Detecting Data Visualization Preferences Using Games

Lori L. Scarlatos, Kwame Nti, Alice Wong
Department of Technology and Society
Stony Brook University
Stony Brook, NY
Lori.Scarlatos@stonybrook.edu

Abstract— **In visualizations of large multivariate data sets, discrete data can be effectively represented using glyphs. Glyphs have the advantage of allowing for rapid visual comparison, using differing visual dimensions to represent the different variables in the data. Some types of glyphs accommodate even more variables by using shape to represent the data. Yet the characteristics of these shapes may have underlying perceptual meanings. The purpose of this study was to determine whether certain shape characteristics are commonly viewed as good or bad. We conducted a study using two methods to gather data: a traditional survey, and a casual game. The results of this study strongly suggest that there are certain shape characteristics that are generally perceived as positive/negative, although they are not necessarily what might be expected.**

Keywords-visualization; glyph; shape; perception

I. Introduction

"Big data" is a big problem, as indicated by President Obama's speeches, and subsequent requests for proposals from federal agencies such as DARPA and NSF [1]. This is because the quantity and variety of data collected by today's technologies are too massive and diverse for people to analyze using traditional techniques. Visual analytics has the potential to aid big data analysis by revealing patterns that identify trends, relationships, and points of interest, which may then be investigated more closely. Yet to do this effectively, care must be used in selecting visual representations for the various data dimensions.

One strategy for representing discrete variables is to use glyphs. These iconic representations of the data, which use differing visual dimensions to represent the different variables, are particularly useful for layering of information and showing small multiples. Glyphs can be made to contain even more information by using data plots to determine their shapes; some examples are whisker, star, and parallel coordinate plots. Numerous papers have examined the effective use of glyph representations for various applications [2, 3, 4]. Yet these papers do not consider the unconscious associations that visual attributes – such as shape – may hold. For example, if the data visualization produces a shape with negative connotations, the analyst may be predisposed to view that data negatively. Understanding these connotations will enable us to eliminate unintended bias in the visualizations. We therefore pose the question: are there shape characteristics that can be used to represent the goodness of the data values?

Some related, albeit limited, research has already been conducted on people's perceptions of shapes in the area of psychology. Regarding humans' preferences for curved versus sharp shapes, some studies show that people prefer the curved shapes [5, 6]. However, other studies have suggested that this preference may be moderated by valence [7] or even fashion [8]. In any case, it is not clear that this preference, whatever the cause, will translate to biases about the data in a visualization. Other studies have shown that people have a preference for, and more readily recognize, symmetrical shapes. Although most of these studies focus on the preference for symmetry in human faces [9], it has also been shown that this is also true with more abstract images [10]. Yet this does not necessarily mean that data analysts will therefore have a negative impression of data represented with asymmetric shapes.

The purpose of our study was to determine whether there are particular shape characteristics that are commonly perceived as beneficent (good) versus maleficent (bad). In addition to looking at rounded versus sharp corners, we looked at other shape characteristics: convexity, symmetry, orientation, and sliveriness. The remainder of this paper describes the characteristics we looked at, our methodology, and our results. The contribution of this paper is that it suggests that there are certain shape characteristics with strong positive/negative connotations, although they are not necessarily the ones we would expect. It also suggests that gamified surveys can increase completion rates while delivering results consistent with those of the traditional survey.

II. Shape Characteristics

Five shape characteristics were examined in this study: rounded versus sharp corners; convex versus concave shapes; symmetry versus asymmetry; slivery versus chunky shapes; and shapes with varying orientations, pointing upward, downward, left and right. The purpose of this study was to determine, within those categories, whether specific attributes have positive or negative connotations. For example, if a multivariate data point is represented as a shape with corners that are either sharp points or rounded spline curves (with tension of the curve determined by another variable), will the roundness of the corners imply that the data values being

978-1-4799-2546-9/13/$31.00 ©2013 IEEE

represented are either good or bad? In all cases we used two-dimensional shapes where only the shape characteristic in question was varied. All shapes shown are neutral, i.e. not representing objects with strong intrinsic attractiveness or aversiveness.

Because of previous studies suggesting that people preferred rounded shapes to those with sharp corners [5, 6], we expected that this might emerge as a bias in visualizations. Related to this is the issue of sliveriness: long, splintery shapes as opposed to chunky, more solid shapes. Sliverness of a shape can be measured as the ratio area/perimeter. Shape choices for these two characteristics, as they appear in the survey, are shown in Figure 1.

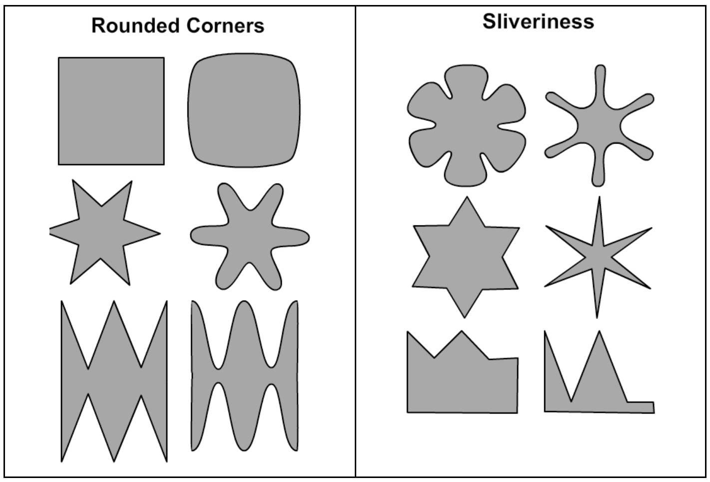

Figure 1. Shapes used to measure preference for a) rounded versus sharp corners, and b) slivery versus chunky shapes.

Because of the studies that have shown people having a preference for symmetrical shapes [9, 10], we decided to examine that characteristic. We also decided to look at convexity versus concavity, expecting that people would prefer convex shapes to those with indentations. Shape choices for these two characteristics, as they appear in the survey, are shown in Figure 2.

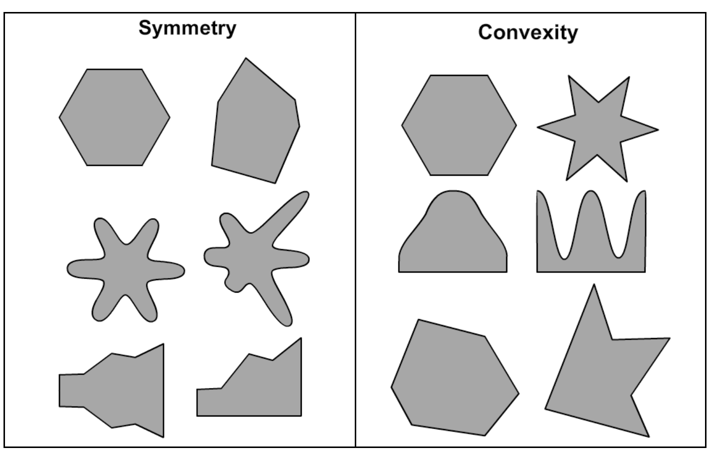

Figure 2. Shapes used to measure preference for a) symmetry versus asymmetry, and b) convex versus concave shapes.

Finally, we also wanted to look at the people's preferences for various orientations. We suspected that people's experiences – with gravity, and with reading time-based charts – might cause them to "read" glyphs as having positive or negative trends. The choices we offered in the survey are shown in Figure 3.

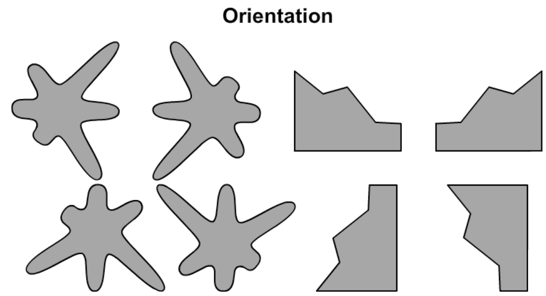

Figure 3. Shapes used to measure preference for varying orientations

III. METHODOLOGY

We designed an online survey to measure people's perceptions of the beneficence (positive) or maleficence (negative) of the shape characteristics described. The survey had four parts to it. First came a consent form, created in accordance with IRB requirements. This was followed by questions for collecting generalized demographic data about the subjects including gender, age group, ethnicity, and education level. The third part was a traditional web-based survey. The fourth part was a casual game, where players were given the same shape choices, but with opportunities to choose the same shape over and over again.

Links to the survey were sent to the instructors of six different undergraduate classes at Stony Brook University. These instructors then forwarded the links to their students. Half of the classes were given a link where they were asked to play the game before answering the traditional web-based survey (G-S); the other half were given a link where the game was played after taking the survey (S-G). Each instance of the game and of the survey recorded the date and time when they were completed, and the IP address of the machine they were accessed from. These data were collected solely for the purpose of correlating the game and survey results.

A. *Survey Design*

The traditional web-based survey used radio buttons to ensure that only one choice is made per question. Qualtrics' survey tool was used to design and collect the data for this, as well as the demographic data. Figure 4 shows an example of such a survey question.

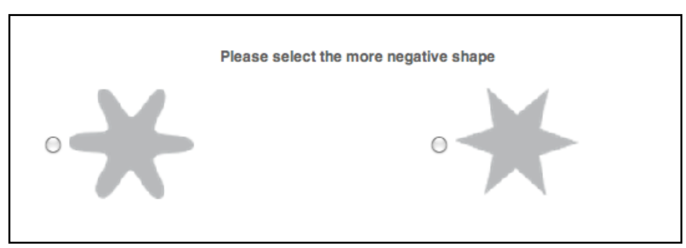

Figure 4. Question from the traditional web-based survey.

B. Game Design

We decided to develop a gamified version of the survey for two reasons. First, we hoped to get the subjects more engaged in making shape choices. Gamification, the strategy of using game design elements to motivate greater participation and retention in non-game contexts, has been successfully used in a variety of other applications that range from marketing to education [11]. Games are currently being used to collect other types of useful information. Zynga has gained notoriety for its use of game data to enhance gameplay and, ultimately, their bottom line through data analytics [12]. We wanted to see how well this would translate to a survey.

Second, we hoped to improve our survey completion rate. Shifting modes of survey research – from face-to-face interviews to telephone surveys to web-based questionnaires – have seen dramatic declines in response rates among potential participants [13]. Low response rates are problematic because of concerns about data integrity as well as cost of collecting the surveys [14].

The casual game version of the survey was developed in JavaScript using the canvas, and can therefore be played on any system with a modern web browser (computer, tablet, or phone) without having to download any software. It collects the player's choices, a time stamp, and an IP address in a MySQL database. A snapshot of the game interface appears in Figure 5.

Figure 5. Snapshot from the gamified version of the survey.

This game begins by instructing players to prevent shapes with negative connotations from hitting the ground. Subsequent screens, representing different levels, show shapes falling from the sky. Shapes captured by the player are shown in the jars below. Feedback from the game (i.e. score, sound effects) encourages players to be consistent in their choices. The game keeps track of all of the shapes selected by each player, enabling us to check for consistency later on.

We developed a scoring scheme that encourages the player to be consistent with his/her shape selection. The game has one level corresponding to each question in the survey. Scores from each level accumulate. For each level we maintain two vector values, V1 and V2. There are at most 4 types of shapes, A, B, C, and D. When the player clicks on shape type A, V1 increments, and type B causes V1 to decrement. Type C causes V2 to increment, and type D makes V2 more negative.

The score of a level is calculated as

$$S_i = | \sqrt{V1^2 + V2^2} | \qquad (1)$$

When the absolute score increases, a short ascending melody is played, and when the score decreases, a descending flat sound is played. A string representing every shape selected by the player for that level is recorded in a database at the end of the level.

IV. Results

A total of 371 students consented to complete the survey of which 82 participated in the G-S order while the other 289 did so in the S-G order. Five students declined to participate. All respondents completed both the web-based survey and played the game over a five-day period. The overall completion rate for both approaches was about 87.6%. The completion rate for the G-S order was 67% while that of the S-G order was 92%. For the S-G order, the response rate for each question was about 94% except for the question that asked respondents to rank their preference for the game or survey, which recorded a response rate of about 90%. The G-S order was in the following sequence: demographic questions, game, and web-based survey. Here, the response rate for each demographic question was about 96%. However, the question response rate dropped to about 69% on the web-based survey. These results suggest that the idea of playing the game acted as an incentive to survey participation and completion. In the G-S order, response rate dropped substantially after the game had been played. In the S-G order however, the response rate was high and consistent at about 92% probably because respondents knew that they had to complete the web-based survey in order to get to play the game.

49.56% or all respondents identified themselves as female. In regards to race, 8.3% identified themselves as African-American/Black, 0.3% as American Indian/Alaska Native, 44.5% as Asian American, 12.76% as Hispanic, 6.2% as Mixed Race, and 27.9% as Caucasian White. 47.7% of respondents were between the ages of 16 and 19 years, 50.9% between 20 and 29 years, and 1.4% identified themselves as between 30 to 39 years old.

TABLE I. PERCENTAGE OF SHAPE CHARACTERISTICS THAT RESPONDENTS PERCEIVED AS NEGATIVE WITHIN FOUR CATEGORIES IN THE SURVEY

Game / Survey Order	**Shape Characteristics Viewed as Negative**			
	Rounded Corners	***Not Convex***	***Slivery***	***Not Symmetrical***
G-S	53.2%	67.6%	72.6%	84.8%
S-G	53.2%	60.9%	74.6%	86%

Looking at results from the traditional survey, respondents viewed shapes that were slivery, not convex, not symmetric, and oriented downwards or to the left as more negative. However, it does not appear that the sharpness of the corners has strong negative connotations. Table 1 summarizes this observation.

Yet even within the categories there were variations. For example, although asymmetric shapes were deemed more negative in two of the cases, the asymmetric shape was considered to be less negative when it showed an upward orientation. Figure 6 shows the ambiguous cases for three categories. If we eliminate these, then the results are as shown in Table 2.

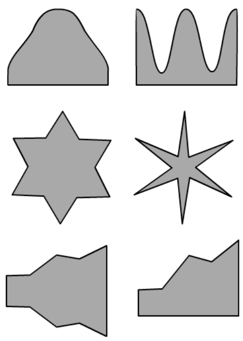

Figure 6. Some shape comparisons produce results differing from the others.

TABLE II. PERCENTAGE OF SHAPE CHARACTERISTICS THAT RESPONDENTS PERCEIVED AS NEGATIVE WITHIN FOUR CATEGORIES IN THE SURVEY, ELIMINATEING AMBIGUOUS CASES

Game / Survey Order	Shape Characteristics Viewed as Negative			
	Rounded Corners	***Not Convex***	***Slivery***	***Not Symmetrical***
G-S	56.5%	70.4%	75%	86.1%
S-G	53.3%	66.5%	76.6%	85.5%

In the game, respondents had the opportunity to choose the shape they viewed as negative more than one time. Because of this, we consider a respondent to have made a selection if more than 75% of their choices were the same. Using this criterion, more than one third of the respondents did not settle on one shape. Of those who did settle on one shape, the results were similar to those obtained by the survey but not as strong. Table 3 shows a summary of these results.

TABLE III. RESULTS OF USING A GAME TO REPLICATE THE SURVEY

	Shape Characteristics			
	Rounded Corners	***Not Convex***	***Slivery***	***Not Symmetrical***
% viewed as negative	52%	63.4%	50.8%	58.2%
% viewed as negative, without ambiguous cases	54.1%	66.1%	54.8%	68.7%
% making inconsistant choices	39.9%	41.5%	34.6%	32.9%

V. CONCLUSIONS

Considering the results of the survey alone, it appears that sliveriness, asymmetry, and concavity are viewed as negative by most people. Downward and leftward trends are also viewed negatively. We suspect that this is due to the training that all of our subjects have had in learning to read graphs. Yet although previous studies had suggested that people perceive sharp corners as negative, our study – using neutral shapes – did not confirm this finding.

Our results suggest that glyphs can be designed to convey goodness – or badness – by taking advantage of these preferences for certain shape characteristics. For example, axes on a star glyph could be arranged so that the optimal values produce a symmetric shape. Furthermore, the sliveriness of the glyph could be modified algorithmically to convey positive or negative information. However, it will always be important to be aware of potentially contradictory shape characteristics, as they will cancel each other out.

We found that the promise of playing a game later seemed to motivate greater participation among respondents, as more completed the survey in the S-G ordering than in the G-S ordering. Playing the game itself also seemed to encourage completion, as more than 90% of respondents made at least one choice for all question levels in the game.

Finally, although the game produced similar results to those from the survey, the preferences indicated are not as strong. To further confound this, a large number of respondents did not consistently choose the same shapes. This may be due to a problem in the game design, where perhaps the respondents were not clear on the objective of the game. Alternatively, it is possible that the consistent choices were the only valid results, and that respondents to the survey may just be picking anything to get it over with; playing the game may help us to identify participants who are really thinking about the answers they are giving. This deserves a great deal more investigation.

ACKNOWLEDGMENT

The authors would like to thank the instructors who participated in this study by giving the links to the survey/game to their students. They are Paul Siegel, Rita Reagan-Redko, Marypat Taveras, Matt Henigman, and Edwin Tjoe.

REFERENCES

[1] Office of Science and Technology Policy, “Big Data Press Release”. Downloaded from http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf.

[2] Johannes Fuchs, Fabian Fischer, Florian Mansmann, Enrico Bertini, and Petra Isenberg, “Evaluation of alternative glyph designs for time series data in a small multiple setting”, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13), pp. 3237-3246, April 2013.

[3] Timo Ropinski, Steffen Oeltze, Bernhard Preim, Survey of glyph-based visualization techniques for spatial multivariate medical data, Computers & Graphics, Volume 35, Issue 2, pp. 392-401, April 2011.

[4] Lanka, Y.S. and Swan, J.E., “An Evaluation of Glyph Perception for Real Symmetric Traceless Tensor Properties”, Computer Graphics Forum, vol. 29, no. 3, pp. 1133-1142, June 2010.

[5] Paul J. Silvia and Christopher M. Barona, "Do people prefer curved objects? Angularity, expertise, and aesthetic preference", Empirical Studies of The Arts, volume 27, issue 1, pp. 25-42, 2009.

[6] Bar, Moshe, and Maital Neta. "Humans prefer curved visual objects." Psychological science 17.8 (2006): 645-648.

[7] Leder, Helmut, Pablo PL Tinio, and Moshe Bar. "Emotional valence modulates the preference for curved objects." Perception-London 40.6, p. 649, 2011.

[8] Bohrn I, Nabecker G, Carbon C C, "Are curved visual objects always preferred?" Perception 37 ECVP Abstract Supplement, p. 75, 2008.

[9] Reber, R, Schwarz, N and Winkielman, P, "Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience?", Personality And Social Psychology Review, Volume 8, Issue 4, pp. 364-382, 2004.

[10] Makin, ADJ, Pecchinenda, A and Bertamini, M, "Implicit Affective Evaluation of Visual Symmetry", Emotion, Volume 12, Issue 5, pp. 1021-1030, October 2012.

[11] Deterding, S., Dixon, D., Khaled, R. and Nacke, L. From game design elements to gamefulness: Defining "gamification". *Proc. of MindTrek '11*, ACM Press (2011), 9-15.

[12] Wingfield, N. Virtual products, real profits. *The Wall Street Journal*, Sept. 9, 2011.

[13] Groves, R.M. Three eras of survey research. *Public Opinion Quarterly*, Vol. 75, No. 5, 2011, pp. 861–871.

[14] Fan, W. and Yan, Z. Factors affecting response rates of the web survey: A systematic review. *Computers in Human Behavior* 26(2010) 132-139.