Evaluating the Impact of Shape and Metric Selection on Human Perception in Geospatial Data Visualizations

Nicole R. Schneider*
nsch@umd.edu
University of Maryland
College Park, Maryland, USA

Harsh Patel*
hpatel01@terpmail.umd.edu
University of Maryland
College Park, Maryland, USA

Hanan Samet
hjs@cs.umd.edu
University of Maryland
College Park, Maryland, USA

Abstract

Visualizations such as bar charts, scatter plots, and objects on geographical maps often convey critical information, including exact and relative numeric values, using shapes. The choice of shape and method of encoding information is often selected arbitrarily, or decided based on convention or common practice. However, past studies have shown that the human eye can be fooled by visual representations. The Ebbinghaus illusion demonstrates that the perceived relative sizes of shapes depends on their configuration, which in turn can affect judgements, especially in visualizations like proportional symbol maps. In this study we evaluate the effects of varying the type of shapes and metrics for encoding geospatial data in visual representations on a spatio-temporal map interface. We find that some combinations of shape and metric are more conducive to accurate human judgements than others, and we provide recommendations for applying these findings in future spatial visualization designs.

CCS Concepts

• **Human-centered computing** → **Empirical studies in HCI**; Graphical user interfaces; **Geographic visualization**; Information visualization; **Empirical studies in visualization**; • **Applied computing** → *Health informatics*; • **Information systems** → Geographic information systems.

Keywords

Visual perception, proportional symbol maps, visual illusion

ACM Reference Format:
Nicole R. Schneider, Harsh Patel, and Hanan Samet. 2024. Evaluating the Impact of Shape and Metric Selection on Human Perception in Geospatial Data Visualizations. In *5th ACM SIGSPATIAL International Workshop on Spatial Computing for Epidemiology (SpatialEpi'24), October 29-November 1, 2024, Atlanta, GA, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3681777.3698468

1 Introduction

Visualizations such as bar charts, scatter plots, and objects on geographical maps often convey critical information, including exact and relative numeric values, using shapes. Given a dataset, there are many valid choices that can be made about how to present the information, but there are relatively few definitive conclusions about what shapes and metrics should be used to promote accurate perception of the data shown in the visualization. In fact, the choice of shape and method of encoding information is often made arbitrarily, or based on convention. However, past studies have shown that the human eye can be fooled by visual representations. The famous Ebbinghaus illusion indicates that the perceived relative sizes of shapes depends on their configuration, which in turn can affect judgements, especially in visualizations like proportional symbol maps. However, there is a gap between the known visual illusion and the application of that knowledge to selecting appropriate representations for visualizing dynamic spatio-temporal data.

In other words, there is a lack of concrete knowledge about the impact of choices in visual representation, such as selection of shape and encoding metric on a proportional symbol map. Given that these choices may influence visual perception of the data being portrayed, seemingly benign decisions may impact the message a reader takes away from a visualization. In the worst case, this could have significant impacts in a variety of domains, including infectious disease mapping, where funds are often appropriated based on disease spread and severity. During pandemics, when many areas are in need of funding, it is critical that decision-makers have an accurate understanding of case, death, and vaccination rates, which are often conveyed using proportional symbol maps. By studying visual perception of COVID-19 data, we can better understand how visualization choices affect whether the reader can accurately interpret the information being presented, which will lead to better-informed visualization choices for future epidemics.

We present a study that evaluates the effects of varying the type of shape and metric for encoding data in visual representations on a spatio-temporal map interface. We use real data encoded in various ways on a production data visualization and exploration system for tracking COVID-19 related statistics through space and time. We perform a user study to determine which combinations of symbol shape, encoding metric, and type of variation targeted (spatial or temporal) promote the most accurate perceptions in the context of this system. We measure participant responses to a survey of multiple choice questions and analyze the results to draw conclusions about the conditions under which people make more or less accurate judgements of the relative sizes of shapes on a map visualization.

With respect to encoding metrics, we hypothesize that metrics requiring little mental manipulation, such as diameter of a circle, will yield better participant performance than metrics requiring substantial mental manipulation, such as circumference of a circle.

*Both authors contributed equally to this research.

SpatialEpi'24, October 29-November 1, 2024, Atlanta, GA, USA

ACM ISBN 979-8-4007-1153-4/24/10
https://doi.org/10.1145/3681777.3698468

Intuitively, we are positing that because diameter simply consists of estimating the distance between two points on the circle, participants will be able to make this estimation more accurately than they can do for more complex metrics like perimeter or area. Regarding spatio-temporal variation, we hypothesize that questions depicting spatial variation on a single map will yield more accurate estimates of relative size than questions depicting temporal variation using side by side map snapshots, which require participants to look back and forth between two separate maps with identical background configurations.

Our main contributions are the findings we show that indicate some combinations of shape and metric are more conducive to accurate human judgements than others. We also detail our methodology which can easily be extended to test different shapes and metrics, or new attributes altogether, depending on the context under study. We also provide recommendations for applying our findings in future map visualization design, especially in spatio-temporal mapping of epidemiological data.

The rest of the paper is organized as follows. In Section 2 we present a review of previous work. We then describe the system used to contextualize our research questions in Section 3, and our methodology in Section 4. Finally, we present results in Section 5, a discussion of the findings in Section 6, avenues for future work in Section 7, and conclusions in Section 8.

2 Related work

In this section we outline related work in visual perception as it pertains to mapping and map visualization.

2.1 Visualization and Visual Perception

Visualizations are a common method for representing data in an easily digestible manner, which has gained new interest in recent monitoring of disease spread and pandemics [32, 33]. Visualized data can be encoded in any number of ways, which varies depending on the type of visualization and the data. However, all visualizations rely on visual perception as a key underlying principle.

Visual Perception. There is a sizeable body of work dealing with visual perception and visualizations, including perceptual correspondence between data and its visualization [8]. Healy et al. [19] gives a survey of visual attention and memory, explaining principles related to the visual system, and what it sees and misses in different scenarios. Heer et al. [20] shows that Amazon Mechanical Turk (MTurk), the same crowdsourcing platform we use, is a viable way to conduct many visualization perception studies.

Illusions in Visualizations. When presenting data in visual form, one must be cognizant of the phenomenon of visual illusions. There are many such illusions, including the widely studied Ebbinghaus Illusion, where the perceived size of a circle can be influenced by the surrounding circles, as well as a number of other factors [26]. Visual illusions have been studied under a variety of conditions in the psychology literature [9], as well as in the data visualization field [22, 28] and more recently in the context of virtual reality (VR) [1, 13]. In the latter, characteristics of data are represented visually using shape, surface properties, and motion through VR. This can lead to illusions in how geometric structures are perceived due to their properties in the VR world [1]. From the data visualization perspective, there is recent work [22] detailing the systematic bias in tri-variate scatter plots, when encoding a third dimension of information in size or color. Visual perception is shown to be sensitive to choice in size or color range, which leads to misjudgements. All of this work serves as motivation for studying which kinds of objects and object properties lead to the most accurate visual perceptions in data visualization.

2.2 Spatial, Temporal, and Cartographic Visualizations

Many works have focused on developing or modifying spatial visualizations in novel ways, to allow for easier human perception [2, 12]. In particular, Drocourt et al. [12] develop an algorithm for visualizing the advancement/retreat of glaciers in Greenland using radial lines and nested rings. They use a nonlinear mapping to generate angular coordinates from Cartesian coordinates, which allows for consistent spatial perception. Their work represents a unique use of circles in spatial visualization, where arc length (a segment of the circumference of a circle) conveys vital information. In our work, we find that the circumference of a circle is one of the most visually challenging metrics for participants to estimate accurately on maps (out of the 3 metrics and 2 shapes we tested across both spatial and temporal questions).

Temporal visualizations can be constructed in a number of different ways [7], but often treat time as an additional axis, or include animation or interactivity to convey changes over time. Examples include Hao et al. [17], which develops an interactive display of large molecule datasets in biology. Other previous work has studied the effectiveness of different temporal encoding mechanisms, including small multiples, cartograms, and proportional symbols [29]. In our study, the system we use to generate the images we present to participants is a spatio-temporal map interface that uses proportional symbols to convey metrics, meaning it incorporates aspects of both spatial and temporal visualization. We further describe the context in which we compare spatial and temporal visualization queries in Section 4.1.1.

Cartographic visualization has evolved substantially with the rise of modern post-computing mapping [24]. In particular, the rapid development of mapping applications spawned the study of cartographic interaction, which is the subject of Roth [31]. The link between scientific visualization and cartographic visualization is thoroughly discussed in Maceachren and Kraak and Fairbairn et al. [14, 25].

2.3 Role of Shapes and Encoding Metrics in Map Visualization

One of the key aspects of cartographic visualization that we test in our study is choice of shape. We know that the observable size of a circle can be influenced by factors such as the size difference between a target circle and adjacent circles in a close proximity [15]. However, many visualizations use circles as the primary shape to represent data, especially in a geographical context [3, 27]. There are several works that study the perception of shapes in a geographical context [6, 23, 36]. Many of these works indicate that choice of shape is important, in addition to other factors like background and dimensionality of encoding metric which may also influence perception. Cleveland et al. studies visual judgements of the relative areas of circles on a plain background, finding that participant judgements were consistent [4]. On the other hand, Raidvee et al. finds that humans tend to perceive the size of a circle proportionally to its diameter [30]. Stachon et al. [36], studies the effect of shape (circle vs. triangle) on the speed of processing when a map background is present and not present. Another study [6] finds that graduated squares built on the basis of area rather than a linear dimension were estimated accurately. Groop et al. [16] finds that overlap of circles also affects the perception of relative sizes and proposes transparent overlap rather than partial occlusion to help combat these effects. Legend values have also been explored as a way to improve the effectiveness of a geographic display that uses circles to represent data [10]. However, if done incorrectly, the legend values can inhibit the conversion of information. Cox et al. assesses the efficacy of value scaling against the use of numerous legend symbols for both circles and squares [5]. Results showed that the use of various legend symbols on a map yielded more correct shape estimations compared to the apparent value scaling.

2.4 Map Animation

Although map animation is outside the scope of our study, it is a common way to convey temporal (and non-temporal) geographic data. Early work on map animation to represent geographic-scale change was done by Harrower [18]. See also Slocum et al. [35] for an in-depth analysis of geovisualization and the evolution of map animation. Studies have also devised methods to improve comprehension of multivariate geographic data [11, 39] and make animated maps easier to comprehend by reducing cognitive overload [41].

3 CoronaViz System

CoronaViz [1] [2] [33] is a dynamic COVID-19 disease visualization system that was created in light of the coronavirus pandemic as a new way to track and visualize pandemic-related data over time. The system displays various data including confirmed cases, deaths, recoveries, hospitalizations, positivity rate, and vaccination rate on a single interactive and multi-layer geographical display. The data displayed corresponds to specific locations on a map that allows the user to select, hover, zoom in or out, and pan. This system builds on many of the key principles of interactive map interfaces described in Teitler et al. and Sankaranarayanan et al. [34, 37], and differs from many of the COVID-19 reporting visualizations [38, 40] in that it enables spatial and temporal comparisons in one interface, while also supporting zooming which increases the resolution of data presented, as additional smaller units become visible.

In its base configuration, the graphical interface for CoronaViz represents data using hollow circles which we call 'geocircles' whose diameter are determined by the values of the variables they represent. Animation control buttons allow users to search through time manually or view the data unfolding in accelerated time, giving a summary of the temporal changes in the data. However, for users to glean an accurate picture of pandemic status and progression through time, they must be able to accurately estimate the relative size of a geocircle, given the other geocircles visible on the map, as well as the size of the same geocircle at a different (no longer visible) point in time. This raises the main question of our study:

What shape and encoding metric should be used to allow for the most accurate perception of numerical values or relative numerical values on a dynamic map interface like CoronaViz?

To address this question, in this study we vary the shape of the geocircles (circle or square) and the encoding metric (diameter/side length, circumference/perimeter, and area) used to visualize data in the interface, and measure participant responses.

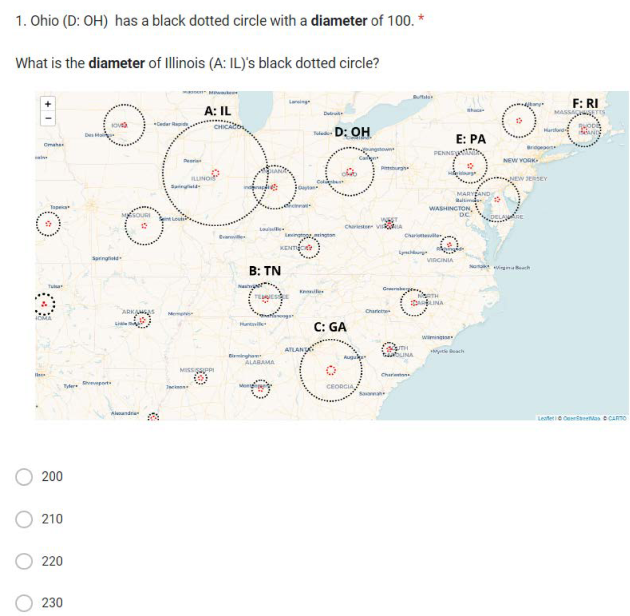

Figure 1: Map query question using a circle to encode diameter across spatial variation.

4 Method

In this section we describe our study methodology, including the survey design and participant recruitment.

4.1 Description of the Survey

For this study we developed a survey [3] that is made up of 9 parts. Parts 1 and 2 consist of the consent form and a question asking the user to provide their MTurk Worker ID. Parts 3, 4, and 5 consist of the 24 main survey questions, which are multiple choice style. These questions present the participant with one or two map visualizations and ask them to estimate, by eye, the relative sizes of shapes on the map(s). These questions are described in full detail in Section 4.1.1. There is also an attention check question mixed in with the aforementioned 24 questions. The attention check is used to decide which responses are good faith attempts, and which are the result of random guessing (we discard these responses). More details about the content of this question are given in Section 4.1.2. Part 6 contains two questions which inquire participants'

[1] https://coronaviz.umiacs.io/
[2] https://coronaviz.umiacs.io/squares/
[3] https://github.com/nicoleschneider/coronaviz

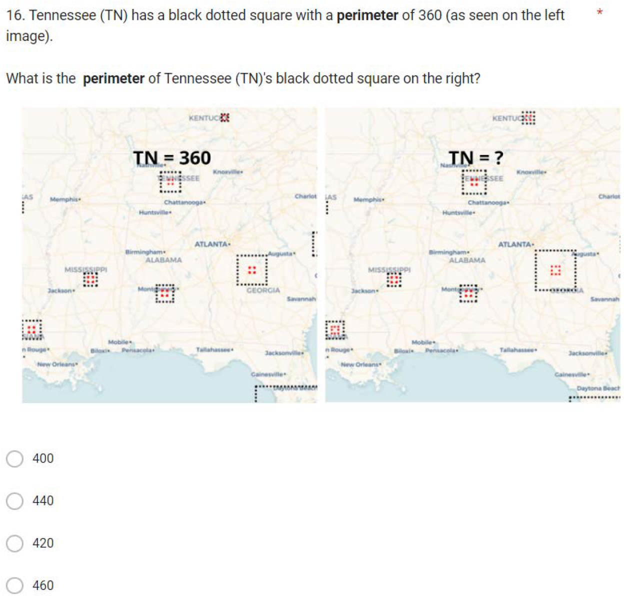

Figure 2: Map query question using a square to encode perimeter across temporal variation.

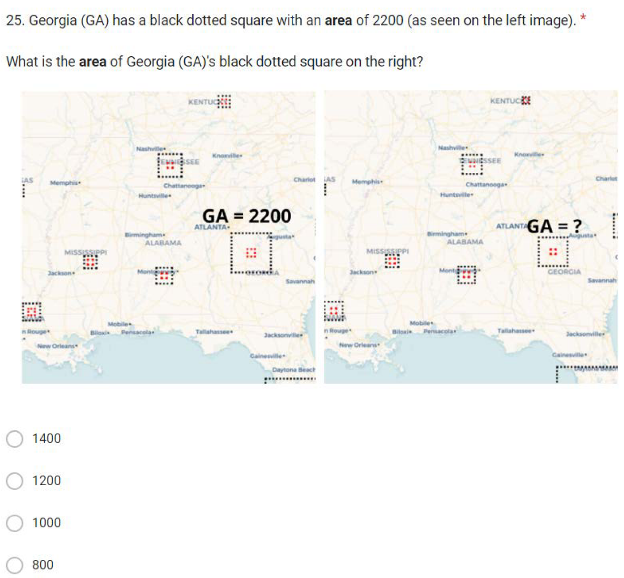

Figure 3: Map query question of using a square to encode area across temporal variation.

opinions about which shape types and metrics they found easiest to estimate. Part 7-9 record additional information, including demographic breakdown, and provide a survey completion code.

4.1.1 Map Query Questions. The map query questions comprise the majority of the survey. Each of these questions provide the participant with one or two images of the CoronaViz map interface, which includes several shapes representing COVID-19 statistics by location. However, the underlying statistics that drive the sizes of the shapes are hidden from the participants. Instead, the only numerical value visible on the map is a label for the *reference shape*. The participants are also told which metric was used to encode this value (area, circumference/perimeter, or diameter/side length). Using those two pieces of information, the participants are asked to visually estimate the relative size of a second shape on the map, termed the *query shape*, which is labeled with a question mark '?'. Each map query question is designed to evaluate the participant's ability to estimate the relative size of shapes in situations that vary across several attributes of interest, which are described in further detail below.

Spatial Questions. All of the map query questions that we presented to the participants fell into one of two categories, spatially-focused questions and temporally-focused questions. The first type (termed spatial questions), present the participant with scenarios where they must estimate the size of a shape in one location, given a reference shape in a *different location* on the same map. An example of one such question is given in Figure 1. These questions measure the participant's ability to visually estimate variation in shape size across space.

Temporal Questions. The other type of map query questions, temporal questions, present the participant with two distinct maps that have identical background perspectives (Figure 2). These two maps represent two snapshots of one location undergoing temporal animation. In other words, we use static side-by-side images that, when considered in tandem, convey a temporal variation in shape size. This design choice allows us to isolate an important difference between spatial and temporal variation: the fact that temporal variation consists of a single center point, around which a shape is changing size, whereas spatial variation consists of both changing shape size, and changing center location.

Shapes. To represent the numeric values associated with different locations on the map, each of the map query questions presents one of two types of shapes, whose sizes are scaled according to their encoding metric (see Section 4.1.1). These values are either presented using *circles* (ex. Figure 1) or *squares* (ex. Figure 3). Circles are the canonical choice for map symbols [3, 27], but squares have been shown to allow for good visual estimation under certain conditions [6].

Encoding Metric. We also vary the metric used to encode the numerical values for the shapes presented in our survey. For each shape type, we encode numbers using diameter/side length (ex. Figure 1), circumference/perimeter (ex. Figure 2), and area (ex. Figure 3). These three encoding metrics cover the typical ones used in spatio-temporal map visualizations based on our review of the geovisualization literature.

4.1.2 Attention Check Question. Disguised within the map query questions is one attention check question (Figure 4). This question is designed so that the correct answer is easily derivable from the information presented in the summary box overlaid on the map image. We include this question to identify responses in which the participant randomly guessed from those where the participant made a good faith attempt to answer each question. We eliminate from consideration the entire survey response for any participant who did not answer the attention check question correctly.

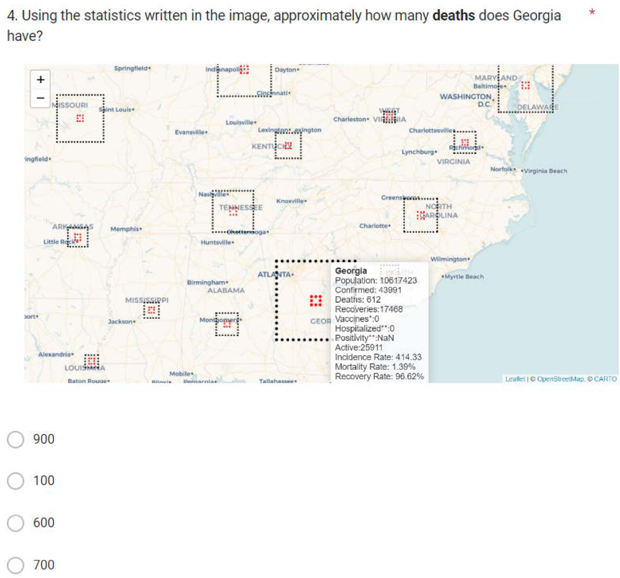

Figure 4: Attention check question used to discard survey responses made via random guessing. The question has an obvious correct answer of 600.

4.1.3 *Opinion Questions.*

After completing the map query questions we ask participants to provide their opinions about which shape and encoding metric they found easiest to estimate. These questions are designed to gather participant feedback that can be directly compared to their actual success rates for the map query questions, to determine how well the attitudes about shapes and encoding metrics align with the actual performance across these attributes. Participants were also asked to provide optional demographic information. This includes a question for gender, highest degree completed or in progress, and age range.

4.2 Recruiting Participants

We recruited 24 participants for the study using the Amazon Mechanical Turk (MTurk) crowd-source platform. We limited the task to only allow crowd-workers located in the United States. When workers accepted our task, they used the survey link provided in the task to access our survey. [4] After answering all of the questions in the survey, workers received a completion code that they then provided via the MTurk interface to complete the job. We paid workers $2.50 for completing the task, which is a target of about $8-10 per hour based on the number of questions in the survey and our estimates for the time it would take to complete the survey.

4.3 Selecting Reference and Query Values

We chose the reference and query values for the questions so that the two questions per condition (i.e. the two questions for Circle, Area, Spatial) cover both the case where the reference value is larger than the query value and also the case where the reference value is smaller than the query value. We did so to neutralize any effects due changes in difficulty of estimating a larger value given a smaller one, versus estimating a smaller value given a larger one.

We also chose values appropriately sized to the task. For instance, values encoded with the Area metric were larger across the board, so that the overall sizes of the shapes stayed relatively similar across all questions. While we allowed some variation in shape size which is natural within the CoronaViz platform, we ensured that no reference or query shape consisted of more than approximately one quarter of the map background, to help combat any effects that may arise out of difficulty estimating very large shapes on the map.

4.4 Selecting Reference and Query Positions

We recognize that distance between reference and query shapes may impact the difficulty of estimating their relative sizes. For temporal questions, the reference and query shapes were presented in two identical side by side map backgrounds. This means the distance between the centers of these two shapes is constant across all temporal questions, since the map sizes are held constant from question to question. In our survey this distance was approximately 1040 pixels. Of note, the nature of the two maps side by side requires a small visual break between the maps, which in turn means that the distances for temporal questions were higher than for spatial questions. We discuss the implications of this in Section 6.

For spatial questions, it is more difficult to keep a consistent distance between reference and query shapes, while also maintaining a variety of locations to combat learning effects from question to question. We settle on a middle ground by varying the locations on the map, but ensuring that all distances from reference to query shape (measured center to center) are between 200 and 900 pixels. Further, for any one condition tested, there are always two questions for which the scores are averaged. We ensure that the average distance between reference and query for any condition is between 450 and 600 pixels.

4.5 Setting Multiple Choice Options

For simplicity we rounded all correct answers to the nearest multiple of 10, and chose incorrect (distractors) that were also multiples of 10. We used a pixel ruler [5] to measure the actual diameter or side length of each reference and query shape, and used that to mathematically calculate circumference/perimeter or area if applicable for the question. To account for annotator error in measuring the values, we used the pixel ruler five times for each question computed the average before rounding to the nearest 10. We also scaled the answers and distractors down to a range we thought participants could reasonably be asked to estimate: at or below 4000 for area, 2000 for perimeter/circumference, and 600 for diameter/side length.

5 Results

We collected responses from twenty-four participants, of which we retained and report results on the sixteen responses that passed the attention check question. We summarize the results from parts 3, 4, and 5 of the survey in Tables 1 and 2.

[4] https://github.com/nicoleschneider/coronaviz

[5] https://www.rapidtables.com/web/tools/pixel-ruler.html

		Success Rate (%)
Shape type	Circle	25.5
	Square	**28.1**
Encoding metric	Diameter/Side length	**29.7**
	Perimeter	21.1
	Area	**29.7**
Variation	Spatial	**32.3**
	Temporal	21.4

Table 1: Overall summary of participant performance across attributes of interest for the sixteen participants who passed the attention check question. Note these attributes are not disjoint.

	Spatial Results		Temporal Results	
	Circle	Square	Circle	Square
Diameter/Side length	40.6	37.5	25.0	15.6
Perimeter	15.6	43.8	12.5	12.5
Area	37.5	18.8	21.9	40.6

Table 2: Success rate (%) of participants estimating shapes across 3 attributes of interest. Results are given as the average over 2 questions for each combination of attributes. There are 12 disjoint combinations tested.

5.1 Opinion Questions

Out of the sixteen participants, seven said circles were the easiest to estimate, five said squares were easiest, and the remaining four said they were of equal difficulty. For the encoding metrics, six participants found circumference/perimeter to be the easiest, five said diameter/side length was easiest, one thought area was easiest, and four found them to be of the same difficulty.

5.2 Analysis of Map Query Questions

A summary of the overall participant performance is presented in Table 1 and a summary of the performance for each disjoint combination of conditions is presented in Table 2. Each participant answered two questions for each combination of shape, metric, and spatial/temporal question type. This means that for each scenario (such as Circle-Area-Spatial or Circle-Diameter-Temporal) a participant could have answered 0, 1, or 2 of the questions correctly. We aggregate these to determine the number of correct responses per participant for each of the tests we perform (Circle vs. Square, Spatial vs. Temporal, etc.).

For all statistical tests we apply the Bonferroni correction [21] to adjust the significance level required to reject the null hypothesis, since we are performing multiple hypothesis tests on the same dataset. Rather than testing all possible combinations of shape, metric, and question type, we select a few based on the RQs outlined in Section 1 and the participant feedback discussed in Section 5.1. We first perform the Shapiro-Wilk test for normality for each pair of scenarios we test, which showed in each case that the data was not normally distributed ($p < 0.05$). As a result, we used the non-parametric test Wilcoxen Signed Rank test to test each of the following hypothesis.

Shape. In our first test we aimed to find out if squares are easier to estimate than circles. For each participant we count the number of correctly answered questions which used the circle shape, and then number of correctly answered questions which used the square shape. We then use the Wilcoxen Signed Rank test to determine if the median difference is zero (null hypothesis) or if it is not zero (alternative hypothesis). We discard the ties and find $p > \alpha$, indicating no significant difference in median between the two groups. *This means our survey showed that varying the shape type, circle vs. square, did not significantly affect viewers' ability to glean accurate information from the map visualization.*

Spatial/Temporal. In the next test we aimed to find out if spatial type questions are easier to estimate than temporal ones. For each participant we count the number of correctly answered spatial and temporal questions and use the one tailed Wilcoxen Signed Rank test to determine if the median difference between scores for spatial and temporal questions is zero or greater than zero. We discard the ties, and find that $p < \alpha$, meaning that spatial type questions are significantly easier than temporal type questions. *This means when presented with side by side maps capturing temporal variation, users have more trouble discerning the information encoded, as compared to a similar visualization that varies spatially (on the same map).*

Encoding Metric. In the opinion questions we found that participants reported having the easiest time estimating circumference/perimeter and diameter/side length for metrics and circle for shape. To explore this phenomenon in the context of spatio-temporal map visualization, we test to see if one of these two metrics (diameter or perimeter) is easier to estimate accurately for circles. We again use the Wilcoxen Signed Rank test to determine if the median difference between scores for Circle-Diameter and Circle-Perimeter questions is zero or greater than zero. We discard the ties and find that $p < \alpha$, meaning that the diameter of a circle is significantly easier to estimate than perimeter of a circle. *Our finding that diameter of a circle in our spatio-temporal interface is easier to estimate than circumference of a circle in the same setting is consistent with the literature on visualization that indicates that people tend to perceive the size of a circle proportionally to its diameter [30].*

5.3 Demographics

The majority of the participants, eleven, where between the ages of 21-30. Three were between the ages of 31-40 and two were between 41-50. Half of the participants were male and the other half are female. We had one individual with a high school degree or equivalent, three who had an associates degree, seven with a bachelors degree, and five who hold a masters or professional degree.

5.4 Summary

Based on the results of our three tests and the opinions of the participants, it is clear that no particular shape or metric is easier to estimate across the board. We found that squares were slightly (but not statistically significantly) easier to estimate than circles. On the other hand, the opinion questions indicated that more participants thought circles were easier to estimate than squares. This potential disconnect between what participants think is easier to estimate and what they are better at estimating in practice is an interesting avenue of future study. For spatial and temporal questions, we found that the spatial questions were significantly easier to estimate than temporal questions, supporting our hypothesis that the visual separation and extra distance between the snapshots provided in the temporal questions made them more difficult than the spatial questions, which presented only a single map to look at. For circle questions in particular, we observed that participants estimated diameter significantly more accurately than they estimated perimeter. This is also in line with our hypothesis that metrics requiring less mental manipulation, like diameter, would be easier to estimate than metrics requiring more complex manipulations, like unfolding the circumference of the circle and estimating its length.

6 Discussion

Looking at the overall results, squares yield slightly higher performance than circles, diameter/side length and area have better performance than perimeter, and spatial questions have better performance than temporal ones. Looking at the disjoint combinations, we see that side length and perimeter perform similarly for squares in both spatial and temporal questions. This fits with the intuition that for squares, perimeter is simply a 4x multiple of side length, which should be just as easy (or difficult) to estimate. No such simple relationship exists for circles, which show more mixed results depending on the encoding metric and question type.

6.1 Limitations

For this study we surveyed 25 participants using MTurk. However, with 9 participants failing the attention check question, we only retained 16 responses to use in analysis. With this relatively small sample size, we were only able to find significant differences between a few combinations of shape, metric, and spatial/temporal question type. By design, our survey is also limited in its ability to test temporal queries in particular. We chose to design the temporal question to test one aspect of temporal changes, the change in shape size while center point holds steady, and ignore other aspects that make temporal questions challenging. This includes the need to remember, rather than reference, the previous representation. In our study participants could look back and forth between one snapshot and another to estimate the difference. However, even with this advantage, we found that performance on spatial questions was significantly better than on the temporal snapshots. We attribute this to the added distance between the query and reference shape for the temporal questions, which came about as a result of presenting two maps side by side with a small visual break between them. A future study could be designed to incorporate animation, which better captures the complexities of temporal queries, and is already supported in the CoronaViz map query interface.

We also consider that by deploying the survey on a crowd-source platform like MTurk, we have no control over the resolution of the screens used while taking the survey. This is an inherent limitation to all visualization studies deployed in this manner, and is discussed extensively in Heer and Bostock [20]. In the context of spatio-temporal visualization on a production COVID-19 visualization system, variation in the size and quality of screen used to view the system is expected, and so this setting for the study is natural to the context of interest, even though it introduces uncertainty.

Finally, in designing the questions, we made trade-offs with respect to allowing or controlling variation of distance between the reference and query point. We decided to keep the distances within some reasonable bounds, rather than allowing complete variation. Ideally, the distances should be held constant from question to question to eliminate possible confounding effects, but this undermines the natural variability intrinsic to a real system like CoronaViz. Since we did notice that temporal questions led to significantly worse performance than spatial ones, we suggest as future work a study that explicitly measures the effects of reference-query distance in a map setting like this one.

7 Future Work

There are a few avenues of future work that we believe would enhance the results presented in this study. One aspect of the CoronaViz interface that we did not address directly here is the presentation of multiple metrics per location using concentric shapes of different colors. It would be interesting to study how well people are able to estimate the relative sizes of the outer and inner shape, to determine if this is indeed a useful way to convey multiple data values per location. For the purposes of this study, we used static images taken from a graphical interface. This gave the user a visual to compare the encoded object to another one at all times. To account for this, future studies can be conducted where a user is shown a temporal animation or a GIF from the graphical interface and asked to determine how the object changed over time instead of having the original reference object statically viewable. For spatial questions, a user might be asked to compare the relative sizes of objects that are either close together or distant. Further research should be conducted to find what shapes and metrics allow for the most accurate perception, as we have shown that data encoded as circles by convention may not be the best in all scenarios.

8 Conclusion

Previous work has shown that visual perception can be influenced by a number of factors, including the type of shape being viewed and the background it is viewed on. With this in mind, we studied how well people are able to visually estimate the relative sizes of different shapes in cartographic visualizations taken from a real system for visualizing COVID-19 data. We varied the choice of shape, metric by which numbers are encoded visually, and type of variation depicted across the reference and query shapes: either spatial variation or temporal variation (via side by side snapshots). We found that when using circles as the visualization shape, diameter was significantly easier to estimate than circumference. We also found that participants more accurately estimated relative sizes for spatial queries than for temporal ones, which we believe is attributable to the increased distance between the reference and query object in the temporal question setup. Ultimately, we have shown that choice of shape and metric makes a measurable difference in how map visualizations are perceived by viewers. As a result, we hope that these findings spur further research along the lines we have suggested and encourage scientists as well as cartographers to consider carefully how they present numerical data in spatio-temporal map visualizations.

Acknowledgments

This work was sponsored in part by NSF Grants IIS-18-16889, IIS-20-41415, and IIS-21-14451. Special thanks to Terry Slocum and Brian Ondov for their input.

References

[1] 2003. *Looking Just Like It Doesn't: Perception and Illusion in Scientific Visualization.* Fluids Engineering Division Summer Meeting, Vol. Volume 2: Symposia, Parts A, B, and C. https://doi.org/10.1115/FEDSM2003-45198 arXiv:https://asmedigitalcollection.asme.org/FEDSM/proceedings-pdf/FEDSM2003/36975/1629/2593161/1629_1.pdf

[2] Richard A. Becker, Stephen G. Eick, and Allan R. Wilks. 1995. Visualizing Network Data. *IEEE Trans. Vis. Comput. Graph.* 1 (1995), 16–28.

[3] William S. Cleveland, Charles S. Harris, and Robert McGill. 1982. Judgments of circle sizes on statistical maps. *J. Amer. Statist. Assoc.* 77, 379 (1982), 541–547. https://doi.org/10.1080/01621459.1982.10477844

[4] William S Cleveland, Charles S Harris, and Robert McGill. 1982. Judgments of circle sizes on statistical maps. *J. Amer. Statist. Assoc.* 77, 379 (1982), 541–547.

[5] Carleton W. Cox. 1976. Anchor effects and the estimation of graduated circles and squares. *The American Cartographer* 3, 1 (1976), 65–74. https://doi.org/10.1559/152304076784080195

[6] Paul V. Crawford. 1973. The perception of graduated squares as cartographic symbols. *The Cartographic Journal* 10, 2 (1973), 85–88. https://doi.org/10.1179/caj.1973.10.2.85

[7] Chaouki Daassi, Laurence Nigay, and Marie-Christine Fauvet. 2005. A taxonomy of temporal data visualization techniques. *Information-Interaction-Intelligence* 5, 2 (2005), 41–63.

[8] Mehdi Dastani. 2002. The Role of Visual Perception in Data Visualization. *Journal of Visual Languages & Computing* 13, 6 (2002), 601–622. https://doi.org/10.1006/jvlc.2002.0235

[9] Matthieu M. de Wit, John van der Kamp, and Rob Withagen. 2015. Visual illusions and direct perception: Elaborating on Gibson's insights. *New Ideas in Psychology* 36 (2015), 1–9. https://doi.org/10.1016/j.newideapsych.2014.07.001

[10] Micheal W Dobson. 1974. Refining legend values for proportional circle maps. *Cartographica: The International Journal for Geographic Information and Geovisualization* 11, 1 (1974), 45–53. https://doi.org/10.3138/ktw4-7562-6811-8h21

[11] Daniel Dorling. 1992. Stretching Space and Splicing Time: From Cartographic Animation to Interactive Visualization. *Cartography and Geographic Information Systems* 19, 4 (1992), 215–227. https://doi.org/10.1559/152304092783721259 arXiv:https://doi.org/10.1559/152304092783721259

[12] Yoann Drocourt, Rita Borgo, Kilian Scharrer, Tavi Murray, Suzanne Bevan, and Min Chen. 2011. Temporal Visualization of Boundary-based Geo-information Using Radial Projection. *Comput. Graph. Forum* 30 (06 2011), 981–990. https://doi.org/10.1111/j.1467-8659.2011.01947.x

[13] Mie Egeberg, Stine Lind, Niels C. Nilsson, and Stefania Serafin. 2021. Exploring the Effects of Actuator Configuration and Visual Stimuli on Cutaneous Rabbit Illusions in Virtual Reality. In *ACM Symposium on Applied Perception 2021* (Virtual Event, France) *(SAP '21)*. Association for Computing Machinery, New York, NY, USA, Article 1, 9 pages. https://doi.org/10.1145/3474451.3476230

[14] David Fairbairn, Gennady Andrienko, Natalia Andrienko, Gerd Buziek, and Jason Dykes. 2001. Representation and its Relationship with Cartographic Visualization. *Cartography and Geographic Information Science* 28, 1 (2001), 13–28. https://doi.org/10.1559/152304001782174005

[15] Patricia P. Gilmartin. 1981. Influences of map context on circle perception. *Annals of the Association of American Geographers* 71, 2 (1981), 253–258. https://doi.org/10.1111/j.1467-8306.1981.tb01351.x

[16] Richard E Groop and Daniel Cole. 1978. Overlapping graduated circles / magnitude estimation and method of portrayal. *Cartographica: The International Journal for Geographic Information and Geovisualization* 15, 2 (1978), 114–122. https://doi.org/10.3138/q5q5-n244-8462-ng25

[17] Xuejun Hao, Amitabh Varshney, and Sergei Sukharev. 2004. Real-Time Visualization of Large Time-Varying Molecules. (08 2004).

[18] Mark A. Harrower. 2002. *Visual benchmarks: Representing geographic change with map animation.* Ph.D. Dissertation. https://www.proquest.com/dissertations-theses/visual-benchmarks-representing-geographic-change/docview/275797248/se-2 Copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated - 2022-01-18.

[19] Christopher Healey and James Enns. 2012. Attention and Visual Memory in Visualization and Computer Graphics. *IEEE Transactions on Visualization and Computer Graphics* 18, 7 (2012), 1170–1188. https://doi.org/10.1109/TVCG.2011.127

[20] Jeffrey Heer and Michael Bostock. 2010. Crowdsourcing graphical perception: using mechanical turk to assess visualization design *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 203–212. https://doi.org/10.1145/1753326.1753357

[21] Yosef Hochberg and Yoav Benjamini. 1990. More powerful procedures for multiple significance testing. *Statistics in medicine* 9, 7 (1990), 811–818.

[22] Matt-Heun Hong, Jessica K. Witt, and Danielle Albers Szafir. 2022. The Weighted Average Illusion: Biases in Perceived Mean Position in Scatterplots. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2022), 987–997. https://doi.org/10.1109/TVCG.2021.3114783

[23] H. Howard, R. McMaster, T. Slocum, and F. Kessler. 2008. Thematic cartography and geovisualization. (2008).

[24] A.M. MacEachren and D.R.F. Taylor. 2013. *Visualization in Modern Cartography.* Elsevier Science. https://books.google.com/books?id=3cP-BAAAQBAJ

[25] Alan M. Maceachren and Menno-Jan Kraak. 1997. Exploratory cartographic visualization: Advancing the agenda. *Computers & Geosciences* 23, 4 (1997), 335–343. https://doi.org/10.1016/S0098-3004(97)00018-6 Exploratory Cartograpic Visualisation.

[26] Dom Massaro and Norman Anderson. 1971. Judgmental model of the Ebbinghaus illusion. *Journal of experimental psychology* 89 (08 1971), 147–51. https://doi.org/10.1037/h0031158

[27] Hans-Joachim Meihoefer. 1973. The visual perception of the circle in thematic maps/experimental results. *Cartographica: The International Journal for Geographic Information and Geovisualization* 10, 1 (1973), 63–84. https://doi.org/10.3138/2771-5577-5417-369t

[28] Sebastian Mittelstädt, Andreas Stoffel, and Daniel Keim. 2014. Methods for Compensating Contrast Effects in Information Visualization. *Computer Graphics Forum* 33. https://doi.org/10.1111/cgf.12379

[29] Vanessa Peña-Araya, Emmanuel Pietriga, and Anastasia Bezerianos. 2019. A comparison of visualizations for identifying correlation over space and time. *IEEE transactions on visualization and computer graphics* 26, 1 (2019), 375–385.

[30] Aire Raidvee, Mai Toom, Kristiina Averin, and Jüri Allik. 2020. Perception of means, sums, and areas. *Attention, Perception, & Psychophysics* 82 (2020), 865–876.

[31] Robert Roth. 2011. *Interacting With Maps: The Science And Practice Of Cartographic Interaction.* Ph.D. Dissertation.

[32] Ariel Saffer, Laura Tateosian, Amanda C Saville, Yi-Peng Yang, and Jean B Ristaino. 2024. Reconstructing historic and modern potato late blight outbreaks using text analytics. *Scientific Reports* 14, 1 (2024), 2523.

[33] Hanan Samet, Yunheng Han, John Kastner, and Hong Wei. 2020. Using animation to visualize spatio-temporal varying COVID-19 data. In *Proceedings of the ACM SIGSPATIAL Workshop on Modeling and Understanding the Spread of COVID-19 (COVID-19)*. Association for Computing Machinery, New York, NY, USA, 53–62. https://doi.org/10.1145/3423459.3430761

[34] Jagan Sankaranarayanan, Hanan Samet, Benjamin E. Teitler, Michael D. Lieberman, and Jon Sperling. 2009. TwitterStand: News in Tweets *(GIS '09)*. Association for Computing Machinery, New York, NY, USA, 42–51. https://doi.org/10.1145/1653771.1653781

[35] Terry A Slocum, Robert B McMaster, Fritz C Kessler, and Hugh H Howard. 2022. *Thematic cartography and geovisualization.* CRC Press.

[36] Zdeněk Stachoň, Čeněk Šašinka, Jiří Čeněk, Stephan Angsüsser, Petr Kubíček, Zbyněk Štěrba, and Martina Bilíková. 2018. Effect of Size, Shape and Map Background in Cartographic Visualization: Experimental Study on Czech and Chinese Populations. *ISPRS International Journal of Geo-Information* 7, 11 (2018). https://doi.org/10.3390/ijgi7110427

[37] Benjamin E. Teitler, Michael D. Lieberman, Daniele Panozzo, Jagan Sankaranarayanan, Hanan Samet, and Jon Sperling. 2008. NewsStand: A New View on News. In *Proceedings of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (Irvine, California) *(GIS '08)*. Association for Computing Machinery, New York, NY, USA, Article 18, 10 pages. https://doi.org/10.1145/1463434.1463458

[38] The New York Times. 2022. New York Times COVID-19 dashboard. Retrieved May 28, 2022 from https://www.nytimes.com/interactive/2021/world/covid-cases.html

[39] Cagatay Turkay, Aidan Slingsby, Helwig Hauser, Jo Wood, and Jason Dykes. 2014. Attribute Signatures: Dynamic Visual Summaries for Analyzing Multivariate Geographical Data. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 2033–2042. https://doi.org/10.1109/TVCG.2014.2346265

[40] Johns Hopkins University. 2022. Johns Hopkins COVID-19 dashboard. Retrieved May 28, 2022 from https://coronavirus.jhu.edu/

[41] Kamila Łucjan. 2016. Perception of the contents of animated maps. *Polish Cartographical Review* 48 (12 2016). https://doi.org/10.1515/pcr-2016-0015