Color Enhanced Star Plot Glyphs – Can Salient Shape Characteristics be Overcome?[1]

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

GeoVISTA Center, Department of Geography, Pennsylvania State University, PA, USA[a]

Dutton e-Education Institute, Pennsylvania State University, PA, USA[b]

Center for Spatial Analysis and the School of Computer Science, University of Oklahoma, OK, USA[c]

Abstract. This paper reports two experiments that address the question of how the shape characteristics of star plot glyphs influence classification tasks and whether additional graphic features such as color can be used to counterbalance the effects of shape characteristics. In a previous study (Klippel, Hardisty, & Weaver, to appear, 2009) we found that salient shapes of star plot glyphs such as "has one spike" influence the classification of the data represented by star plot glyphs compared to star plot glyphs that represent the same data but do not exhibit salient shape characteristics. Hence, we found a similar phenomenon that has been reported for Chernoff faces: that assigning variables to facial characteristics differently can change the outcome of classifications tasks. The shape differences in star plot glyphs are induced by assigning variables to rays in different ways. While the first two experiments showed shape influences, we conducted two follow up studies to shed more light on the influence of shape. A third experiment (the first new experiment) simply addressed the question of how the classification of star plot glyphs would be affected if they are stripped of their meaning, that is, instead of classifying glyphs which represented data, participants were asked to group star plot glyphs as shapes. The forth experiment (the second new experiment) used the same icons as experiment one (from the previous study) and experiment three. This time the star plot glyphs were meaningful; they represented data but additionally, the rays in the star plots were color coded to focus attention on differences in salient shapes. The shape characteristic "has one spike" as a general characteristic cannot be applied as easily as all spikes have different colors. The results show that color enhanced star plot glyphs have positive effects on the processing speed and that they reduce the influence of salient shape characteristics.

[1] Accepted by *Cartographica*, to be published 2009.

Introduction

The design of graphic representations of both spatial and non-spatial data is a challenging endeavor. Perhaps surprisingly, it is not getting less challenging in the light of new information technologies. One of the main questions to solve is how to select the ‘best’ or ‘most correct’ representation for a given data set from the perspective of supporting human information processing. For some areas, such as the groundbreaking work of Bertin (1983) we have found solutions that describe the assignment of data types to specific visual variables to optimally support the interpretation of data.

For other areas, however, we are still facing the challenge of a one-to-many relationship between a data set and potential graphical representations (see Figure 1). This problem can be further differentiated into macro decisions and micro decisions. For example, in the case of multivariate point symbols, a macro decision would address the question of which (point) symbol to use to represent the data. In this area we find several behavioral studies that address this question. For example Nelson and Gilmartin (1996) compared the use of Chernoff faces (Chernoff, 1973) to other multivariate point symbols such as boxed letters. They found that Chernoff faces encounter problems in cases in which only a few variables are available and that boxed letters function best overall. Nonetheless, they conclude that these results are isolated and that a robust theoretical basis for understanding and predicting people’s perceptions of multiple-cue graphics is essential.

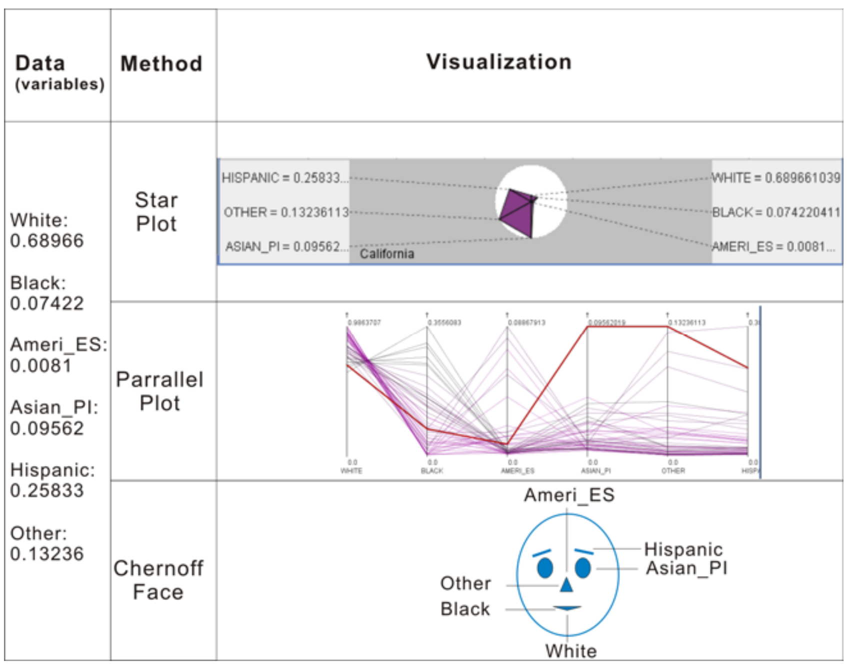


Figure 1. Alternatives to represent the same multivariate data set. Depicted are a star plot glyph, a parallel coordinate plot and a Chernoff Face, all showing the same data set (population characteristics for California).

In another comparative study on which graphic representation to use, Lee and his colleagues (Lee, Reilly, & Butavicius, 2003) investigated several representations of binary data (including star plot glyphs). A distinction that they made separated representation techniques that allow for presenting raw data compared to representation techniques that filter and preprocess the data before visualizing them. They classify Chernoff faces and star plot glyphs as raw data representations (Chambers, Cleveland, Kleiner, & Tukey, 1983). Preprocessing techniques on the other hand are, for example, spatializations (Skupin & Fabrikant, 2007) such as multidimensional scaling (MDS) or self-organizing maps (SOM). The latter type of representation reduces the dimensionality of a multivariate data set to an extent that the data can be represented in two or three dimensions. This reduction in dimensionality is meant to provide an easier access to the data than presenting raw data. In their experiments (using binary data), Lee and colleagues indeed found an advantage of spatialization techniques compared to raw data representations. Other studies (Borg & Staufenbiel, 1992), however, show that the type of preprocessing also has an influence on how data are interpreted. Hence, a new set of questions may be raised, urging answers to what kind of preprocessing should be applied by the designer.

Decisions on the micro level deal with the design of individual multivariate point symbols once the type of point symbol has been selected. For example, early research by Chernoff and Rizvi (1975) showed that the way variables are assigned to facial characteristics has an impact on how the data represented by Chernoff faces is interpreted in a classification task. This observation has triggered research, both theoretical and behavioral, on how to decide on the best way to assign variables to facial characteristics. Nelson (2007), for example, discusses the concept of feature salience for Chernoff faces. Feature salience is defined as "[…] the perceptual ordering of facial features from those that produce the most noticeable changes to those that produce the least noticeable changes." (p. 5). She discusses extensively the importance of this concept in the assignment of variables to facial characteristics, noting in particular that the transformation of a data space into a graphic (facial) space has the potential to affect the way the data is interpreted, for example in a classification task (Chernoff & Rizvi, 1975).

In this context it is important to note that many representations, not only Chernoff faces or star plot glyphs, have component parts. Research on the salience of objects and object parts has shown that not all object parts are equally salient. This holds for both the emotionally important parts of faces as well as the general salience of object parts (Hoffman & Singh, 1997).

A question for star plots is, for example, how many parts does a star plot have? Looking at a *classical, ideal* star plot representing 8 variables (see, for example, Figure 2), we could suggest that this star plot has four distinct parts. This assessment would correspond to the minima rule defined by Hoffman and Singh (1997) who state that "[…] human vision defines part boundaries at negative minima of curvature on silhouettes, and along negative minima of the principle curvatures on surfaces." (p. 29). Huffman and Singh (1997) propose that the salience of an object part is defined by three factors: (1) its size in relation to the whole object, (2) the degree to which it protrudes, (3) the strength of its boundaries. They found evidence that these factors indeed influence the visual processing of objects and object parts. In case of star plot glyphs, minima, the size of an object part, and the degree of protrusion are defined by low and high variable values of neighboring axes. For example, the star plot glyphs in Figure 2 that are star shape have four distinguishable parts.

Shape is extremely powerful in providing a first index of an object (Palmer, 1999; Biederman, 2007). A characteristic shape alone allows the human cognitive system to recognize, essentially instantly, an object even if it is only a silhouette. Hence, something similar should happen for star plots, too.

As a preview on our experiments we are depicting several star plot glyphs in Figure ##2##. This Figure visually accompanies the preceding discussion on decisions on the micro-level. While a designer (in the broadest sense) might have settled on star plot glyphs as the best visual representation for a data set at hand, this still leaves manifold design choices. The columns in Figure 2 show different visualizations for identical data sets, the rows show how a design choice manifests itself visually across different data sets. The design choices presented here are: first, how to assign variables to rays, and second, whether or not to additionally color the rays of a star plot glyph. Figure 2 should give an idea how representations of the same data set are affected by decisions on the micro-level.

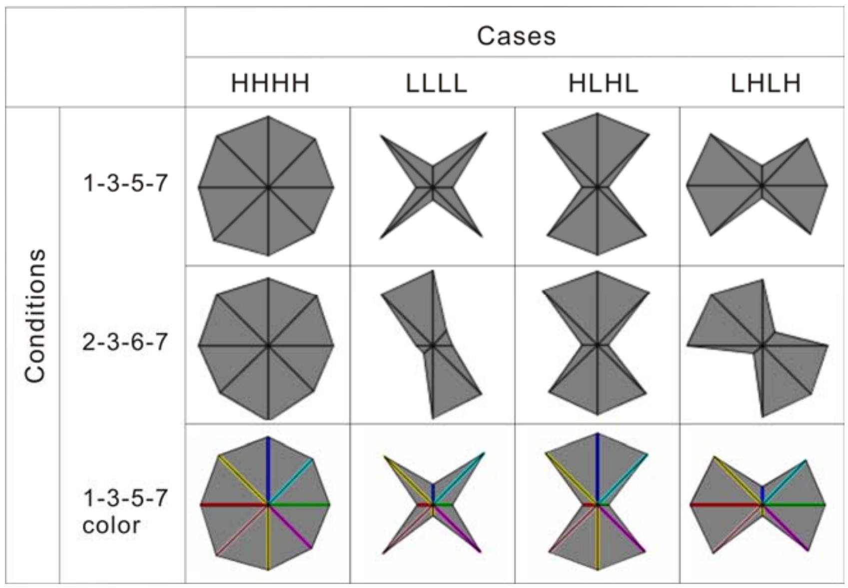


Figure 2. Examples of the star plot glyphs used in our experiments, demonstrating how different visual representations can look as the result of micro-level decisions such as how to assign variables to rays or whether or not color should be used to distinguish rays. The attribute values are qualitatively classified into variations along 4 rays (variables): HHHH – all values are high, LLLL – all values are low, and HLHL as well as LHLH are combinations of high and low values. More details can be found in the Experiment section.

In general, when people encounter sets of new or known entities, they are able to categorize them spontaneously into groups (Lakoff, 1987; Rosch, 1978; Margolis & Laurence, 1999). Given the continuous stream of information people encounter, categorization is an efficient way to reduce the amount of information a cognitive system has to deal with as it establishes equivalent classes (categories). Entities within an equivalent class (a category) are treated as similar to one another but different from entities in other equivalent classes. A distinction is made between the case where participants learn category structures, that is, placing entities into appropriate categories through feedback, and the case where participants are asked to spontaneously divide a given set of entities into groups without feedback. Given the omnipresence of nearly continuous categorization, feedback is regarded as an artificial influence (Milton & Wills, 2004). Several studies have excluded feedback by allowing participants to categorize provided entities in a way they think is most appropriate (Pothos & Chater, 2002; Milton & Wills, 2004) or most natural without providing any a priori information on the number or characteristics of the groupings. The key to this approach is that no feedback is provided about whether the participant-created groupings adhere to previously established category structures. Several names have been used for this approach, including unsupervised human categorization (Pothos & Chater, 2002), free sorting (Billman & Davies, 2005), category construction (Medin, Wattenmaker, & Hampson, 1987), and free classification (Handel & Imai, 1972).

One of the most intriguing questions is how perceptual and conceptual aspects interact with each other in cases where visual representations (such as star plot glyphs) are used for abstract, multivariate data spaces.

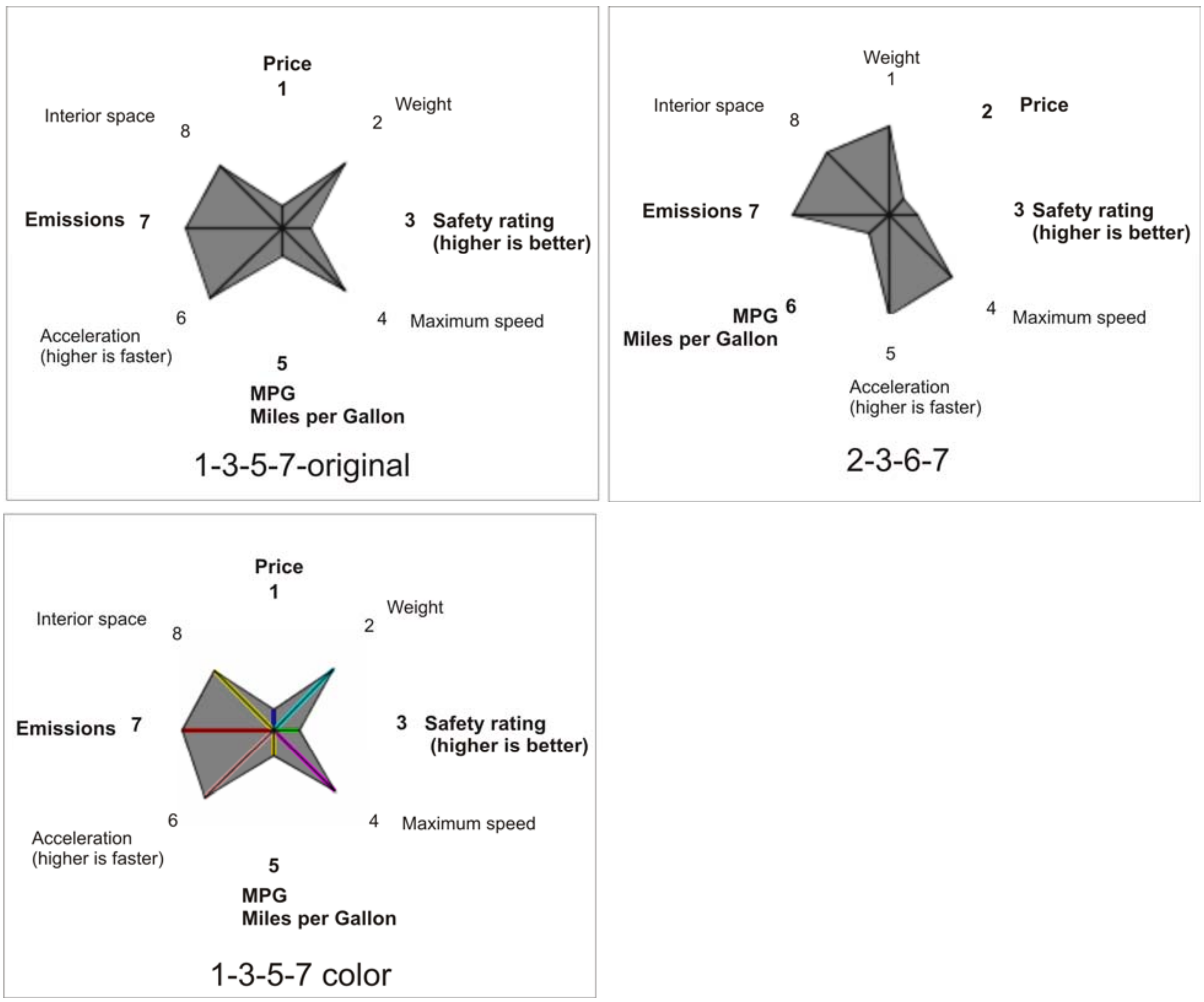


Figure 3. Assignment of variables to rays in a star plot glyph. Depicted are all three conditions that represent data using different star plot glyph characteristics: condition 1-3-5-7, condition 1-3-5-7 color, and condition 2-3-6-7. We varied 4 variables: price, safety rating, miles per gallon, and emissions (bold font). Differences in appearance occur due to the different assignment of variables to rays (1-3-5-7 versus 2-3-6-7) and the use of color (1-3-5-7 versus 1-3-5-7 color). The forth condition, 1-3-5-7 geom, is identical to 1-3-5-7 but without the variables, that is, only the shape is used.

Experimental setting

The experimental setting follows a previously conducted experiment (Klippel et al., to appear, 2009). To ensure that this article is self contained, we will present the details about the methodology and important findings from previous experiments here in as much detail as necessary.

In both experiments (experimental conditions) that are the primary focus of this article, as well as in the previous two experiments reported in (Klippel et al., to appear, 2009), we used eight ray star plots glyphs (compare Figure 2 and 3). In three experiments (of the four) the star plot glyphs represent data about future cars, in one experiment the star plot glyphs are simply introduced as geometric figures (details will be provided below). To be able to refer to the experiments (later on also referred to as conditions) unambiguously, we are using the following terminology that we will explain in detail in this section. Please note that the numbers refer to the rays as depicted in Figure 3 and indicate the rays that exhibit major variation in their values:

1-3-5-7-original: This refers to the experiment reported in (Klippel et al., to appear, 2009). The main characteristic of this experiment was that the variation of the data values occurred along the axis 1, 3, 5, and 7 (see Figure 3)

2-3-6-7: This refers to the experiment reported in (Klippel et al., to appear, 2009). The main characteristic of this experiment was that the variation of the data values occurred along the axis 2, 3, 6, and 7 (see Figure 3). Otherwise the experiment and the data used were identical to 1-3-5-7-original. This allowed us to compare differences which occur only due to the arrangement of the axes.

1-3-5-7-geom: In this experiment we used the same star plot glyphs that were used in experiment 1-3-5-7-original. The only difference was that the star plot glyphs were not introduced as being semantically meaningful, i.e. they were not representations of data sets but simply geometric shapes.

1-3-5-7-color: In this experiment we used again the same star plot glyphs that were used in experiment 1-3-5-7-original. The only difference was that the rays were enhanced with color. Each ray was colored using a categorical color scheme, using the primary colors in the red-green-blue and cyan-magenta-yellow color spaces, as well as two additional contrasting colors (pink and orange).

The three experiments in which the star plots represent data about future cars (1-3-5-7-original, 2-3-6-7-geom, 1-3-5-7-color) all have the same data set using eight variables to characterize cars (compare Chambers et al., 1983): price, weight, safety rating, maximum speed, miles per gallon, acceleration, emissions, and interior space. The data set is designed such that only four of these variables exhibit major changes in their values (price, safety rating, miles per gallon, and emissions). The other four variables are always high. To design the star plot glyphs, we used three distinct qualitative classes, i.e. low, medium, and high. Within these classes the values were randomized but it was ensured that the classes could be distinguished. Figure 4 details the design of the star plot glyphs. The differences between these three experiments lies either in the way the variables are assigned to the rays: In the two conditions 1-3-5-7- original and 2-3-6-7 two pairs of rays were exchanged (1 with 2 and 5 with 6). This simple change had a dramatic effect on the shape characteristics of the star plot glyphs (see, for example, Figure 2,3, and 6). In the two conditions 1-3-5-7-original and 1-3-5-7-color the only difference was that the rays in the star plot glyphs were color coded (see as an example Figure 6). The rationale for this procedure was that we found influences of shape characteristics that cannot be explained by the represented data (i.e. we found statistically significant differences in the classification of star plot glyphs in experiment 1-3-5-7-original and 2-3-6-7).

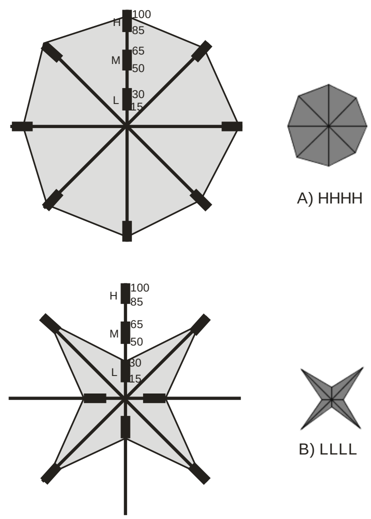


Figure 4. Design of the star plot glyphs. Four variables were varied on three different levels: high (100-85), medium (65-50), low (30-15). The remaining four variables that were not varied on three levels were randomized on the high level (i.e., between 100 and 85). Exemplarily the design of two star plot glyphs for the category (A) all-values-high (HHHH) and (B) all-values-low (LLLL) are depicted for condition 1-3-5-7. Other conditions work accordingly.

Hence, a logical follow up study to explore the influence of shape characteristics on classification tasks (and how they might be overcome) was to enhance the *semantic clarity* of the star plots by coloring their rays (see, for example, Figure 3 and 6). This procedure should allow for focusing the attention of the participants on the fact that the star plots represent data and to more easily associate a particular ray with a variable. To give a concrete example, the most prominent, unexpected classification that was found in experiment 1-3-5-7-original was “has one spike” (see also Figure 6)[2]. This grouping based on the shape characteristics of the star plot glyphs cannot be explained by the characteristics of the represented data. Consequently, the grouping of the represented data set of the star plot glyphs with one spike was different in condition 2-3-6-7 where two variables were assigned to different rays resulting in very different shape characteristics. Color, as one of the most effective tools in designing graphic representations (e.g., Brewer & Pickle, 2002) might be able to leverage the influence of shape by focusing attention on the differences between the star plot glyphs. To take up the example again: All spikes in the “has one spike” group of star plot glyphs have now one spike, too, but the spikes have different colors.

[2] The description “has one spike” does only apply to the three conditions 1-3-5-7. The description is colloquially used to describe this specific shape.

To round up the assessment of shape influences, the fourth experiment (1-3-5-7-geom) was designed to further deepen the understanding of shape influences as such. We again used the same data set and the same star plot glyphs that we used in experiment 1-3-5-7-original. This time though, instead of using the star plot glyphs as representations of a data set, the icons were introduced to participants as geometric shapes without any further meaning. This way we created a baseline for classification behavior that relies completely on shape characteristics.

All experiments used the same number of icons, 81. This number results from the variation we introduced: four variables were varied using three distinct classes (low, medium, and high). Using all possible combinations results in 81 icons. The four remaining variables were randomized within the class of high values.

Procedure

To realize the classification (category construction) experiments we are using a custom made software solution that allows for presenting and collecting classification data. After a warm up task to acquaint participants with the software tool as well as the idea of sorting icons based on their similarity, all icons are shown to the participants at the same time (requiring scrolling) on the left side of the screen. As explained in the introductory section, we are using a free classification (category construction) approach that does not provide the participants with feedback regarding their grouping choices. Additionally, the participants are free to create as many groups as they consider appropriate. The experiments were designed as group experiments, that is, up to 11 participants could perform the experiment at the same time. In reality this number was between 3 and 9. The experiments took place in a GIScience laboratory at the Geography Department at the Pennsylvania State University using Dell Pentium D, 3 GHz computers with 2 GB RAM and wide screen 20-inch monitors. The lab was prepared such that it was not possible for participants to see each other's screens. After the grouping task participants were again presented with the groups that they created and had to (a) provide a linguistic label for the groups and (b) provide an explanation for their grouping rational.

Participants. A total of 40 participants took place in the new experiments, 20 in each condition. The participants were undergraduate students in geography primarily recruited from introductory classes. In condition 1-3-5-7-geom 7 out of 20 participants were female, the average age was 20.7. In condition 1-3-5-7-color 9 out of 20 participants were female, the average age was 21.2. The participant characteristics are comparable to the ones in the previous experiments (1-3-5-7-original and 2-3-6-7) that we will use as a comparison.

Results and analyses

The grouping behavior of the participants is captured in similarity matrices for each participant. Each similarity matrix contains 81x81=6561 cells that encode for each icon whether it has been placed together with the other icons or not (0 = two icons are in different groups, 1 = two icons are in the same group). A summary matrix is obtained by adding up all individual matrices. Thereby a similarity measure is created that can be subjected to various analyses methods such as cluster analysis or multi-dimensional scaling. We performed cluster analysis on all four experiments to gain first insight into the clustering structure as a result of the classification task. We used CLUSTAN™, squared Euclidian distance as a similarity measure, and various clustering algorithms such as Ward's method and average linkage. We used cluster analysis primarily as a way to gain insights into the clustering structure. We are not interested per se in the number of clusters and will not report the details of the cluster analysis here.

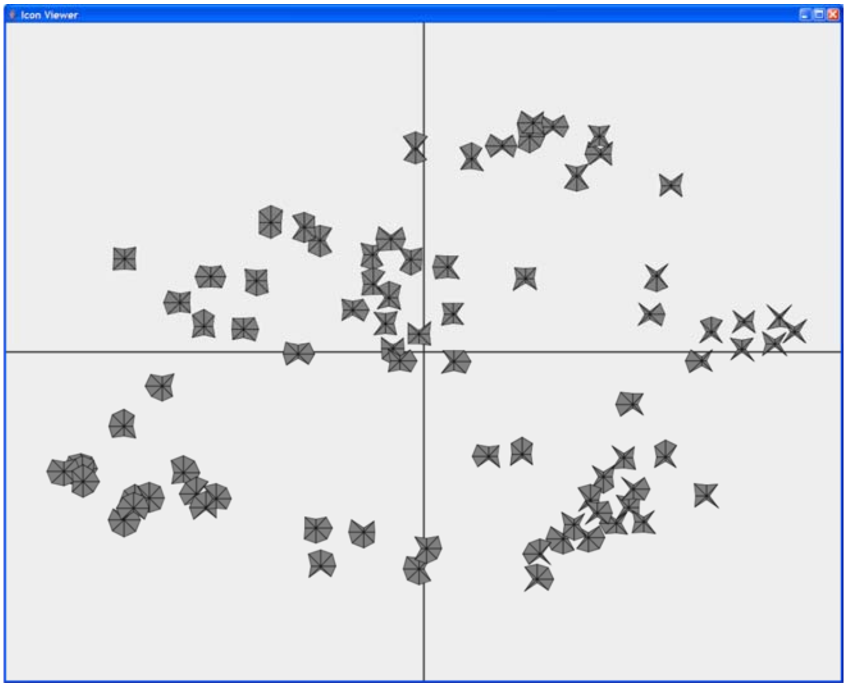


Figure 5. Exemplary depiction of the MDS plot for condition 1-3-5-7-geom

As an alternative way to gain insights into the clustering structure, we performed multidimensional scaling. The results (exemplarily for condition 1-3-5-7-geom) of this process (using again CLUSTAN and a custom made software tool to visualize the MDS plots showing the star plot glyphs) are depicted in Figure 4. These MDS plots provide a first visual indication that the similarity ratings are partially induced by salient shapes and not only by the data characteristics alone. We quantified these differences in (Klippel et al., to appear, 2009) for experiments 1-3-5-7-original and 2-3-6-7 and showed statistically significant differences for salient shape groups. The two new experiments will allow further exploration of the influence of shape and how shape influences may be reduced. For most of the analysis we not only focus on the two newly conducted experiments but compared all four conditions with each other to provide a deeper understanding of the influences on the classification behavior of the participants.

Time to finish the task. While there was no time constraint explicitly introduced, that is, participants were not told to finish the task as fast as possible, we previously found a statistical trend in the comparison of experiments 1-3-5-7-original and 2-3-6-7, indicating that salient shapes have a positive influence on the time required to finish the task. With now four experiments that are based on the same data set we are able to compare them together using a one-way analysis of variance (ANOVA). We found a highly statistically significant effect of presentation condition between all four experiments ($F(3,76) = 6.907$, $p<.0005$). A planned contrast confirmed the statistical trend for the comparison of condition 1-3-5-7-original and 2-3-6-7 ($t = -1.622$, $df = 76$, $p = .055$, one tailed). Additionally, a planned contrast between the condition 1-3-5-7-geom (geometric shapes without meaning) and all other conditions was significant (t = 3.899, df = 76, p < .0005). Likewise, a planned comparison of the identical shape experiments (contrasting shapes without meaning, 1-3-5-7-geom, and shapes with meaning, 1-3-5-7-original and 1-3-5-7-color) showed that semantics significantly slowed down the classification process (t = 2.924, df = 76, p = .005). The last planned contrast compared the non-salient shape condition 2-3-6-7 with the color enhanced salient shape experiment. For this comparison we found a statistically significant time advantage for the color coded star plot glyphs (t = 2.283, df = 76, p = .0125, one-tailed).

Number of groups. We repeated the analysis for the number of groups that participants created in all four experiments. None of the analyses showed up being statistically significant.

Correlation analysis of the similarity matrices. We performed a correlation analysis on the similarity values of all pair-wise star plot glyphs. Each similarity matrix contains 3240 meaningful cells encoding the similarity from 0 (two star plot glyphs were never placed into the same group) to 20 (two star plot glyphs were always placed into the same group). The results of this analysis are depicted in Table 1. The Table shows for all possible combination of the 4 experiments that the correlation coefficients are statistically significant. This, to some extent, may be a byproduct of the high number of cells (N=3240). The more interesting observation from this table is that both new experiments show higher correlation values with the experiment 1-3-5-7-original than with condition 2-3-6-7, with the correlation between 2-3-6-7 and 1-3-5-7-geom being the lowest (.474). In other words, the conditions with identical shapes correlate higher than the ones with identical meanings (but different shapes).

Table 1. Similarity coefficients for all four conditions (1-3-5-7-original, 2-3-6-7, 1-3-5-7-color, and 1-3-5-7-color)

			1357-original	2367	1357-color	1357-geom
Spearman's rho	1-3-5-7-original	Correlation Coefficient	1.000	.576(**)	.684(**)	.691(**)
		Sig. (2-tailed)	.	.000	.000	.000
		N	3240	3240	3240	3240
	2367	Correlation Coefficient	.576(**)	1.000	.515(**)	.474(**)
		Sig. (2-tailed)	.000	.	.000	.000
		N	3240	3240	3240	3240
	1357-color	Correlation Coefficient	.684(**)	.515(**)	1.000	.583(**)
		Sig. (2-tailed)	.000	.000	.	.000
		N	3240	3240	3240	3240
	1357-geom	Correlation Coefficient	.691(**)	.474(**)	.583(**)	1.000
		Sig. (2-tailed)	.000	.000	.000	.
		N	3240	3240	3240	3240

** Correlation is significant at the 0.01 level (2-tailed).

In-depth analysis of the classification task. Although statistical methods have proven widely applicable for identification and analysis of patterns in data taken as a whole, they are much less useful when it is desirable to explore both coarse- and fine-grained structure at different scales or across different subsets of multiple dimensions. In contrast, the relatively new field of visual analytics (Thomas & Cook, 2005) is starting to produce techniques that support exactly this sort of flexible exploration inside highly interactive visual user interfaces. These two broad families of analytic approaches nevertheless promise to be complementary. Visual tools provide analysts with a means to "drill-down" into heterogeneous data in order to identify and isolate data subsets of suitable form for input into particular statistical methods. The results of such methods can not only stand by themselves, but may also provide further intuition that spurs continued visual analysis in a virtuous cycle. Often, visualization techniques alone prove sufficient for exploration and analysis down to very fine levels of detail, relieving the need for any assistance from statistical methods at all.

There are an increasing number of interactive builders and programmer toolkits for implementing visualization software designed for use with data from particular knowledge domains. We used the Improvise visualization environment (Weaver, 2004) to develop a visual tool, KlipArt, designed for the purpose of identifying and analyzing individual differences in icon classification. The flexibility and speed of interactive visualization design in Improvise has enabled us to modify and extend KlipArt in response to our own evolving models of icon classification behavior and our corresponding data analysis needs and desires. The tool is currently in its third major version. One variant of the current visualization (see Figure 5), designed for use with typical modest screens sizes, displays data from only a single experiment at a time. An alternate variant takes advantage of larger screens by displaying data from two experiments side-by-side for comparison.

The focal point of the tool (see Figure 6) is a graph that shows associations between participants and icon groupings. The graph thus contains two classes of nodes: one node for each participant (yellow squares), plus one node for each unique icon grouping (white rectangles). Arrows connect a participant node with a grouping node whenever that participant classified icons into exactly that grouping. As an additional visual guide, 'packs' enclose the grouping nodes of each participant using a visual shrink-wrapping metaphor.

The exploratory utility of the graph arises not from static representation of all participant-grouping relationships, but rather the ability to drill-down into the data by selecting arbitrary subsets of icons, participants, and icon groups. The graph shows only selected items from each set. Moreover, icon groups are filtered so as to display only selected icons. As a result, it is possible to visually 'slice' the data by posing subset queries about specific icon characteristics shared (or not) by the selected set of icons, and thereby explore hypotheses about how participants classify icons as a function of such characteristics. Importantly, arbitrary icon set selection means that the unifying icon "characteristics" are defined by the experimenter, possibly but not necessarily in response to the characteristics that different sets of participants appear to use to classify icons (or by using MDS to identify interesting groupings, see Figure 5).

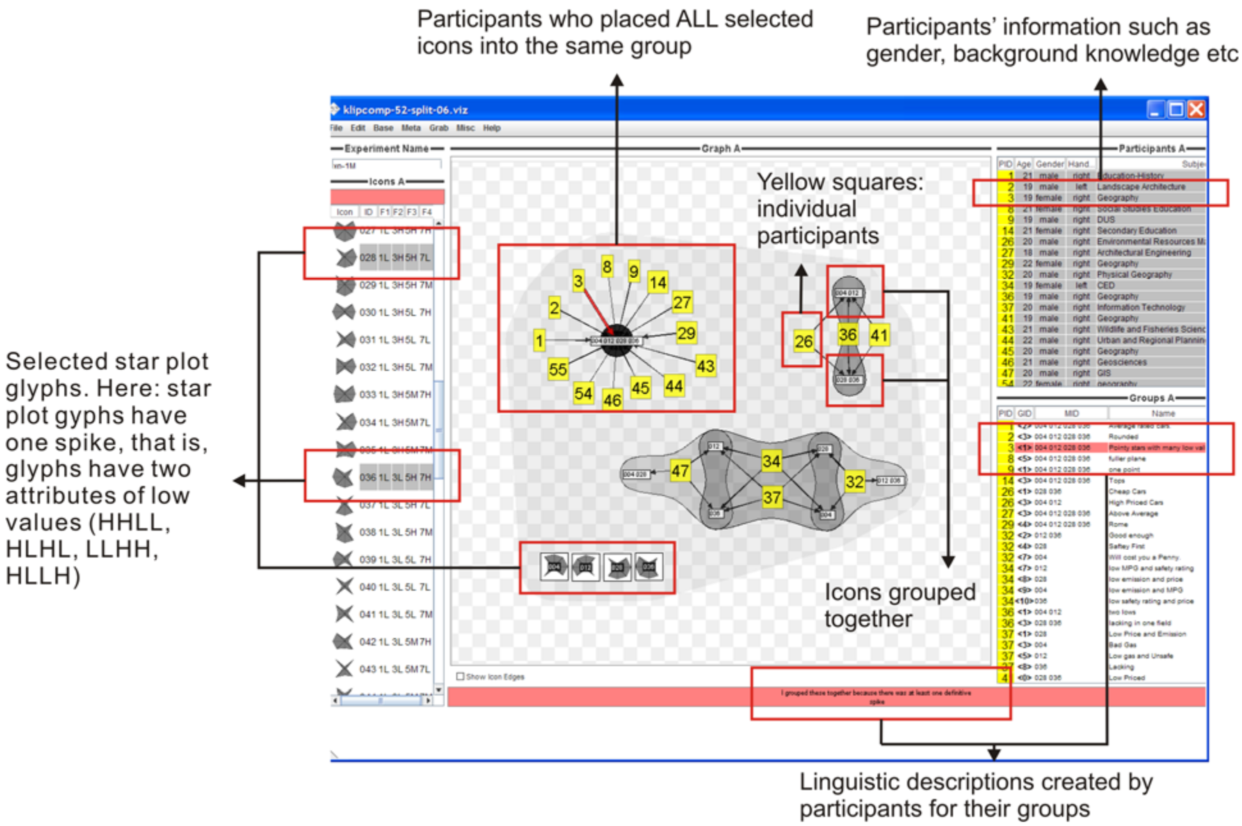


Figure 6. A snapshot of the KlipArt tool. It shows (on the left side, highlighted in gray) the star plot glyphs that are selected. The central part shows the grouping (classification) behavior of the participants for the selected glyphs. The right side provides information on the participants (upper part) and the linguistic descriptions that participants provided (lower part).

Finally, to facilitate rapid exploration of the complex graphs that result when many items are selected, the user can "brush" participants, icons, or groups by mousing over the tabular rows of icons, participants, or groups, respectively. Brushing causes the corresponding nodes in the graph to highlight one at a time, allowing the user to locate items quickly. (Brushing in the reverse direction, from nodes to rows, is planned for the next version of KlipArt.) A participant's linguistic description of an icon group is additionally displayed in full at the bottom of the visualization when the user brushes the corresponding group row. Searching and filtering icon groups on keywords is a potentially useful future feature of the tool.

Our first comparison used the group of star plot glyphs identified as "has one spike". The most prototypical examples of this group are those with two high and two low values. The corresponding star plot glyphs are depicted in Figure 7. The numbers of participants who placed the four icons into the same group are 14 for 1-3-5-7-original, 2 for 2-3-6-7, 8 for 1-3-5-7-color, and 12 for 1-3-5-7-geom. The difference between 1-3-5-7-original and 2-3-6-7 is statistically significant (χ^2 = 12.601, 1, p < .0005) and is an example of how salient shape characteristics can influence classification tasks. Our subsequent analysis of all four experiments showed that the differences are still statistically significant over all experiments, too (χ^2 = 16.97, 3, p = .001). Addressing now the previously posed question of whether color reduces the effect of salient shapes, we compared the two conditions 2-3-6-7 and 1-3-5-7-color. The analysis showed (applying continuity correction as advised for 2x2 tables) a reduced significance below the .05 threshold (χ^2 = 3.333, 1, p = .068).

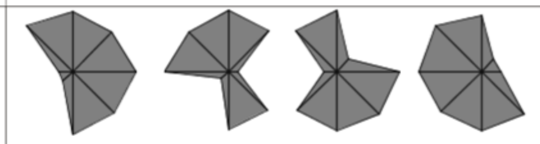
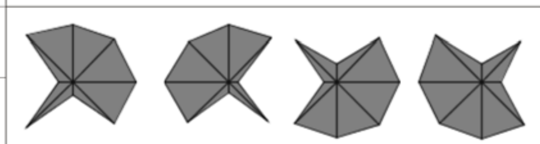
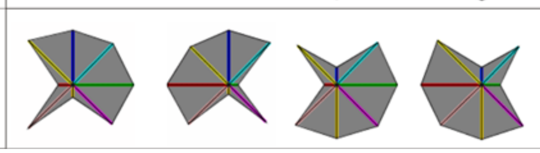
Conditions	Icons	Number of Participants
2-3-6-7		2
1-3-5-7 original		14
1-3-5-7 geom		12
1-3-5-7 color		8

Figure 7. Grouping behavior for star plot glyphs. Depicted is the analysis of the grouping behavior for a subset of icons (see also Figure 5 and 6) that are presenting four identical data sets across the four conditions. In the case here, the data is characterized by a variation of two variables (out of the four we varied) that show low values and two that show high values. The result is a shape we refer to as “has one spike”. Please note that the icons for 1-3-5-7 original and 1-3-5-7 geom are identical and therefore presented only once. The number of participants indicates how many participants classified all four star plot glyphs as belonging into the same category.

We compared the grouping behavior for various shapes families (using the KlipArt tool). Another example that repeated this pattern is shown in Figure 8. In this example, 11 participants in condition 1-3-5-7-original, 3 in 2-3-6-7, 1 in 1-3-5-7-color, and 7 in 1-3-5-7-geom placed all five star plot glyphs into the same group. Again the overall comparison of icons placed into the same group showed a statistically significant difference (χ^2 = 14.796, 3, p = .002). More interesting though is that the color enhanced version of experiment 1-3-5-7-original is more in line with experiment 2-3-6-7 than with experiment 1-3-5-7-original.

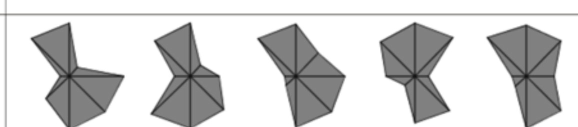
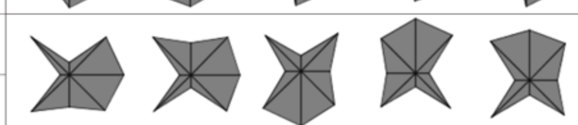
Conditions	Icons	Number of Participants
2-3-6-7		3
1-3-5-7 original		11
1-3-5-7 geom		7
1-3-5-7 color		1

Figure 8. Color is reducing shape influences.

An interesting question is the one of semantics versus shape characteristics. Informally, a semantic difference would be one caused by comparing the glyphs based on their meaning (that is, the data they represent), while shape characteristics are features like "has one spike" irrespective of the data meaning. To look into this relationship, we paired five star plot glyphs together: the ones representing the lowest possible values for all 4 variables and the semantically closest ones, i.e. those having three variables with lowest values and one variable with a medium value (see Figure 9). The results are that 13 participants in condition 1-3-5-7-original, 11 in 2-3-6-7, 13 in 1-3-5-7-color, but only 6 in 1-3-5-7-geom classified the star plot glyphs as belonging in the same category. The statistical analysis shows a statistical trend (χ^2 = 6.587, 3, p = .086, two-sided).

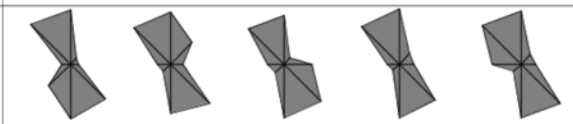
Conditions	Icons	Number of Participants
2-3-6-7		11
1-3-5-7 original		13
1-3-5-7 geom		6
1-3-5-7 color		13

Figure 9. Classification of semantically similar values.

An interesting observation is that the saliency of shapes works on different levels. Sometimes similar shapes are in line with the represented data similarities, sometimes they work in the opposite direction. In other words, in some such cases the data similarity does not match the shape similarity. This becomes apparent in cases where conditions 1-3-5-7-original, 1-3-5-7-color, and 1-3-5-7-geom are creating shape differences but condition 2-3-6-7 does not. One example is depicted in Figure 10, showing a group of 5 star plot glyphs (icons are placed together 5 times in condition 1-3-5-7-original, 10 in 2-3-6-7, 2 in 1-3-5-7-color, and 1 in 1-3-5-7-geom). There are only comparatively small differences in the shape characteristics in condition 2-3-6-7 but there are large differences in the other three conditions (χ^2 = 14.143, 3, p = .003, two-sided). Another example is presented in Figure 11 (χ^2 = 16.207, 3, p = .001, two-sided) showing again how salient shape characteristics of some icons lead to differences in the grouping behavior between condition 2-3-6-7 (12), and the other three conditions (1-3-5-7-original (3), 1-3-5-7-color (4), and 1-3-5-7-geom (2)).

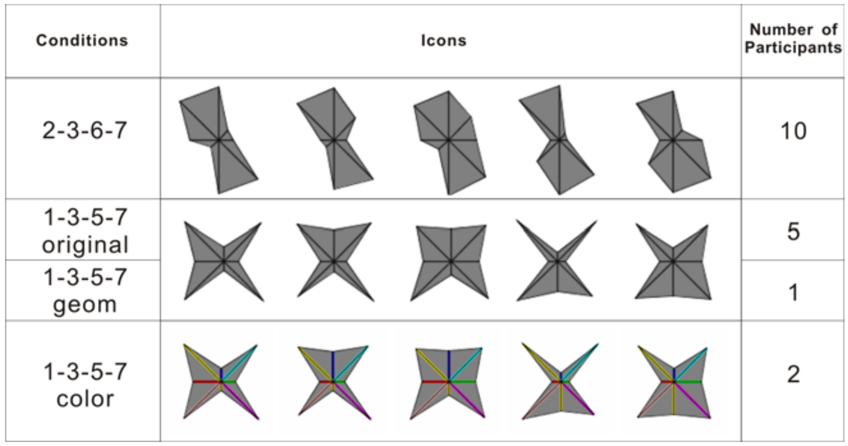


Figure 10. Effects of semantics and non-salient shapes.

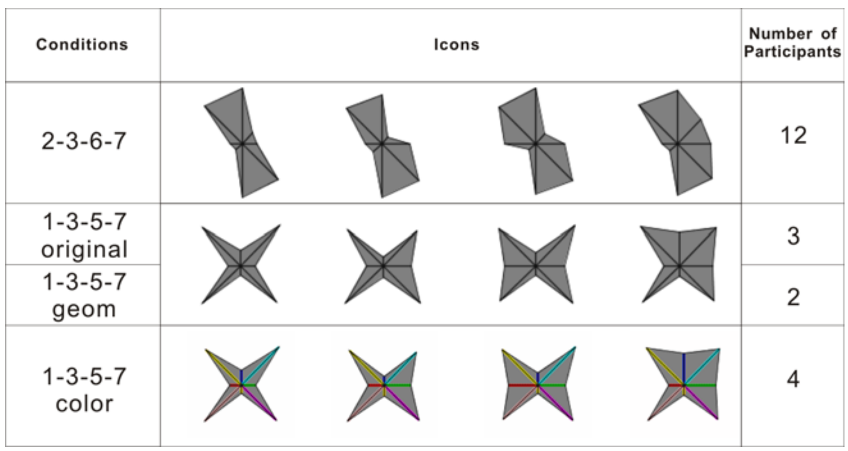


Figure 11. Non-salient shapes closer to semantic similarity.

One last, somewhat unexpected finding should be reported, too. It has been found, likely as an exception, that in cases where all four conditions show similar shape characteristics, the color coding of the rays can lead to a classification behavior distinct from the other three conditions (see Figure 12, χ^2 = 8.095, 3, p = .044, two-sided).

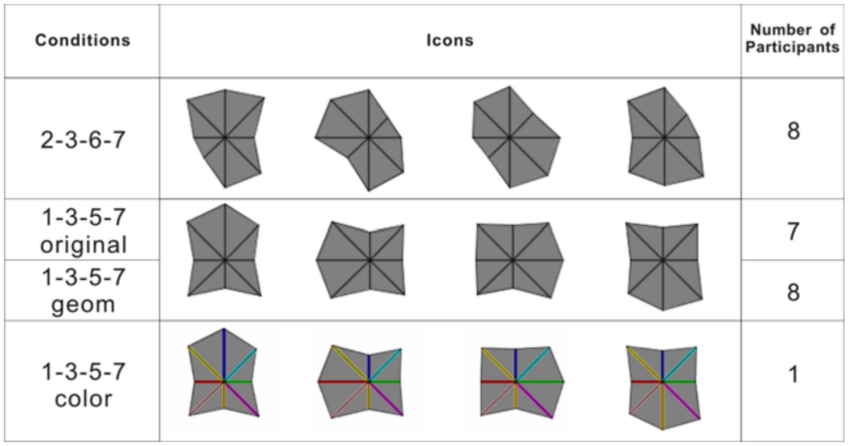


Figure 12.Effects of colors on similar shapes.

Discussion

The results of these experiments reveal several aspects of salient shape characteristics and color in designing graphic representations of multivariate data spaces. One of the most obvious findings is that adding meaning (i.e. semantics) to a representation slows down decision making. This aspect is clearly demonstrated by the results of the ANOVA of the time it took participants to finish the task. Condition 1-3-5-7-geom, that is, the condition where the star plot glyphs are stripped off their meaning and participants classify them on the basis of their shape characteristics, was significantly faster compared to all other conditions.

Beyond this observation the analysis of the time it took participants to classify all items reveal one further important aspect. In the original comparison of condition 1-3-5-7-original and 2-3-6-7 we found a statistical trend indicating an advantage for the condition that shows more salient shapes, that is, condition 1-3-5-7-original with the classic star-shaped star plot glyphs. This trend was reinforced by adding color to the star plot glyphs; comparing the times of condition 2-3-6-7 and 1-3-5-7-color shows a statistically significant advantage for the color enhanced star plot glyphs. While our experiments were not only focusing on aspects of visual processing, these results can be explained by looking into the literature on visual perception and particularly pre-attentive processing. Research on pre-attentive processing has revealed that a limited number of visual characteristics have the potential to be detected very rapidly by the visual system in early stages of processing (low-level vision). It is believed that these processes are automatic and, importantly, spatially parallel (Healey, Booth, & Enns, 1996; Ware, 2004). Normally, work addressing questions of pre-attentive processing employs reaction times and accuracy measures. Nonetheless, two of the features employed in our research are also frequently mentioned as features that guide pre-attentive processing: color and shape (the latter, as discussed in the introduction, has to be looked at from the perspective of what makes a shape salient, for example, convexity; see Hoffman & Singh, 1997). The process of pre-attentive processing can be thought of as a guidance mechanism in a signal detection problem (Wolfe, 2005). Targets function as a signal while other icons create noise. This situation occurs in our experiments once a participant has started to create categories and is looking for icons (star plot glyphs) to be placed into the category. This situation also occurs in the very beginning when a participant is scanning the set of star plot glyphs to decide on which star plot glyph to choose as a seed for a new category. Obviously, if icons can be distinguished more easily from one another or if the visual system is able to structure the search process by employing features pre-attentively processed, these icons will be placed into groups faster than those that have to be searched for (i.e., star plot glyphs that are less salient).

While we find that a combination of more salient shape characteristics and color enhances information processing by combining features that can be pre-attentively processed, the original question addressed the influence that salient shape characteristics have on the classification behavior of participants. This question is one that is omnipresent in research on graphic representations, that is, the question of how the choices a designer makes (consciously or not) impact on the way visual information is interpreted. To answer this question we used a custom made software solution (KlipArt) to identify groupings of salient shape characteristics and analyzed how the grouping behavior changes across the four different conditions. As the most prominent example, we looked at a shape characteristic that could be labeled "has one spike". For this characteristic we found a statistically significant difference between condition 1-3-5-7-original and 2-3-6-7 indicating that, just like for Chernoff faces (Chernoff & Rizvi, 1975), the assignment of variables to rays in a star plot glyph has an impact on the classification behavior of participants. In comparison to this result, the color enhanced star plot glyphs reduced this effect and pushed it below the classically assumed .05 significance threshold. We found other examples in our analysis where the raw counts showed the same tendency but the overall results did not surface as being statistically significant.

These results can be taken as an indication that color does not only enhance the speed of classification but additionally stresses the semantic focus of the data making it more 'difficult' for participants to classify icons on the basis of salient shape characteristics alone. However, the result of comparing the similarity values of the proximity matrices directly between all four conditions draws a slightly different picture in that it identifies the three conditions with identical shapes (1-3-5-7-original, 1-3-5-7-color, and 1-3-5-7-geom) as being more highly correlated among each other than in comparison to condition 2-3-6-7. However, all correlations are statistically highly significant and therefore the overall similarity might not be the best indicator for assessing the classification behavior of the participants on the level of fine-granularity classifications.

The latter findings are, however, reinforced by comparing situations in which shape characteristics lead to a classification behavior that places semantically similar star plot glyphs into different categories because a certain star plot glyph might exhibit characteristics not shared by other star plot glyphs. In condition 2-3-6-7 where the shape characteristics are not as salient, the grouping behavior is closer to the semantic similarity of the data represented by the star plot glyphs (see Figures 9). It also has to be noted that color, just like other visually salient features, seems to be able to induce a certain classification behavior.

Conclusions

The comparison of the four experimental conditions that we applied reveals several important aspects for theoretical consideration when designing graphic representations of multivariate data spaces as well as for the design of experiments in cartographic sciences. One of the most striking effects of using color in star plot glyphs is the gain in time. Even without requiring participants to finish the task as quickly as possible, comparing the color enhanced star plots (experiment 1-3-5-7-color) with the regular non-salient star plots made the statistical trend (we found comparing 1-3-5-7-original and 2-3-6-7) statistically significant. Hence, one promising line of future research should be to evaluate the possibilities to combine features of point symbols that can be pre-attentively processed to facilitate search processes and decision making.

The use of color also reduced the effect of shape influences on the classification task, in some cases statistically significantly, in others at least in raw numbers. On the other hand, color seems to introduce its own rules and can, in cases where the overall shape across all four conditions is very similar, lead to different classification behaviors. This finding definitely deserves more attention and a more detailed analysis.

An interesting question to be addressed is that of visual clutter: How it can be defined? One way of defining (and then reducing) clutter in visual representations is to order dimensions (variables) according to their similarity. This ordering is regarded as being crucial for many multivariate visualization techniques such as parallel coordinate plots, scatter plots, recursive patterns (Ankerst, Berchtold, & Keim, 1998), and also for star plot glyphs (Peng, Ward, & Rundensteiner, 2004). While there definitely is some truth in this line of thought, it is also worthwhile to rethink this assumption of having to order dimensions according to their similarity in the light of findings from vision research on salient objects' parts. To make data distinguishable it might be best (at least for star plot glyphs) to arrange data such that the resulting shapes have as many distinguishable parts as possible to allow for easier identification in search processes and decision making. We need to define features and *feature salience* (Nelson, 2007) for star plot glyphs.

What our research shows is that there are various factors influencing classification of visually represented data. Shape is certainly one of the most prominent ones, but color has also shown to have an influence in certain situations: either positive, in that color further reduced the time the classification task required and also reduced the effect of salient shapes on classification behavior, or negative, in situations where shape characteristics are absent, as it seems to bias classification behavior, too.

It is therefore important to continue this line of research to bring out the best of visual features but without the negative influences, one could say manipulations (Monmonier, 1999), that may be unwillingly introduced by choosing (or letting the software chose) certain design parameters. A research framework often called for (Nelson & Gilmartin, 1996; Ward, 2008) is necessary to systematically analyze cartographic and other design choices to provide the creators of visual artifacts, whether human or machine, with guidelines that at the least make them aware of the consequences of their design choices.

An important contribution to this field comes from research employing eye movement studies (Fabrikant, Rebich-Hespanha, & Hegarty, in press). In experiments on interpreting weather maps Fabrikant and collaborators addressed the relationship between thematic relevance and perceptual salience. Their empirical findings document the cognitive adequacy of cartographic design principles that align thematic relevance with perceptual salience. It is important to (a) explicate these design principles and (b) behaviorally validate them to establish their cognitive adequacy.

One approach in this direction, in the case of star plot glyphs, is to formally characterize the shape characteristics of star plot glyphs and to compare them with the classification behavior of participants. This can be accomplished by either following work done in the area of qualitative spatial reasoning which seeks to identify shape primitives (Galton, 2000) or by more quantitative approaches (Latecki & Lahkämper, 1999). Galton (2000), for example, discusses shape families that he identifies on the basis of different shape primitives and that can be associated with linguistic labels. Latecki and colleagues (1999), on the other hand, extend the notion of convexity by using discrete curve evolution to identify salient object parts on coarser granularities. These two approaches could be used to define what makes certain shape characteristics more salient than others.

Additionally, it would be interesting to compare the influence of shape characteristics across different semantic domains to establish whether the choices of variables in our experiment had an effect, too. For example, an experiment could use variables that characterize a country (population, area, etc.) or provide information of the quality of life in a place (number of crimes, days of sunshine, etc.). Would we observe the same effects? Our hypothesized answer to this question is: Yes. We would still find advantages in processing the information presented if shape characteristics are more salient. We also would expect that shape would influence the classification behavior of participants and that this influence can be leveraged by using colored rays. The general goal should therefore be to find an alignment between perceptual salience and semantic importance to maximize information processing.

Acknowledgement

We would like to thank the reviewers for their valuable comments and feedback on the article. We are also indebted to participants of the Conference on Research and Training in Spatial Intelligence 2008 for their inspirational comments on this work.

This research is supported by the National Visualization and Analytics Center, a U.S. Department of Homeland Security program operated by the Pacific Northwest National Laboratory (PNNL). PNNL is a U.S. Department of Energy Office of Science laboratory. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agency.

References

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.

Bertin, J. (1983). *Semiology of graphics: Diagrams, networks, maps.* Madison, Wis.: Univ. of Wisconsin Press.

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.

Billman, D., & Davies, J. (2005). Consistent contrast and correlation in free sorting. *American Journal of Psychology*, *118*(3), 353–383.

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.

Chernoff, H. (1973). The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association, 68*, 361–368.

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.

Fabrikant, S. I., Rebich-Hespanha, S., & Hegarty, M. (in press). Cognitively adequate and perceptually salient graphic displays for efficient spatial inference making. *Annals of the Association of American Geographers*,

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

Handel, S., & Imai, S. (1972). The free classification of analyzable und unanalyzable stimuli. *Perception and Psychophysics, 12*, 108–116.

Healey, C. B., Booth, K. S., & Enns, J. T. (1996). High-speed visual estimation using preattentive processing. *ACM Transactions on Computer-Human Interaction, 3*(2), 107–135.

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

Klippel, A., Hardisty, F., & Weaver, C. (to appear, 2009). Star plots: How shape characteristics influence classification tasks. *Cartography and Geographic Information Science,*

Lakoff, G. (1987). *Women, fire and dangerous things.* Chicago: Chicago University Press.

Latecki, L. J., & Lahkämper, R. (1999). Convexity rule for shape decomposition based on discrete contour evolution. *Computer Vision and Image Understanding, 73*(3), 441–454.

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).

Margolis, E., & Laurence, S. (Eds.) (1999). *Concepts: Core readings.* Cambridge, MA: MIT Press.

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology, 19*(2), 242–279.

Milton, F., & Wills, A. J. (2004). The influence of stimulus properties on category construction. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*(2), 407.

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

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

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.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology. A Bradford book.* Cambridge, MA: MIT Press.

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.

Pothos, E. M., & Chater, N. (2002). A simplicity principle in unsupervised human categorization. *Cognitive Science, 26*(3), 303–343.

Rosch, E. L. B. (Ed.) (1978). *Cognition and categorization.* Hillsdale, NJ: Lawrence Erlbaum.

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.

Thomas, J. J., & Cook, K. A. (Eds.) (2005). *Illuminating the path: Research and Development Agenda for Visual Analytics.* Richland, WA: IEEE Press.

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* .

Wolfe, J. M. (2005). Guidance of visual search by preattentive information. In *Neurobiology of attention* (pp. 101–104). Amsterdam, Boston: Elsevier Academic Press.