 1992; Pothos & Close, to appear). The main idea of such tasks is that conceptual knowledge plays the central role in rating the similarity of a given stimulus. Stimuli are assessed as similar if they are instances of the same concepts, and assessed as dissimilar if they are instances of different concepts. According to Quine (1969: 116): “*there is nothing more basic to thought and language than our sense of similarity; our sorting of things into kinds*”. If other aspects of the stimulus are controlled, grouping experiments can provide insight into the internal structure of conceptual knowledge (Knauff, Rauh, & Renz, 1997; Klippel & Montello, 2007; Mark, 1999; Goldstone, 1994).

Figure 8. Screenshot of the grouping tool user interface, taken during a mimicked experiment under condition *1-3-5-7*. All icons appear initially on the left half of the screen. The participants subsequently place icons into groups on the right half of the screen. Participants decide on the number of groups.

We therefore chose to employ just such a grouping paradigm. However, we address a more specific question: Does the shape of star plots influence the decision whether to place two cars into the same group?

In other words, will experiments under the two conditions exhibit comparable similarity ratings (classification results)? Figures 8 and 9 show screenshots taken during (mimicked) experiments under the two different conditions.

Figure 9. Screenshot of the grouping tool user interface, taken during a mimicked experiment. Here the condition is *2-3-6-7*.

Each of the conditions was followed by additional tasks to shed more light on the underlying rationale for grouping choices. These tasks consisted of labeling and describing groups:

- “Provide a name for the group you created (no more than three words).”
- “Why did you place these star plots in the same group?”

Experiment, Condition 1-3-5-7

Methods

Participants. 20 undergraduates (6 female) from The Pennsylvania State University participated. Their average age was 20.1.

Materials. Design of icons (see above). Icons were displayed using the grouping tool; Figure 8 provides a screenshot of a typical experiment in progress. The tool divides the screen into two parts. On the left side, the stimulus material consisting of 81 star plots were placed in a random order for each participant. The large number of icons required scrolling to access all items. (Scrolling is a common procedure in interaction with computer interfaces; problems were neither expected nor found during the experiments.) The right side of the screen was empty at the start; participants moved icons to this side in order to group them during the experiment. The interface was kept simple so that participants could perform only the following three actions: Finish, New Group, Delete Active Group.

Procedure. The experiments took place in a GIScience laboratory at the Geography Department at Penn State. Participants were tested in groups. The computers (Dell, Pentium D, 3 GHz, 2 GB RAM) in the lab have 20-inch wide-screen Dell monitors, suitable for grouping tasks. To ensure individuality of results, the lab was prepared with view blocks, with computers arranged so that participants could not see the screens of other participants.

After arriving and providing consent, participants were introduced to the general concept of star plots. For this purpose, we created an explanation sheet that participants could refer to during the experiment. We placed these sheets in front of the computer screens. After this introduction, participants entered their personal data into the computer.

The grouping tool provides participants with additional instructions on the star plot grouping task. An introduction to the task used animal icons to familiarize participants with the grouping tool such as creating new groups and to give them an example of creating similarity ratings from a different domain. After this warm-up task, participants were shown 81 star plots (on the left side of the screen) that they had to group on the right side of the screen according to the similarity rating they considered appropriate (see Figures 6 and 7). Participants were advised explicitly that there is no right or wrong way of grouping the icons.

After the main task, participants performed a second task in which they were asked to: (1) label the groups they created, and (2) provide a rationale for their grouping. The groups the participants created were shown to them one at a time. The results of this part of the experiment are not discussed in this paper.

Results

For each participant, grouping results in an individual similarity matrix. Matrix size corresponds to the number of star plot glyphs used; each symmetric matrix encodes all possible similarity ratings of star plot glyphs taken pairwise. Of the 81x81=6561 total cells, n*(n-1)/2=3240 cells are relevant for the similarity assessment. Individual similarity is binary encoded; any pair of star plot glyphs is coded as true (1) if the participant placed them in the same group, or false (0) if not. An overall similarity matrix, across all participants, is the sum of individual similarity matrices. For example, if 20 participants place two icons (called A and B) into the same group, the overall similarity matrix encodes a similarity of 20x1 = 20 at positions AB and BA. The maximum overall similarity rating is thus equal to the number of participants.

To identify and analyze categories within the grouping results, we subjected the overall similarity matrix to a hierarchical cluster analysis. Cluster analysis allows for identifying "natural" groupings within data by minimizing within-group and maximizing between-group variation (Aldenderfer & Blashfield, 1984). Agglomerative cluster analysis starts with each entity (here star plot glyphs) in its own group, i.e. 81 individual clusters, then recursively combines the most similar clusters as determined by a clustering method. We used the software CLUSTAN, squared Euclidean distance as a similarity measure, and we compared different clustering methods. However, we will only briefly discuss the clustering structure resulting from Ward's method (see below) as the main purpose lies not within identifying clusters but to compare the two conditions.

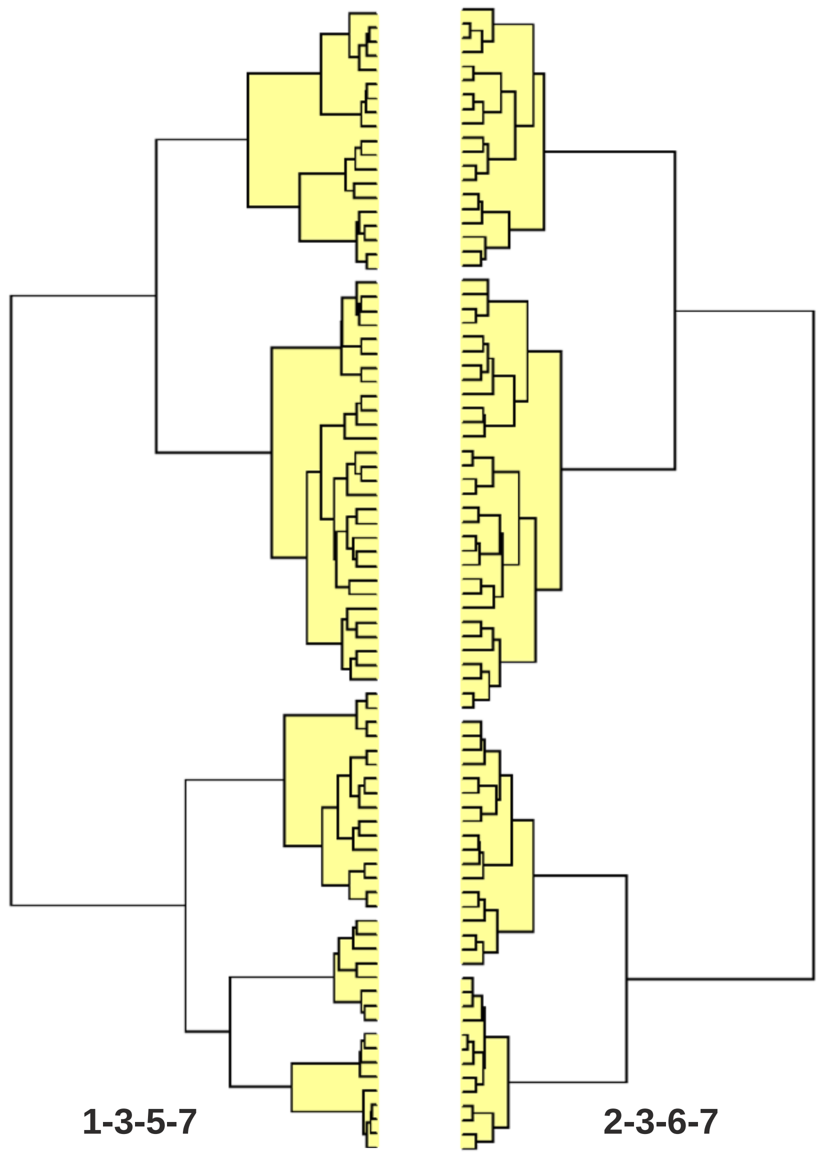


Figure 10. Clustering structure in the overall similarity matrix from condition *1-3-5-7* (left) and *2-3-6-7* (right), using Ward’s method. Corresponding 5- and 4-cluster solutions are shown here.

As an example, Figure 10 shows the result of applying Ward's method to the overall similarity matrix from the two conditions. While the identification of clusters would normally be the main priority, it is only one step in our analysis as the main goal is the comparison of the results of the two conditions (*1-3-5-7* and *2-3-6-7*).

Experiment, Condition 2-3-6-7

Methods

Participants. 20 participants (6 female) from The Pennsylvania State University participated. Their average age was 19.5.

Materials and Procedure. Except for the design of the star plots, this condition was identical. We randomly assigned students to one of the two conditions, but did so as to keep the number of female participants approximately equal in both conditions.

Results

The analysis of data collected under condition *2-3-6-7* follows the same patterns as under condition *1-3-5-7*. Figure 10 shows an exemplary result of applying Ward's method to the overall similarity matrix from condition *2-3-6-7*.

Overall Results – Comparing Conditions 1-3-5-7 and 2-3-6-7

Correlation of similarity values. To get a first idea of whether the two similarity matrices are similar or different, we performed a correlation analysis on individual similarity values obtained from both experiments for star plots representing the same data. That is, each of the 3240 cells provides a similarity value for the star plot pairs in both experiments. The values for this similarity can range from 0 to 20 (0 meaning that two star plots were never placed into the same group and 20 meaning that all participants placed the two star plots into the same group). We performed a Spearman's rank correlation analysis (see Figure 11). There was a significant positive correlation between the similarity values under conditions *1-3-5-7* and *2-3-6-7* ($r = .581$, $N = 3240$, $p < .0005$, two-tailed).

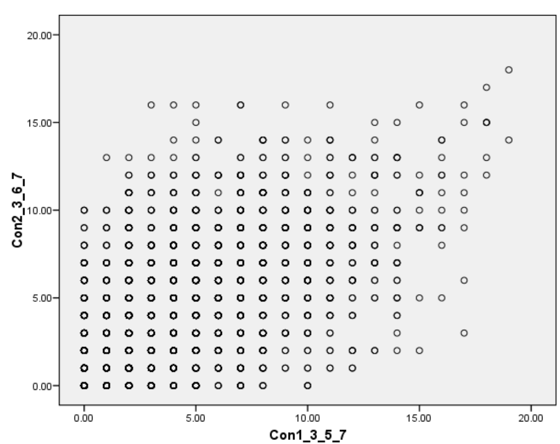


Figure 11. Correlation between similarity ratings extracted from the overall similarity matrix for conditions *1-3-5-7* and *2-3-6-7*.

Time to complete the task. We hypothesized that participants in condition *1-3-5-7* would be faster than participants in condition *2-3-6-7*. The gain in speed could be attributed to the fact that in condition *1-3-5-7*, the variation is taking place along the two main axes, which results to some extent in more characteristic shapes. These characteristic shapes might offer an advantage in identifying similar shapes. An analysis of the time participants required to complete the task showed that participants in condition *1-3-5-7* (mean time: 669.93 seconds) were indeed faster than participants in condition *2-3-6-7* (mean time: 786.26 seconds), resulting in a mean difference of nearly two minutes (116.33 seconds). However, an independent t-test turned out to be not significant but showing a statistical trend (t = -1.533, df = 38, p = .068, one-tailed).

Number of groups. We tested whether the number of groups differed between participants, as an indication of whether certain shape characteristics might have led to the creation of more distinct categories in condition *1-3-5-7* than in condition *2-3-6-7*. Participants in both experiments created approximately the same number of groups (categories): 7.5 in *1-3-5-7* versus 7.2 in *2-3-6-7*. The difference is statistically not significant (t = .241, df = 38, p = .811, two-tailed).

In-depth analysis of cluster structure. While the more classical statistical analyses did not indicate statistically significant differences, with the exception of a statistical trend in time to task completion between conditions *1-3-5-7* and *2-3-6-7*, the dendrograms in Figure 10 indicates a somewhat different clustering structure. To grasp these differences visually, we performed a multidimensional scaling analysis (MDS) using CLUSTAN and wrote a tool that allows for visually displaying the results (see Figures 12 and 13). A visual inspection of the MDS plots revealed that there are differences in the clustering structure for specific visually salient star plots. Some star plots that show a characteristic shape, such as those having 1 or 2 'spikes' in condition *1-3-5-7*, cluster together even though they are semantically different because they represent quite different data (see circled star plots in Figure 12). Figure 14 shows the four star plots with one spike in condition *1-3-5-7* and the corresponding star plots in condition *2-3-6-7*.

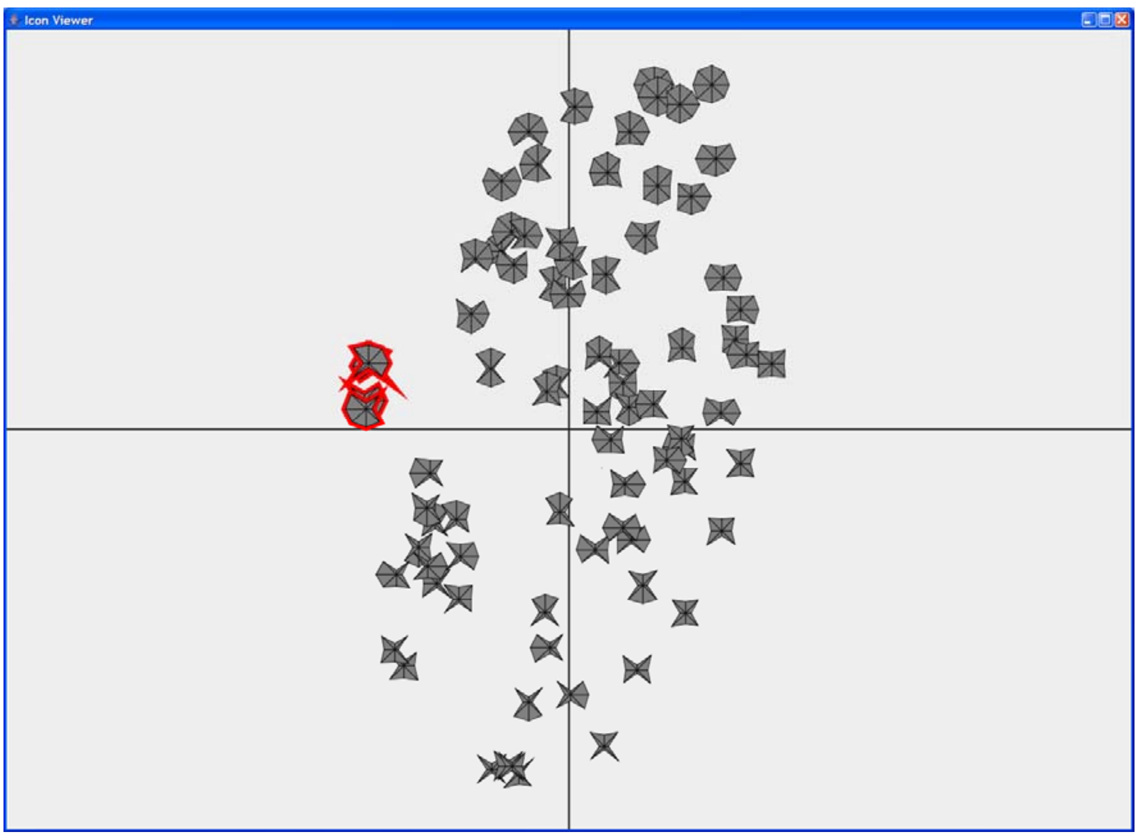


Figure 12. MDS plot of the clustering structure of condition 1-3-5-7. The red circled star plot glyphs marks a group of star plot glyphs that exhibit one spike (see also Table 2a).

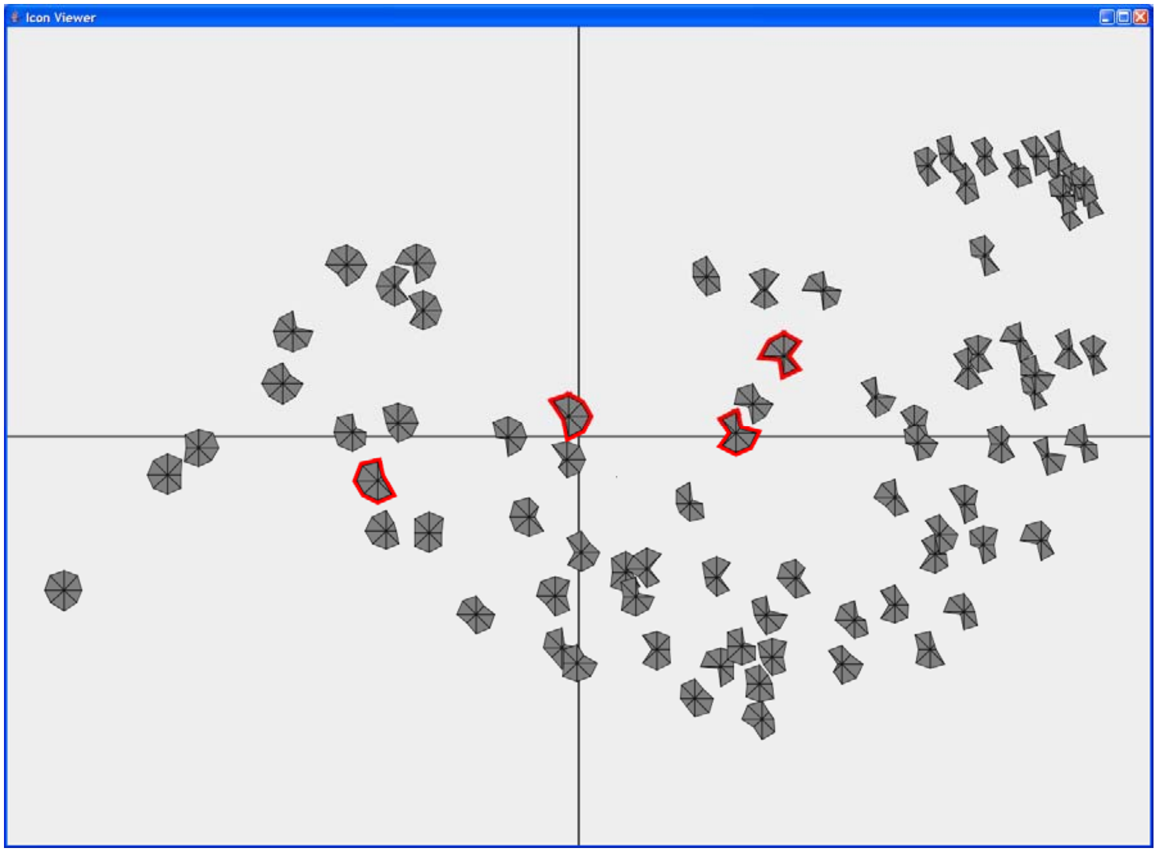


Figure 13. MDS plot of the clustering structure of condition 2-3-6-7. The red circled star plot glyphs correspond to the red circles ones in Figure 12 (see also Table 2a).

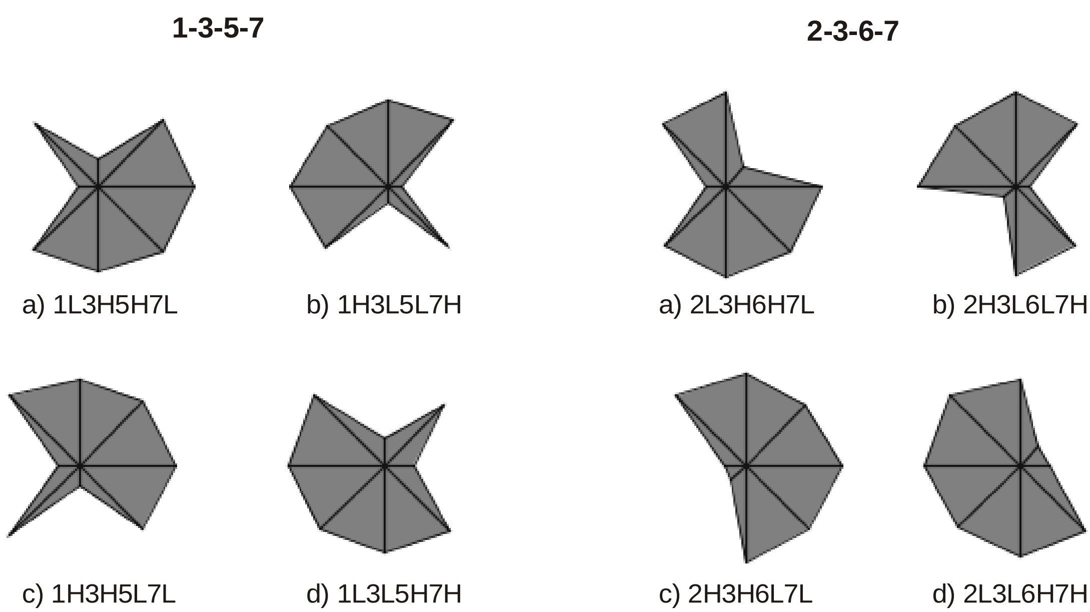


Figure 14. Representations of four pairwise-identical data entries in conditions *1-3-5-7* and *2-3-6-7* (indicated by letters a, b, c, d).

Visual identification and analysis of differences between conditions (*1-3-5-7* vs. *2-3-6-7*). To analyze these differences in greater detail, we used the Improvise visualization development environment (Weaver, 2004) to build an interactive visual analysis tool called KlipArt. The tool displays a graph consisting of a node for each participant plus a node for each *unique* grouping of icons produced by at least one participant. Edges connect each participant to their respective groups. Bubble-like 'packs' encompass the grouping nodes of each participant. The graph supports interactive dragging of any structural element as well as toggling of an iterative spring-based layout algorithm, allowing the experimenter to tease apart even complex grouping relationships. The window is split into two halves allowing for a direct comparison of conditions *1-3-5-7* and *2-3-6-7*. Which icons and which participants are displayed is determined by the user. In our case, we used the insights from the previously discussed cluster analysis and MDS plots to guide our explorations.

Figure 15 shows one interactive state in a visual analysis of the two sets of four star plot glyphs in Figure 14. The graph tells a clear story that we will briefly describe here. The number of participants who grouped all four icons together is very different between the two conditions: 14 participants in condition *1-3-5-7* compared to only two participants in condition *2-3-6-7* (see the upper left corners of the two halves of Figure 16). We will not go into further detail here regarding the analysis tool but instead demonstrate the differences (and similarities) found in both conditions by looking at four shape families in each condition that we identified using the MDS plots and the visual tool.

We analyzed several shape families and exemplarily depict four in Table 2. We used a chi-square analysis on the derived counts of how many participants in each condition placed all selected icons of a shape family into the same group. The results together with the star plot glyphs are summarized in Table 2. They show clearly that the visually salient shapes in condition *1-3-5-7* led to statistically significant differences in the grouping behavior compared to condition *2-3-6-7* (see star plot glyphs with one spike (a)). These differences disappear when the shape differences become less prominent (b) and when the data these shapes represent become less distinct (c). It is interesting to note though that shape salience can go both ways: An example is depicted in row (d). Here condition *2-3-6-7* produces barely distinguishable shapes while condition *1-3-5-7* creates shape differences that lead to a reversed effect. More participants in condition *2-3-6-7* (12) placed these four icons into the same group compared to only three participants in condition *1-3-5-7*.

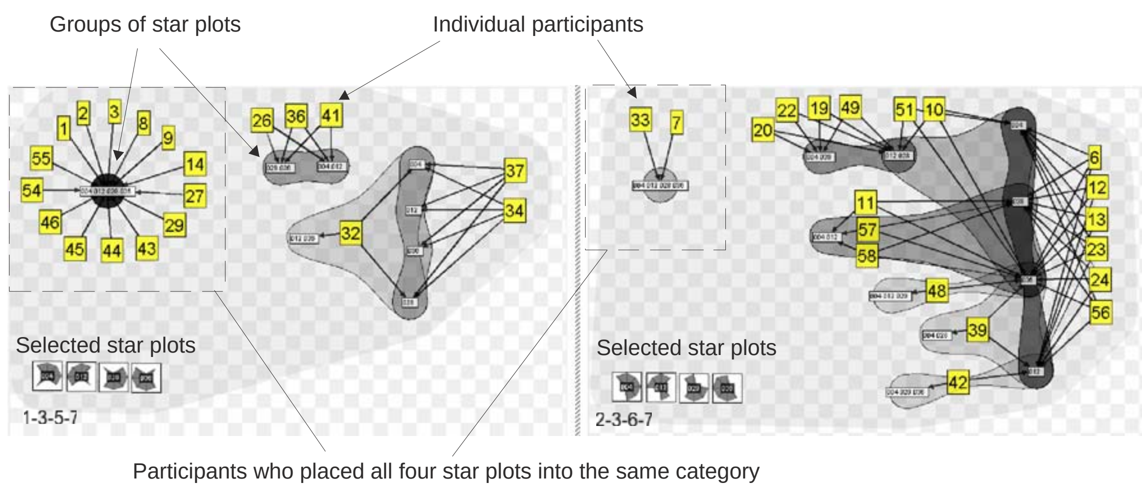
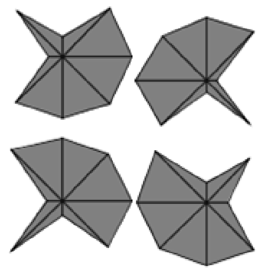
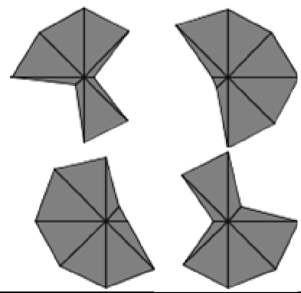
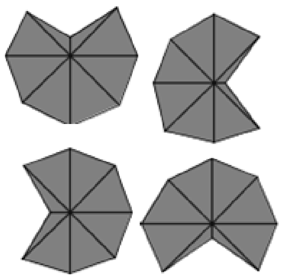
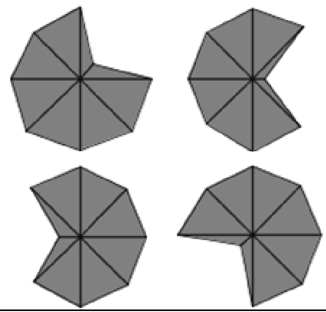
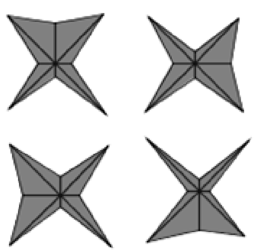
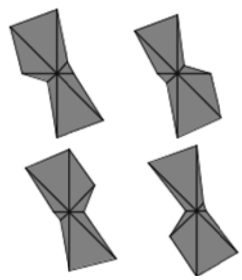
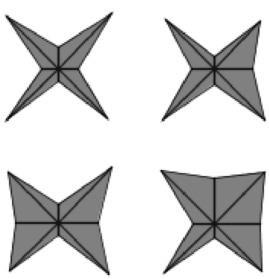
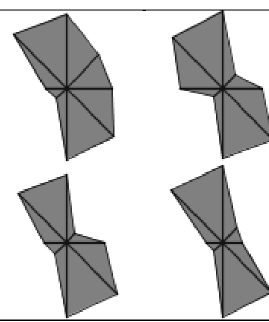


Figure 15. Interactive interface for comparative visual exploration of clustering structure under conditions *1-3-5-7* (left) and *2-3-6-7* (right). Highlighted (dashed boxes) are the two sets of four star plot glyphs in Figure 15 (see also Table 2a).

Table 2. Comparison of how often characteristic star plots, i.e. those with characteristic shapes are placed into the same group in conditions *1-3-5-7* and *2-3-6-7*.

	Star Plots, *1-3-5-7*	Star Plots, *2-3-6-7*	Chi Square Test
a			χ^2 = 15, 1, p < .000
b			χ^2 = 2.51, 1, p. = .11
c			χ^2 = .476, 1, p. = .49
d			χ^2 = 8.64, 1, p. = .003

Conclusions and Outlook

We started our research with the question of how shape influences the classification of data represented by multivariate point symbols, focusing in this case on star plot glyphs. This research is important as the availability and development of visualization tools is increasing rapidly but proper guidelines for designing interfaces are rare and more behavioral research is constantly called for (Ward, 2008; Montello, 2004). The results of our behavioral experiments reveal a diverse and interesting picture. On the global level, few differences surfaced by comparing the number of groups that were created in both conditions (i.e. star plot glyphs with the main variation occurring on rays 1-3-5-7 and star plot glyphs with the main variation occurring on rays 2-3-6-7) and the correlations of the similarity values in both conditions. These results may have been taken as an indication that the shape of the star plots is less influential than we originally hypothesized.

On the other hand, we found a time advantage for condition *1-3-5-7*, even without specific instruction to participants to perform the task as quickly as possible. This condition, with more characteristically shaped star plots (as a result of the main variation taking place on the horizontal and vertical axes) required less time to finish. This time advantage can be seen as a first indication of the differences in cognitive processing between the two conditions.

Additionally, the clustering structure (Figure 10) and the visual representation of similarities using multidimensional scaling (Figures 12 and 13) both revealed differences in the grouping behavior of participants. These differences appear to center on perceptually salient shape features, especially in condition *1-3-5-7*. The KlipArt visual tool is a step forward in analyzing grouping peculiarities in great detail through direct identification of subgroups and comparison of participant grouping behaviors under different experimental conditions in terms of visually salient shapes. A subsequent analysis of the data obtained through KlipArt showed that these differences were statistically significant.

Using this combination of tools and visualization techniques, we have been able to reveal that effects shown to be important for Chernoff faces also hold for star plots, namely the assignment of variables to graphical (facial/ray) characteristics (Chernoff & Rizvi, 1975). This result is a clear call for guidelines on how to assign variables to rays in a star plot, and a first step in a more formal evaluation of the influence of shape characteristics, as called for by Ward (2008).

Given a large body of literature on the salience of shapes in the area of visual search (Hulleman, te Winkel, & Boselie, 2000) and shape change detection (Barenholtz, Cohen, Feldman, & Singh, 2003), Peng et al. (2004) may be right in speculating that the *teardrop* is the simplest shape overall. However, for the detection of differences between star plot glyphs and for breaking down a star plot into subparts for comparison (Hoffman & Singh, 1997), as is necessary in multidimensional data spaces, this simplicity might not be an advantage. The concept of *feature salience* for Chernoff faces as discussed by Nelson (2007) needs to be extended to star plot glyphs, too. The above quoted literature indicates clearly that, for example, concavities have a perceptual advantage over convexities. A teardrop, as the acclaimed most perfect shape, does not have concavities and may therefore be at a disadvantage (although from a symmetry and monotonicity point of view it certainly does seem to be appealing). The shape originally lending its name to our star plots, i.e. the star, may be, from the perspective of comparison-speed and detecting changes in a data set, the better choice. In its most prominent shape it shows pronounced concavities (as well as convexities). It additionally is able to comply with the symmetry criterion and exhibits distinguishable parts (Biederman, 1987).

In our experiment, the shapes resulting from condition 1-3-5-7 show more concavities than in condition 2-3-6-7 (as well as more pronounced convexities). It may be, therefore, that not only the ‘spikes’ contributed to the faster sorting of the icons but also the overall uniqueness of a combination of concavities and convexities. A first recommendation for designing star plot glyphs, for example, in star plot maps, would therefore be to arrange the rays representing variables such that the main variation occurs along the main axes (in case of 8 variables). In other words, the assignments of variables to rays should aim for creating star shaped glyphs instead of teardrop shaped ones.

The downside of perceptually salient shapes seems to be, however, that they do indeed influence the grouping behavior (i.e. the classification) of data represented by star plots. This influence was revealed by our analysis of ‘has one spike’ and other salient shapes using the KlipArt tool. While not completely surprising, this result raises the question of how best to support analysis of multivariate data sets using point-like symbols but not distracting the user from the data by using a specific way of visually representing the data (Monmonier, 1999).

One question more intensely discussed in recent research is that of individual differences (e.g., Coluccia & Louse, 2004; Aubrey, Li, & Dobbs, 1994; Levinson, 2003). Individual differences can occur as general group differences, such as sex, age, or visual impairment, or they can be established on the basis of tests, such as spatial abilities (Hegarty & Waller, 2005). Orthogonal to individual differences though, the human cognitive (perceptual) system exhibits characteristics that most human beings share (e.g., Pinker, 1990; Marr, 1982; Lakoff & Johnson, 1980). An example from the cartographic realm is the early discussion of visual clutter (Phillips, 1979; Dobson, 1980). It can be assumed universal that visual clutter negatively influences information processing. In the case of data represented by star plot glyphs we need to make a distinction between the shape characteristics and the data semantics. Shape, just like visual clutter seems to be universally processed and a salient shape will be a salient shape across individuals. The data that is represented by a star plot glyph, however, might be subject to individual preferences. In the studies presented in this paper we held the semantic aspect constant and focused on the effects of different shapes. On open, yet to explore question would be to whether a particular shape has the same effect across different semantic domains.

One avenue of our ongoing research will seek to address this very question. How can salient shape characteristics be enhanced to (a) allow for the use of salient shapes, yet (b) avoid negative influences of shape characteristics such as 'has one spike'? We have designed a follow-up study in which we will color-enhance star plots. Instead of simply displaying all rays in the same color, rays will be color coded according to a qualitative color scheme. In this manner, we hope to enhance perceptual differences between star plots without sacrificing the benefits of shape characteristics alone. We plan to compare the results of this study with those obtained thus far.

A second avenue for our research will consider a core geographic topic, namely the question of conceptual spatial autocorrelation. How deeply rooted is Tobler's first law (Tobler, 1970) in the human cognitive-perceptual system (Montello, Fabrikant, Ruocco, & Middleton, 2003)? To address this problem we will place star plots on a map just like in Figure 1, but instead of sorting the symbols into groups, participants will be asked to classify icons without moving them, i.e. by annotating the map directly. Our expectation is that this will shed light on the importance of place (Sinton & Lund, 2007) as well as provide a way to explore perceptual interaction of star plots with surrounding star plots.

Reference List

Adrienko, G., Adrienko, N., Dykes, J., Fabrikant, S. I., & Wachowicz, M. (2008). Geovisualization of dynamics, movement and change: Key issues and developing approaches in visualization research. *Information Visualization, 7*(3/4).

Ahn, W.-K., & Medin, D. L. (1992). A two-stage model of category construction. *Cognitive Science, 16*, 81–121.

Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster analysis.* Newbury Park, CA: Sage.

Ankerst, M., Berchtold, S., & Keim, D. A. (1998). Similarity clustering of dimensions for an enhanced visualization of multidimensional data. In *Proceedings of 1998 IEEE Symposium on Information Visualization* (pp. 52-60, 153). Los Alamitos, CA: IEEE Computer Society.

Arnheim, R. (1976). The perception of maps. *The American Cartographer, 3*, 5–10.

Aubrey, J., Li, K., & Dobbs, A. R. (1994). Age differences in the interpretation of misaligned "you-are-here" maps. *Journal of Gerontology, 49*(1), 29–31.

Barenholtz, E., Cohen, E. H., Feldman, J., & Singh, M. (2003). Detection of change in shape: An advantage for concavities. *Cognition, 89*, 1–9.

Bertin, J. (1974). *Graphische Semiologie: Diagramme, Netze, Karten.* Berlin, New York: de Gruyter.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*, 115–145.

Biederman, I. (2007). Recent psychophysical and neural research in shape recognition. In N. Osaka, I. Rentschler, & I. Biederman (Eds.), *Object Recognition, Attention, and Action .* Berlin: Springer.

Bollmann, J. (1981). *Aspekte kartographischer Zeichenwahrnehmung. Eine empirische Untersuchung.* Bonn: Kirschbaum.

Borg, I., & Staufenbiel, T. (1992). Performance of snowflakes, suns, and factorial suns in the graphical representation of data. *Multivariate Behavioral Research, 27*(1), 43–55.

Brewer, C. A., & Pickle, L. (2002). Evaluation of methods for classifying epidemiological data on choropleth maps in series. *Annals of the Association of American Geographers, 92*(4), 662–681.

Chambers, J. M., Cleveland, W. S., Kleiner, B., & Tukey, P. A. (1983). *Graphical methods for data analysis. The Wadsworth & Brooks/Cole statistics/probability series.* Belmont, CA: Wadsworth.

Chang, K.-T., Antes, J. R., & Lenzen, T. (1985). The effect of experience on reading topographic relief information: Analysis of performance and eye movements. *Cartographic Journal, 22*(2), 88–94.

Chernoff, H., & Rizvi, M. H. (1975). Effect on classification error of random permutations of features in representing multivariate data by faces. *Journal of the American Statistical Association, 70*(351), 548–554.

Coluccia, E., & Louse, G. (2004). Gender differences in spatial orientation: A review. *Journal of Environmental Psychology, 24*(3), 329–340.

Cooke, N. (1999). Knowledge elicitation.: Applied Cognition. In F. T. Durso (Ed.), *Handbook of Applied Cognition .* Chichester: John Wiley & Sons.

Dobson, M. W. (1980). The influence of the amount of graphic information on visual matching. *The Cartographic Journal, 17*(1), 26–32.

Eckert, M. (1921). *Die Kartenwissenschaft. Forschungen und Grundlagen zu einer Kartographie als Wissenschaft. 1.* Berlin, Leipzig.

Friendly, M. (2008). A brief history of data visualization. In C.-h. Chen, W. Härdle, & A. Unwin (Eds.), *Springer handbooks of computational statistics. Handbook of data visualization* (pp. 15–56). Berlin: Springer.

Friendly, M., & Kwan, E. (2003). Effect ordering for data displays. *Computational Statistics & Data Analysis, 43*, 509–539.

Galton, A. (2000). *Qualitative spatial change. Spatial information systems.* Oxford: Oxford Univ. Press.

Goldstone, R. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General, 123*(2), 178–200.

Gore, A. (1998). *The digital earth: Understanding our planet in the 21st century.* Retrieved January 07, 2008, from http://www.isde5.org/al_gore_speech.htm.

Griffin, A. L., MacEachren, A. M., Hardisty, F., Steiner, E., & Li, B. (2006). A comparison of animated maps with static small-multiple maps for visually iIdentifying space-time clusters. *Annals of the Association of American Geographers, 96*(4), 740–753.

Harley, J. B., & Woodward, D. (Eds.) (1987). *History of cartography, Vol. 1: Cartography in prehistoric, ancient, and medieval Europe and the Mediterranean.* Chicago, IL: University of Chicago Press.

Harrower, M., & Brewer, C. A. (2003). ColorBrewer.org: An online tool for selecting color schemes for maps. *The Cartographic Journal, 40*(1), 27–37.

Hegarty, M., & Waller, D. (2005). Individual differences in spatial abilities. In P. Shah & A. Miyake (Eds.), *Handbook of Visuospatial Cognition* (pp. 121–169). Cambridge, MA: Cambridge University Press.

Hoffman, D. D., & Singh, M. (1997). Salience of visual parts. *Cognition, 63*, 29–78.

Hulleman, J., te Winkel, W., & Boselie, F. (2000). Concavities as basic features in visual search: Evidence from search asummetries. *Perception & Psychophysics, 62*(1), 162–174.

Klippel, A., & Montello, D. R. (2007). Linguistic and nonlinguistic turn direction concepts. In S. Winter, B. Kuipers, M. Duckham, & L. Kulik (Eds.), *Spatial Information Theory. 9th International Conference, COSIT 2007, Melbourne, Australia, September 19-23, 2007 Proceedings* (pp. 354–372). Berlin: Springer.

Knauff, M., Rauh, R., & Renz, J. (1997). A cognitive assessment of topological spatial relations: Results from an empirical investigation. In S. C. Hirtle & A. U. Frank (Eds.), *Lecture notes in computer science. 1329. Spatial information theory: A theoretical basis for GIS* (pp. 193–206). Berlin: Springer.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago: University of Chicago Press.

Larkin, J., & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 82*(4), 207–225.

Lee, M. D., Reilly, R. E., & Butavicius, M. A. (2003). An empirical evaluation of Chernoff faces, star glyphs, and spatial visualizations for binary data. In T. Pattison & B. Thomas (Eds.), *Proceedings of the Australian Symposium on information visualization .* Sydney: Austrian Computer Society (OCG).

Levinson, S. C. (2003). *Space in Language and Cognition: Explorations in Cognitive Diversity.* Cambridge: Cambridge University Press.

lNelson, E. S. (2007). The face symbol: Research issued and cartographic potential. *Cartographica, 42*(1), 53–64.

MacEachren, A. M. (1995). *How maps work: Representation, visualization, and design.* New York: The Guilford Press.

Mark, D. M. (1999). Spatial representation: A cognitive view. In D. J. Maguire, M. F. Goodchild, D. W. Rhind, & P. A. Longley (Eds.), *Geographical information systems: Principles and applications. Second edition, v. 1* (pp. 81–89).

Marr, D. (1982). *Vision.* New York: Freeman.

Monmonier, M. (1999). *How to lie with maps* (2. ed). Chicago: Univ. 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.

Montello, D. R. (2004). Cognition of geographic information. In R. McMaster & L. E. Usery (Eds.), *A Research Agenda for Geographic Information Science* (pp. 61–91). Boca Raton, FL: CRC Press.

Montello, D. R., Fabrikant, S. I., Ruocco, M., & Middleton, R. S. (2003). Testing the first law of geography on point-display spatializations. In W. Kuhn, M. Worboys, & S. Timpf (Eds.), *Spatial Information Theory: Foundations of Geographic Information Science. International Conference, COSIT 2003, Ittingen, Switzerland, September 24-28, 2003, Proceedings* (pp. 335–351). Berlin: Springer.

Nelson, E. S. (2000). Designing effective bivariate symbols: The influence of perceptual grouping processes. *Cartography and Geographic Information Science, 27*, 261–278.

Nelson, E. S., & Gilmartin, P. P. (1996). An evaluation of multivariate, quantitative point symbols for maps. In C. H. Wood & C. P. Keller (Eds.), *Cartographic design: Theoretical and practical perspectives* (pp. 191–203). Chichester, New York, Brisbane u. a.: John Wiley & Sons.

Nightingale, F. (1858). *Notes on Matters Affecting the Health, Efficiency and Hospital Administration of the British Army. Founded Chiefly on the Experience of the Late War:* Presented by Request to the Secretary of State for War. Privately printed for Miss Nightingale, Harrison and Sons.

Olson, J. M. (1984). Cognitive issues in map use. *International Yearbook of Cartography, 24*, 151–157.

Peng, W., Ward, M. O., & Rundensteiner, E. A. (2004). Clutter reduction in multi-dimensional data visualization using dimension reordering. In *Proceedings of 2004 IEEE symposium on information visualization .* Los Alamitos, CA: IEEE Computer Society.

Phillips, R. J. (1979). Making maps easy to read: A summary of research. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), *Processing of visual language. Vol. 1* (pp. 165–174). New York: Plenum Press.

Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial intelligence and the future of testing* (pp. 73–126). Hillsdale, N. J.: Lawrence Erlbaum.

Pothos, E. M., & Close, J. (to appear). One or two dimensions in spontaneous classification: A simplicity approach. *Cognition*,

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

Robinson, A. H., & Petchenik, B. B. (1976). *The nature of maps.* Chicago, London: The University of Chicago Press.

Rock, I., & Palmer, S. E. (1990). The legacy of gestalt psychology. *Scientific American, 263*(6), 48–61.

Sinton, D. S., & Lund, J. J. (2007). *Understanding place: GIS and mapping across the curriculum /* (1. ed.). Redlands, Calif: ESRI Press.

Skupin, A., & Fabrikant, S. I. (2007). Spatialization. In J. Wilson & A. S. Fotheringham (Eds.), *Blackwell companions to geography: Vol. 7. The handbook of geographic information science* (pp. 61–79). Malden, MA: Blackwell.

Slocum, T. A., McMaster, R. B., Kessler, F. C., & Howard, H. H. (2005). *Thematic cartography and geographic visualization* (2nd ed). *Geographic information science.* Upper Saddle River, N.J.: Pearson Prentice Hall.

Swienty, O., Reichenbacher, T., Reppermund, S., & Zihl, J. (2008). The role of relevance and cognition in attention-guiding geovisualisation. *The Cartographic Journal, 45*(3), 227–238.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography, 46*, 234–240.

Ward, M. O. (2002). A taxonomy of glyph placement strategies for multidimensional data visualization. *Information Visualization, 1*(3-4), 194–210.

Ward, M. O. (2008). Multivariate Data Glyphs: Principles and Practice. In C.-h. Chen, W. Härdle, & A. Unwin (Eds.), *Springer handbooks of computational statistics. Handbook of data visualization* (pp. 179–198). Berlin: Springer.

Ware, C. (2004). *Information visualization: Perception for design /* (2. ed). *The Morgan Kaufmann series in interactive technologies.* San Francisco, Calif.: Morgan Kaufmann.

Weaver, C. (2004). Building highly-coordinated visualizations in improvise. In *Proceedings of the IEEE Symposium on Information Visualization 2004, Austin, TX, October 2004 .*

Wilkinson, L. (1987). *Statistics package SYSTAT-SYGRAPH.* Evanston, IL: Systat.

Wilkinson, L. (2005). *The grammar of graphics* (2nd ed.): Springer.