Recognition and Recall of Geographic Data In Cartograms Paper ID: 104

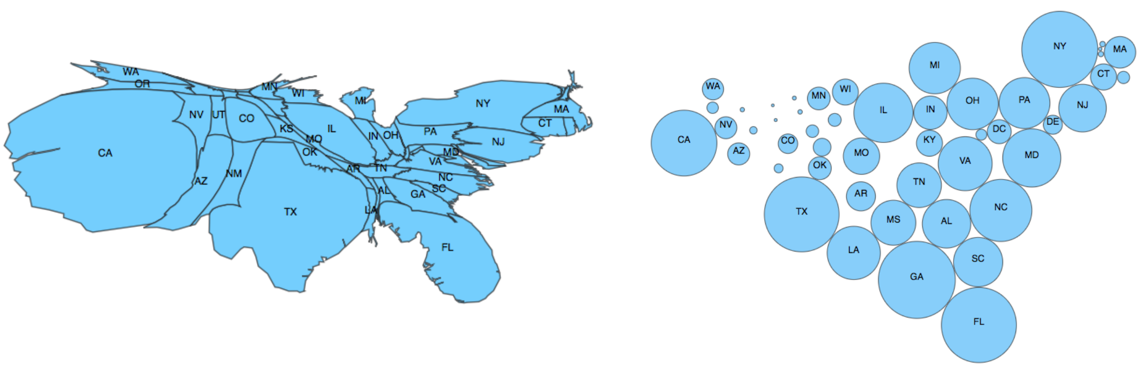


Figure 1: Which cartogram of the USA is more memorable: the contiguous cartogram on the left (showing Hispanic and Latino population) or the Dorling cartogram on the right (showing African-American population)?

ABSTRACT

In this paper we investigate the memorability of two types of cartograms, both in terms of recognition of the visualization and recall of the data. A cartogram, or a value-by-area map, is a representation of a map in which geographic regions are modified to reflect a given statistic, such as population or income. Of the many different types of cartograms, the contiguous and Dorling types are among the most popular and most effective. With this in mind, we evaluate the memorability of these two cartogram types with a human-subjects study, using task-based experimental data and cartogram visualization tasks based on Bertin's map reading levels. In particular, our results indicate that Dorling cartograms are associated with better recall of general patterns and trends. This, together with additional significant differences between the two most popular cartogram types, has implications for the design and use of cartograms, in the context of memorability.

KEYWORDS

geo-referenced visualization, cartograms, user study, memorability

ACM Reference Format:

. 2019. Recognition and Recall of Geographic Data In Cartograms Paper ID: 104. In *International Conference on Advanced Visual Interfaces, June 09–12, 2020, Italy.* ACM, New York, NY, USA, 9 pages.

1 INTRODUCTION

A popular saying goes: "I hear and I forget; I see and I remember." Indeed, one of the intuitive appeals of a good visualization is that it can draw the viewer in, encourage more observation and thought, and hopefully lead to insights that become memorable. Memorability, a basic cognitive concept, is important both due to implications for the design of visualizations that will be remembered and because it could help understand higher cognitive functions such as comprehension [8].

Cartograms, or value-by-area maps, simultaneously show (modified) geography and statistics making it possible to see patterns, trends, and correlations for geo-referenced data of different type, e.g., political, social, economic. Likely due to aesthetic appeal and the possibility to visualize data and put political and socioeconomic reality into perspective, cartograms are widely used in newspapers, magazines, textbooks, blogs, and presentations [1, 2, 23].

The effectiveness with which thematic maps and cartograms communicate spatial information has received considerable attention in the cartographic and visualization literature. McEachren suggests two groups of two criteria that can be used to evaluate the effectiveness of a thematic map: effectiveness for direct acquisition (with the map present) of specific and general information, and effectiveness for remembering (with the map absent) of specific and general information [22]. While there has been some earlier work about the effectiveness of cartograms for direct acquisition of data [18, 21, 25, 34], little is known about the memorability of different types of cartograms. Understanding what types of cartograms are more memorable, both in terms of encouraging the recall of the underlying data and in terms of recognizability, can help us understand how to create more memorable and effective data visualizations using cartograms.

Some early studies in the memorability literature involve thematic maps, however, different cartogram types have not been studied in the context of memorability. In this paper, we study the memorability of the two most popular and frequently used types of cartograms: contiguous and Dorling cartograms. We address several research questions including: Which of the two cartogram types is more effective for information recall? Do people who examine cartograms remember specifics about the data, or just high level patterns? Do novices and experts recall information differently? Does familiarity effect memorability?

With these questions in mind, we designed and performed a human-subjects experiment to measure the memorability of contiguous and Dorling cartograms (see e.g., in Fig 1), both in terms of recognition and recall. We used task-based experimental data and cartogram visualization tasks based on Bertin's map reading levels [5]. Our results indicate that Dorling cartograms are associated with better recall of general patterns and trends. This, together with additional significant differences between the two most popular cartogram types, has implications for the design and use of cartograms, in the context of memorability.

2 RELATED WORK

Cartograms: There are many methods to generate cartograms, which can be broadly categorized into four types: contiguous, non-contiguous, Dorling, and rectangular.

In contiguous cartograms the original geographic map is modified by deforming the boundaries to change areas. Among these cartograms, the most popular is the diffusion-based method proposed by Gastner and Newman [15]. Others of this type include CartoDraw by Keim et al. [19], constraint-based continuous cartograms by House and Kocmoud [17], and medial-axis-based cartograms by Keim et al. [20]. Dorling cartograms schematize regions using circles [13, 14]. Data values are realized by the sizes of the circles. In order to avoid overlaps, circles might need to be moved away from their original geographic locations. Contiguous and Dorling cartograms are among the most frequently used[27] and also among the most effective (based on empirical research that includes qualitative and quantitative experiments) [21, 25] and we focus on these two cartogram types in our study.

Cartograms and Perception: Challenges with area perception in cartograms affect some visualization tasks that involve magnitude estimation. The shape distortion inherent in cartograms makes it difficult to recognize the underlying geography [12]. Bertin [5] provides systematic guidelines to test visual encodings such as area, color, and texture. Cleveland and McGill [11] show that position judgments have higher accuracy than length judgment, which in turn has higher accuracy than area judgment. Stevens [33] shows that subjects perceive length with minimal bias, but underestimate differences in area. This finding is further supported by Cleveland et al. [10], and Heer and Bostock [16].

Dent [12] considers magnitude estimation, specifically with respect to cartograms, and highlights the tendency of humans to estimate lengths correctly, but underestimate areas and volumes in a non-linear fashion. He suggests effective communication strategies when the audience is not familiar with the underlying geography and statistics, e.g., providing an inset map, and labeling the statistical units on the cartogram. These studies indicate that while there are non-trivial issues with area perception in cartograms, these can be alleviated with the help of certain design decisions. Based on these suggestions we provide appropriate labels and inset maps with all the cartograms used in our study.

Memorability of Visualizations: Memorability tests usually involve two types of tasks: recognition and recall [22]. Borkin et al. [7] study what makes visualizations memorable and show that natural visualization (that are similar to scenes, objects, and people) and rounded features are more memorable. They do not specifically study the recall of the underlying data presented in the visualization. In a follow-up study, Borkin et al. [8] investigate how visualizations are recognized and recalled, considering which visualization elements are encoded and consecutively recalled. Their results show that visualizations that are memorable "at-a-glance" (after only 1 second of encoding) are often the same ones that are memorable after "prolonged exposure" (10 seconds of encoding). Titles and supporting text help recall the message of a visualization and even if pictograms do not necessarily help, they do not seem to hinder. Bateman et al. [4] compare embellished charts with plain ones and find that embellished charts are associated with high recall scores for the "value message" (a high-level message communicated through the chart). The effect of embellishments is further supported by the work of Borgo et al. [6] who show that visual embellishments can help participants remember the information depicted in visualization (both in terms of accuracy of recall and time required for recall).

Saket et al. [31] compare node-link visualizations to map-based visualizations focusing on information recall, using immediate and long-term recall conditions. When comparing the results for immediate and long-term recall conditions, the participants who recalled the data after two minutes (immediate) had significantly more accurate recall compared to four days later (long-term). Of particular relevance to our study, the authors conclude that the decay rates from short-term to long-term are the same, indicating that (the simpler to perform) immediate recall experiments might suffice. With these results in mind, we evaluate both recognition and recall using only one time duration (between encoding and testing).

Memorability of Thematic Maps and Cartograms: The effectiveness of choropleth (value-by-color) and isopleth (value-by-contour) maps is evaluated in [22], by a three step process: learning phase, direct acquisition of information with the map shown, and recall of general patterns and specific information. Thei results show that general patterns are better remembered in isopleth maps, while there is no significant difference in the recall of specific details.

Rittschof [28] compares the memorability of cartograms, choropleth maps, and other thematic maps and finds that cartograms perform worse than the other maps, although only one type of manually-created non-contiguous cartogram was used. Some observations and suggestions from a follow-up study [29] include: (i) Cartograms should be used only when learners have a long-term familiarity with the region depicted. (ii) When cartograms are used, the true scale of the depicted region should be emphasized, to prevent misunderstandings.

In summary, in spite of some early studies that involved cartograms, the memorability of different cartogram types has not been considered. Further, several of these studies did not include a proper learning phase. Both of these issues likely affect performance and hence the results. Finally, none of the earlier studies covers the spectrum of possible cartogram tasks [26].

3 MEMORABILITY STUDY DESIGN

Our experiment aims to measure the effectiveness of two major cartogram types (contiguous and Dorling) in terms of cartogram memorability (both recall of underlying data and recognizability).

3.1 Pilot Study

Prior to the main experiment, we conducted an extensive pilot study with 7 subjects to examine experimental parameters such as task difficulty, completion time, and balance between the experimental phases. The participants in the pilot study were graduate students working on data visualization and the experimenter was present during the pilot study to answer questions and collect immediate feedback.

Based on task completion times and accuracy data, we expanded the training phase to include more training tasks for each cartogram type and to ensure that the participants become sufficiently familiar with the types of questions and the experimental interface. The pilot study also helped us determine the time allocated for the recognition phase (see section 3.2). From the initial 5 seconds to "recognize" a map, we gradually increased the time to 12 seconds, in response to comments from the participants that 5 and even 10 seconds was too short to "see the entire map." The pilot study also helped us select a collection of tasks that covers the cartogram task taxonomy and are of comparable difficulty.

3.2 Experiment Overview

We designed and implemented a simple web-based application that guided the participants through the experiment, provided task instructions, and collected data about time and accuracy; see Fig. 2.

At the beginning of the experiment, we briefly introduce cartograms and the idea of encoding value by area, using one Dorling and one contiguous example. We also describe the purpose of the study and mention the expected duration of 15-20 minutes. Next we collect demographic and background information from the participants. We also ask participants if they are interested in data visualization, and whether they like maps and geography in general. The main experiment consists of following four phases.

Phase 1: Training. In the first phase, we show several cartograms to the participants. Each cartogram has a title, label, legend (inset map with labeled US regions), and explanatory text. For each cartogram shown we ask the participants to answer several (multiple choice) questions. After each answer we show the correct answer so that the participants can confirm the correctness, or examine the cartogram and the available answers, if incorrect.

Phase 2: Encoding. In the second phase, a new set of cartograms is shown to the participants. We ask them the same types of questions as in the training phase and request that the participants answer the questions as accurately as possible. We record accuracy and time for each question in this phase. We use this phase to *encode* the data shown in the cartogram in the memory of the participants.

Phase 3: Recognition. In the third phase, we test whether the participants remember the cartograms they encountered in the training or encoding phases. Since we are not interested in short term memory (10 seconds or less), we begin this phase with a distraction. Stevanov et al. [32] advocate using motion illusions to clear short term visual memory and this approach has been used when studying recall of relational data [31]. We show eight visual illusions for two minutes, and explain as follows: *"Before proceeding to the next stage, we need to take a short break. We will show you several optical illusions – please take a moment or two to look at them while we set up the next stage, which is available when the NEXT button is activated."* After the break the participants proceed to the recognition phase. The participants next see several cartograms (one at a time) from a subset of the cartograms shown in phase 1 and phase 2, and an equal number of cartograms that have not been seen in the earlier phases. We ask the participants to tell us whether the current cartogram is one that they have encountered before, or not. The participants have 12 seconds (determined based on the pilot study) to reply to each question and a countdown clock is shown.

Phase 4: Recall. In the recall phase, without showing the cartograms, questions about the data from the cartograms are asked.

3.3 Datasets

It has been shown that cartograms are most effective when the viewers are familiar with the underlying map [29]. Since the majority of our participants are from the USA we used cartograms of the USA.

Different statistical datasets was used to generate the cartograms for our study. These datasets were selected from a much larger pool of possible datasets, so that various geographic patterns are shown in both cartogram types. Specifically, we ensured that many of our examples differed from the typical cartograms of the USA (e.g., population, GDP) which expand the two coasts at the expense of inland states. For all the cartograms we use the contiguous 48 states and the District of Columbia.

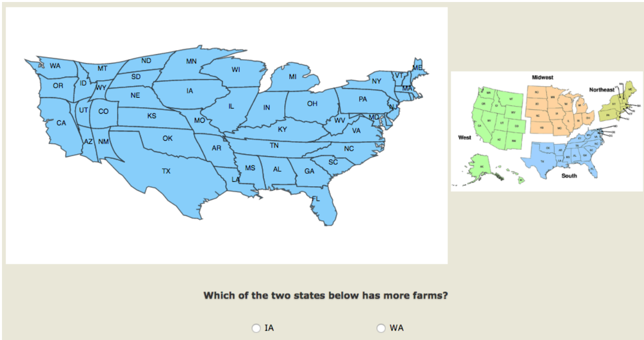
In the training phase we used four cartograms: a contiguous cartogram of the number of farms, a Dorling cartogram of the number of accidental deaths, a contiguous cartogram of the number of Starbucks stores, and a Dorling cartogram of Native American population. For each cartogram we ask four questions (task types are discussed in the next section). Each participant answers 4 cartograms $\times$ 4 questions = 16 questions.

In the encoding phase we also ask 16 questions about a new set of 4 cartograms that show: cattle inventory, African-American population, Hispanic and Latino population, and Starbucks per capita (number of Starbucks stores per 100,000 residents).

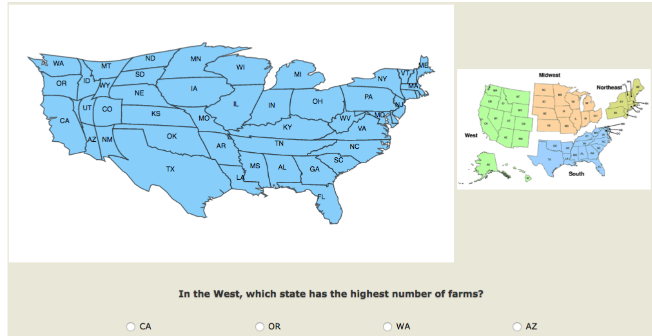
In total, eight different cartograms are seen in phase 1 and phase 2. Four of these, in addition to four "filler" cartograms (cartograms that the participants did not encounter in earlier phases), are used in the recognition phase. In the recall phase, the exact same 16 questions from the encoding phase are repeated, but without showing the cartograms.

3.4 Tasks

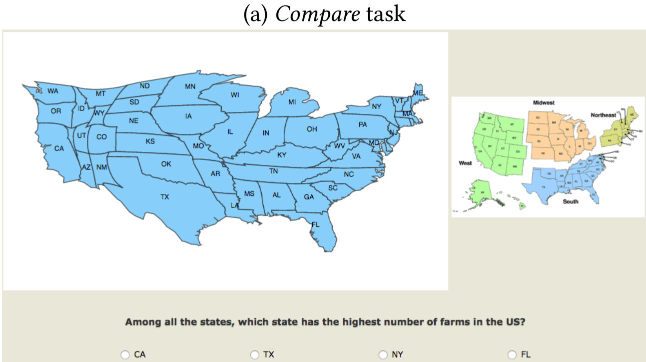
We selected tasks based on several factors. Bertin [5] defines three levels of map reading: elementary, intermediate and overall, which deal with a single data element, multiple elements and all elements of the map, respectively. According to MacEachren [22], maps can provide specific and general information. These are not discrete



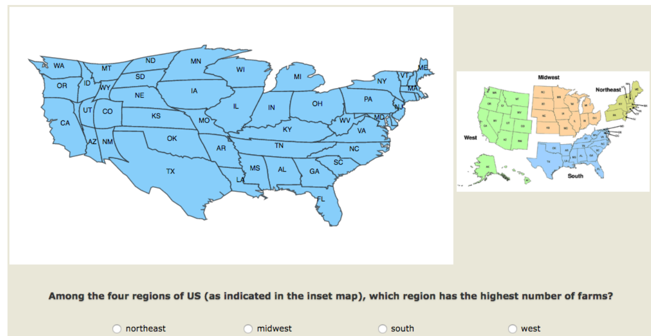
(a) *Compare* task



(b) *Filter* task



(c) *Find top-k* task



(d) *Summarize* task

Figure 2: Sample tasks on a cartogram showing the number of farms in the US in 2012. An inset US geographic map is also shown.

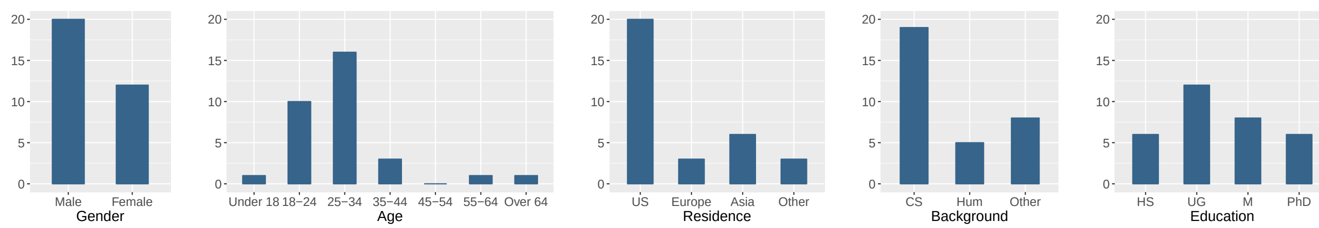


Figure 3: Demographic statistics of the participants.

categories but are two ends of a continuum along which all information contained in a map fall. With this in mind, we chose four tasks from the cartogram task taxonomy [26] that cover Bertin's three levels of map reading, and MacEachren's classification of specific details and general pattern tasks; see Fig. 2.

Compare: The *compare* task is a commonly used task in objective-based taxonomies [24, 30, 35]. This task requires an elementary level of reading, and provides very specific information.

Example Task: Given a population cartogram of the USA, compare two states by size.

Filter: The *filter* task asks to find data cases satisfying some criteria about a given attribute [3, 25]. That is, the viewer can filter out examples that fail the criteria. This task requires an intermediate level of reading, and also provides specific information.

Example Task: In the West, which of the following states has higher cattle inventory than Colorado?

Find top-*k*: This is another commonly used task in visualization [3, 35]. Here the goal is to find k entries with the maximum (or minimum) values of a given attribute. This task requires an overall level of reading, but provides specific information.

Example Task: Given a population cartogram, find out which state has the highest/lowest population.

Summarize: Cartograms are most often used to convey a "big picture," such as US population is denser along the coasts, or Northern European countries have higher GDP than Southern European ones. The *summarize* task is one that asks the viewer to find high level patterns and trends and is associated with overviews of data and global distribution of data on the map [9, 25]. This task requires an overall level of reading, and helps in understanding the general information contained in the map.

Example Task: Among the four regions of US (West, South, Midwest, and Northeast), which region has the highest number of farms?

3.5 Hypotheses

Our hypotheses are motivated by findings reported in the cartogram literature, by earlier cartogram evaluations, and by the initial results from the pilot study.

H1: Dorling cartograms will facilitate recall of general patterns and trends (i.e., *summarize* tasks). This hypothesis is based on the fact that in finding patterns and summarizing data, people perform best using Dorling cartograms [25].

H2: There will be no significant difference between cartogram types in the recall of specific details (i.e., for the other three tasks: *compare*, *filter*, *find top-k*). This hypothesis is based on earlier studies (such as [22]) showing no significant difference in recalling specific information in different thematic maps.

H3: Contiguous cartograms will facilitate recognition. This hypothesis is based on the observation that contiguous cartograms often create peculiar overall shapes (e.g., the barbell-shape for population cartograms of the USA, or the almond-shape for farm cartograms in the USA) which should be easier to recognize.

3.6 Participants

We shared a link to the study via social media and a total of 51 individuals participated. From this set we removed 19 participants who did not complete the entire study. Out of the 32 participants who completed the study, 20 were male and 12 female; 1 was under 18 years of age, 10 between 18–24 years of age, 16 between 25–34 years of age, 3 between 35–44 years of age, 1 between 55–64 years of age, and 1 over 64 years of age; 6 listed high school, 12 listed undergraduate, 8 listed Masters and 6 listed PhD as their highest completed educational level; see also Fig. 3. Familiarity with cartograms also differed: 14 participants were familiar with Dorling and 11 were familiar with contiguous cartograms.

4 RESULTS AND DATA ANALYSIS

All data from the experiment, R scripts to generate the plots for error rates and completion times for each task (by all individual participants and by different demographic groups), as well screenshots of a sample run-through of the entire experiment are provided in the supplementary documents that accompany this submission. We briefly discuss several of the more interesting findings below.

For each recall task we record error rates and average completion times for each participants. For recognition tasks we record the *miss rates* (false reject/(correct recognition + false reject)) and *false alarm rates* (false recognition/(correct reject + false recognition)) for each participant. Similar performance metrics for recognition were used by Borkin et al. [7]. These values are shown in Tables 1 and 2. Since the data does not have normal distribution we look for statistically significant differences between groups, using the Wilcoxon signed-rank test [36] (a non-parametric test that does not assume normality). The within-subject independent variables are the two cartogram types. The two dependent measures are the average completion times and error percentages by the participants.

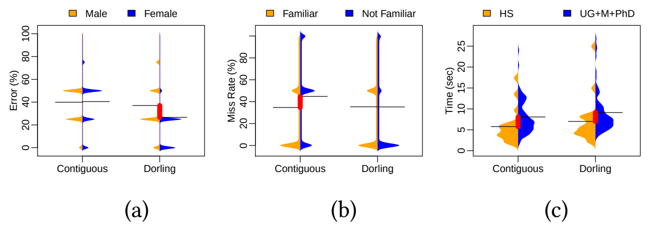


Figure 4: Demographic effects on recognition and recall: (a) gender bias in recognition error, (b) familiarity effect on recognition miss rate, and (c) effect of education level on recall time. The red vertical bars indicate statistically significant difference (based on Wilcoxon signed-rank tests).

In all cases below, the null hypothesis is that the cartogram type does not affect completion times and error rates. When the probability of the null hypothesis (p-value) is less than 0.05, we reject the null hypothesis.

4.1 Support for Hypotheses

H1 is supported by our results. For *summarize* tasks, Dorling cartograms are associated with significantly smaller errors than contiguous cartograms; see the last row in Table 1. This implies that participants can recall general patterns and trends shown in Dorling cartogram more accurately than in contiguous cartograms. There are several possible explanations, including that since the size comparisons with circles are difficult, the participants required more time working with the cartogram and are more likely to remember it. Note the statistically significant differences in time in Table 1.

H2 is also supported by the results: there is no significant difference in accuracy for tasks that require recalling specific details; see Table 1. Note the significant increase in completion time for Dorling cartograms in two of the three specific details tasks (mentioned above). However, the "more difficult, therefore requires more time" explanation is likely not at work here. Since Dorling cartograms have been shown to be efficient (in terms of time and error) for such tasks before [25], a plausible explanation is that the participants enjoyed exploring these cartograms and spent more time. Note that the instructions to the study did not ask the participants to work as fast as possible, but only as accurately as possible.

H3 is only weakly supported. We expected that the more peculiar overall shapes in contiguous cartograms would make them more recognizable. However, for the cartograms that the participants saw during the study, there is no significant difference in recognition between the two cartogram types. On the other hand, for the cartograms that were not shown during the study, the false alarm rate for Dorling cartograms is significantly lower (better); see Table 2.

4.2 Further Analysis

Here we consider more detailed analysis of our data, taking into consideration additional factors such as age, gender, education. Several interesting observations follow.

Gender: In our study, female participants are more accurate in recognition tasks involving Dorling cartograms; see Fig. 4(a). For contiguous cartograms, both male and female participants perform equally, whereas the difference in accuracy between the two

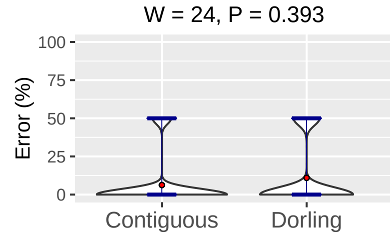
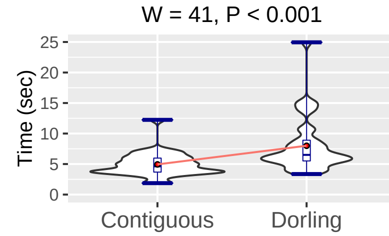
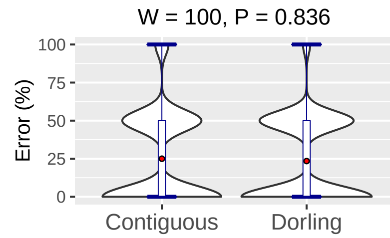
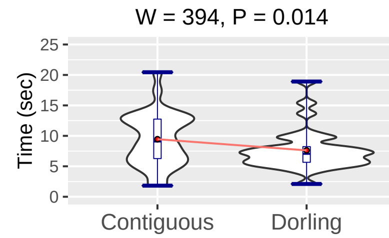
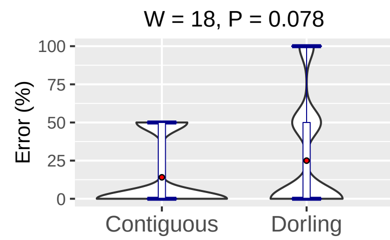
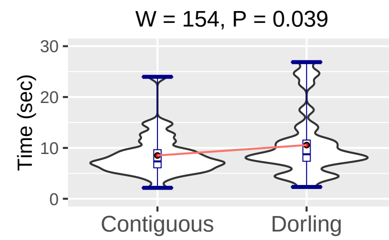
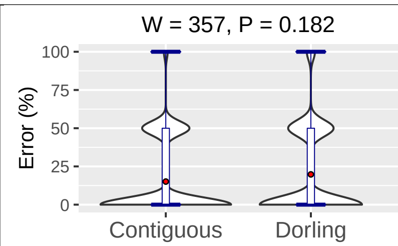
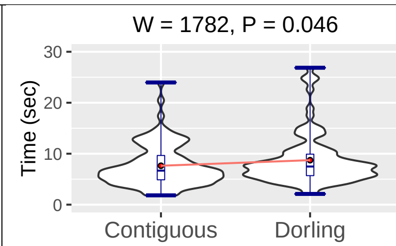
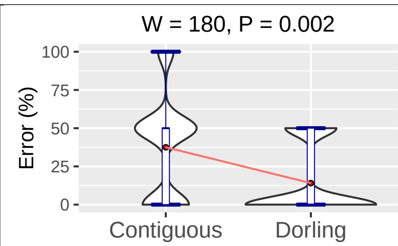
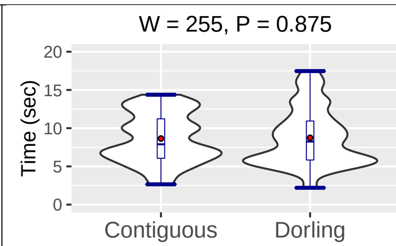
	Sample Question	Error %	Time (s)
Compare	Given a cartogram for the number of Starbucks stores in the US in 2009, which of the two states given has more Starbucks stores?		
Filter	Given a cartogram for the number of Starbucks stores in the US in 2009, which of the four given states in the South of US has the greatest number of Starbucks stores?		
Find-Top-k	Given a cartogram for the number of Starbucks stores in the US in 2009, which state has the greatest number of Starbucks stores among all the states in the US?		
Specific-Information	The previous three rows lists the three specific details tasks: *Compare*, *Filter*, and *Find-Top-k*. Here we show the average error rate and time for the three tasks.		
Summarize	Given a cartogram for the number of Starbucks stores in the US in 2009, which of the four regions in the US (as indicated in an inset map) has the highest number of Starbucks?		

Table 1: For each task, the last two columns show average error percentage for recall accuracy, and average completion time in seconds for contiguous and Dorling cartograms, along with the W and p values from a Wilcoxon signed-rank test. The violin plots show the distributions of error percentages or completion times, while the bottom and top of the boxes and the middle band represent the first quartiles, the third quartiles and the median values. The red circles represent the mean values, and each red line segment indicates that the difference between the two distributions for contiguous and Dorling cartograms is statistically significant. The last two rows summarize the overall error rates and completion times for all the specific details and general pattern tasks, respectively. Note that the error rates for the participants can have only some discrete values, such as 0%, 50%, 100% etc.

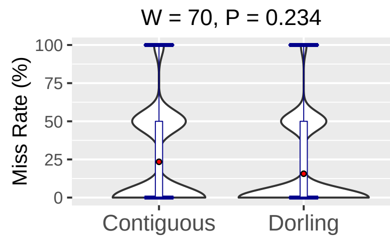
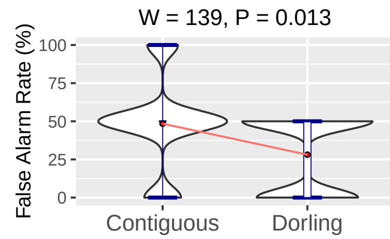
Question	Miss Rate %	False Alarm Rate %
Given a sample cartogram, have you seen this cartogram prior to this point during the study?		

Table 2: The last two columns show average miss rate and false alarm rate in percentages for the recognitions of contiguous and Dorling cartograms, along with the W and p values from a Wilcoxon signed-rank test. The violin plots show the distributions of error (miss or false alarm) percentages, while the bottom and top of the boxes and the middle band represent the first quartiles, the third quartiles and the median values. The red circles represent the mean values, and each red line segment indicates that the difference between the two distributions for contiguous and Dorling cartograms is statistically significant. The error rates for the participants can have only some discrete values, such as 0%, 50%, 100% etc.

groups for Dorling cartograms is statistically significant (using the Wilcoxon signed-rank test). Note that a similar gender bias in the performance on cartograms has recently been observed in [25].

Familiarity: We consider separately data for participants who were familiar with one cartogram type but not the other, and vice versa. While familiarity does not impact recognition in Dorling cartograms, it does seem to impact recognition for contiguous cartograms; see Fig. 4(b). In particular, the miss rate is the same for for Dorling cartograms, but significantly worse for contiguous cartograms. This suggests another possible advantage of the Dorling cartogram, which seems easier to grasp, regardless of familiarity.

Education: Both for contiguous and Dorling cartograms, high school graduates performed recall tasks faster than participants with higher degrees (Fig. 4(c)). A possible explanation is that high school graduates in our experiment are often college students and thus likely younger and quicker.

Country of residence: We considered the possibility that US residents might perform better since we only used US cartograms. We did not find significant difference in performance, however, suggesting that for both recall and recognition, the particular cartogram type might be more influential than the underlying geographic map.

Time and accuracy correlation: Finally, we explicitly tried to find whether taking longer time correlates with higher accuracy. Furthermore, we also verified whether taking more time in the encoding phase helped the participants in the recall and recognition accuracy. We therefore tested the following three pairs of parameters, also see Fig. 5: (a) total time by all participants and total number of correct answer for each recall and recognition task, (b) total time in the encoding phase and error rate in the recall phase by each participant, and (c) total time in the encoding phase and error rate in the recognition phase by each participant. In all three cases, however, there is no clear correlation between the parameters, which is confirmed by the low correlation coefficients: 0.06, 0.26, and 0.04.

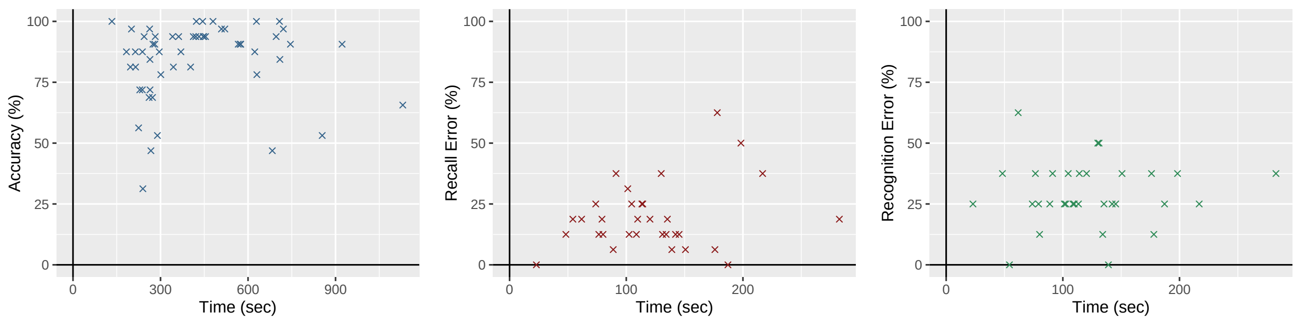
Figs. 6 and 7 show the four cartograms used in the recall phase and the eight cartograms used in the recognition phase along with the percentage of accurate responses by participants on each cartogram; in particular, we can now see that the Dorling cartogram in Fig. 1 was more memorable.

5 LIMITATIONS

We used only maps of the US in our study. We did not consider the relationship between the original geographic area and the statistical data shown: since the number of examples was limited, so was the variation in this relationship. We only considered two types of cartograms, while there are many other options. Further, we used only one representative from each of the two types of cartograms, whereas there are many cartogram variants within these types (e.g., over a dozen different contiguous cartograms).

There are other limitations associated with recognition and recall studies in general and with our study in particular. We attempted to control several variables that typically impact memorability studies, such as the data used for a cartogram type. We used a limited number of datasets to account for the limited time a person would spend with the cartograms during the study. These were selected from a much larger pool of possible datasets, so that various geographic patterns are shown in both cartogram types. Specifically, we ensured that many of our examples differed from the typical cartograms of the USA (e.g., population, GDP) which expand the two coasts at the expense of inland states. We also ran preliminary experiments to eliminate biases associated with particular datasets and to ensure that the tasks were neither too difficult, nor too easy. There remains a possibility, however, that the particular datasets could have confounded some of the results.

Each participant was asked the same set of questions, but we randomized the order of the questions for each participant. We controlled some parameters, such as the states selected for each question type (e.g., Which state has higher cattle inventory – CA or TX?). Our participants were from different parts of the world, but some of them might have used their own background knowledge to answer some of the questions. Our analysis of the results showed no significant impact of country of residence, but participants who



(a) $r = 0.06$ (b) $r = 0.26$ (c) $r = 0.04$

Figure 5: The relationships between (a) total time by all participants and total number of correct answer for each recall and recognition task, (b) total time in the encoding phase and error rate in the recall phase by each participant, and (c) total time in the encoding phase and error rate in the recognition phase by each participant. The correlation coefficients indicate no significant correlations in each pair.

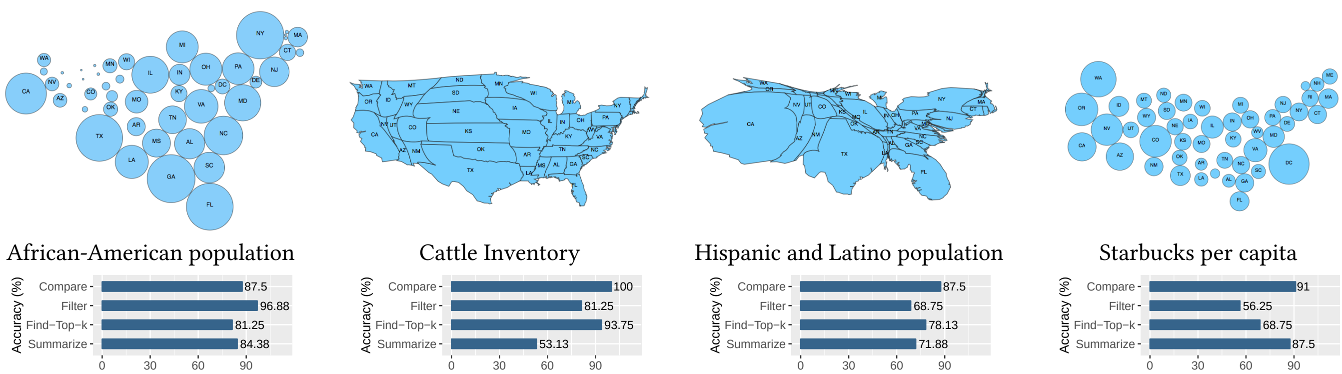


Figure 6: The four cartograms used in the recall phase along with the percentage of accurate answers by the participants in each task in bar-charts.

were more familiar with the US might have used their background knowledge.

Finally, there are limitations regarding the number of participants in our study (32) and the demographics of these participants (younger than average). Possibly associated with the small sample size is the often small effect size. Although some of the effects we report on are high (e.g., .45 and .31 for time and error in summarization), several others are lower. We anticipated that participants in the study would use laptop or desktop computers. One participant, however, mentioned some difficulties when using a mobile phone.

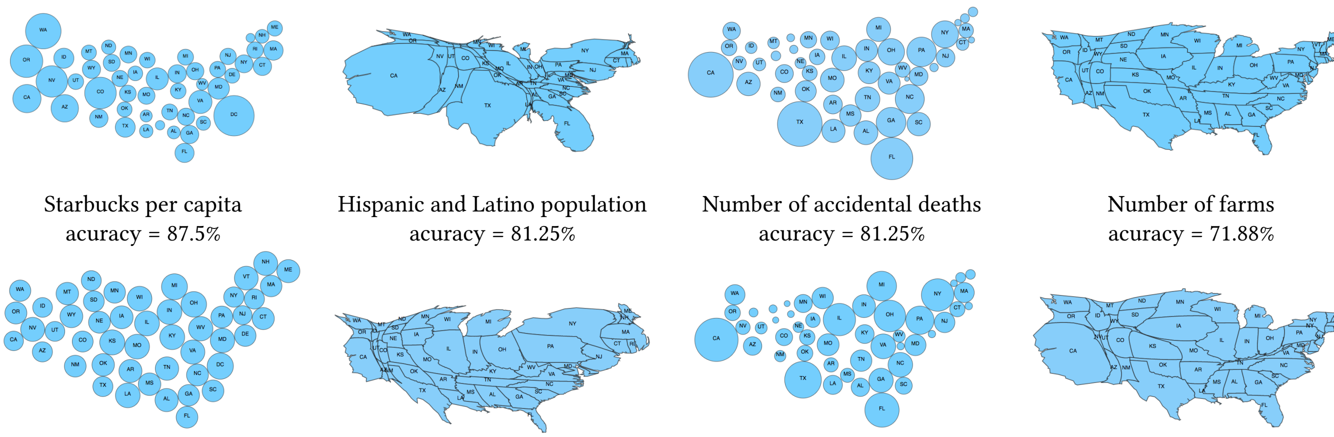
6 CONCLUSION AND FUTURE DIRECTIONS

The use of cartograms has increased through advances in geography, cartography, and visualization. In order to explore their effectiveness, we presented the first study of memorability of two of the most frequently used cartogram types. Our results indicate that Dorling cartograms are more suitable for showing big patterns and trends. Dorling cartograms also appear to be more recognizable. While some demographic factors (such as gender and education level) seem to impact the memorability of cartograms, other factors (such as age and familiarity with cartogram type) do not appear to affect memorability.

Natural directions for future work include studying the effects of a broader spectrum of tasks and a wider coverage of cartogram types. In addition to studying recognition and recall, it would be worthwhile to investigate possible implications of memorability, such as in engagement and ease of learning. Eventually, a solid understanding of what makes data memorable will help us design more impactful cartograms. Finally, broader investigations of geo-referenced visualizations, beyond time and error, are also needed.

REFERENCES

[1] Election News. http://www.nytimes.com/ref/washington/2006ELECTIONGUIDE.html, 2006.
[2] U.S. election results. http://graphics.latimes.com/2012-election-results-national-map/, 2012.
[3] R. Amar, J. Eagan, and J. Stasko. Low-level components of analytic activity in information visualization. *IEEE Symposium on Information Visualization*, pages 111–117, 2005.
[4] S. Bateman, R. L. Mandryk, C. Gutwin, A. Genest, D. McDine, and C. Brooks. Useful junk?: The effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2573–2582, New York, NY, USA, 2010. ACM.



Starbucks per capita acuracy = 87.5% | Hispanic and Latino population acuracy = 81.25% | Number of accidental deaths acuracy = 81.25% | Number of farms acuracy = 71.88%

Mcdonalnd's per capita acuracy = 96.88% | Population in 1930 acuracy = 71.88% | Number of Mcdonalnd's acuracy = 46.88% | Agriculture acuracy = 31.25%

Figure 7: The eight cartograms used in the recognition phase along with the percentage of accurate answers by the participants. The four in the top row had appeared in the study prior to recognition phase, and the four in the bottom row had not appeared beforehand.

[5] J. Bertin. *Semiology of Graphics: Diagrams, Networks, Maps*. The University of Wisconsin Press, Madison, 1983.
[6] R. Borgo, A. Abdul-Rahman, F. Mohamed, P. W. Grant, I. Reppa, L. Floridi, and M. Chen. An empirical study on using visual embellishments in visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2759–2768, 2012.
[7] M. Borkin, A. Vo, Z. Bylinskii, P. Isola, S. Sunkavalli, A. Oliva, H. Pfister, et al. What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2306–2315, 2013.
[8] M. A. Borkin, Z. Bylinskii, N. W. Kim, C. M. Bainbridge, C. S. Yeh, D. Borkin, H. Pfister, and A. Oliva. Beyond memorability: Visualization recognition and recall. *IEEE transactions on visualization and computer graphics*, 22(1):519–528, 2016.
[9] M. Brehmer and T. Munzner. A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2376–2385, 2013.
[10] W. S. Cleveland, C. S. Harris, and R. McGill. Judgments of circle sizes on statistical maps. *Journal of the American Statistical Association*, 77(379):541–547, 1982.
[11] W. S. Cleveland and R. McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American statistical association*, 79(387):531–554, 1984.
[12] B. D. Dent. Communication aspects of value-by-area cartograms. *The American Cartographer*, 2(2):154–168, 1975.
[13] D. Dorling. *The visualization of spatial structure*. PhD thesis, Department of Geography, University of Newcastle-upon-Tyne, U.K., 1991.
[14] D. Dorling. *Area Cartograms: Their Use and Creation*, volume 59 of *Concepts and Techniques in Modern Geography*. University of East Anglia, 1996.
[15] M. Gastner and M. Newman. Diffusion-based method for producing density-equalizing maps. In *The National Academy of Sciences of the USA*, volume 101, pages 7499–7504, 2004.
[16] J. Heer and M. Bostock. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *SIGCHI Conference on Human Factors in Computing Systems*, pages 203–212. ACM, 2010.
[17] D. H. House and C. J. Kocmoud. Continuous cartogram construction. In *IEEE Visualization*, pages 197–204, 1998.
[18] S. Kaspar, S. Fabrikant, and P. Freckmann. Empirical study of cartograms. In *25th international cartographic conference*, volume 3, 2011.
[19] D. Keim, S. North, C. Panse, and J. Schneidewind. Visualizing geographic information: VisualPoints vs. CartoDraw. *Information Visualization*, 2(1):58–67, 2003.
[20] D. A. Keim, C. Panse, and S. C. North. Medial-axis-based cartograms. *IEEE Computer Graphics and Applications*, 25(3):60–68, 2005.
[21] M. R. D. Krauss. *The relative effectiveness of the noncontiguous cartogram*. PhD thesis, Virginia Polytechnic Institute, 1989.
[22] A. M. MacEachren. The role of complexity and symbolization method in thematic map effectiveness. *Annals of the Association of American Geographers*, 72(4):495–513, 1982.
[23] A. Miller. TED talk: The news about the news. https://www.ted.com/talks/alisa_miller_shares_the_news_about_the_news, 2008.
[24] A. M.Maceachren, M. Wachowicz, R. Edsall, and D. Haug. Constructing knowledge from multivariate spatiotemporal data: integrating geographical visualization with knowledge discovery in database methods. *Intl. J. of Geo. Info. Science*, pages 311–334, 1999.
[25] S. Nusrat, M. J. Alam, and S. Kobourov. Evaluating cartogram effectiveness. *IEEE Transactions on Visualization and Computer Graphics, DOI:10.1109/TVCG.2016.2642109*, 2016.
[26] S. Nusrat and S. Kobourov. Task taxonomy for cartograms. In *17th IEEE Eurographics Conference on Visualization (EuroVis–short papers), also available in ArXiv: https://arxiv.org/abs/1502.07792*, 2015.
[27] S. Nusrat and S. Kobourov. The state of the art in cartograms. In *Computer Graphics Forum*, volume 35, pages 619–642. Wiley Online Library, 2016.
[28] K. A. Rittschof. *Comparing thematic map types and region familiarity on recall and inferencing*. PhD thesis, Arizona State University, 1994.
[29] K. A. Rittschof, W. A. Stock, R. W. Kulhavy, M. P. Verdi, and J. T. Johnson. Learning from cartograms: The effects of region familiarity. *Journal of Geography*, 95(2):50–58, 1996.
[30] R. E. Roth. An empirically-derived taxonomy of interaction primitives for interactive cartography and geovisualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2356–2365, 2013.
[31] B. Saket, C. Scheidegger, S. G. Kobourov, and K. Börner. Map-based visualizations increase recall accuracy of data. In *Computer Graphics Forum*, volume 34, pages 441–450. Wiley Online Library, 2015.
[32] J. Stevanov, B. Spehar, H. Ashida, and A. Kitaoka. Anomalous motion illusion contributes to visual preference. *Frontiers in psychology*, 3, 2012.
[33] S. S. Steven. On the psychophysical law. *Psychological Review*, 64(3), 1957.
[34] J. A. Ware. *Using animation to improve the communicative aspect of cartograms*. PhD thesis, Michigan State University, 1998.
[35] S. Wehrend. Appendix B: Taxonomy of visualization goals. In *Visual cues: Practical data visualization.*, pages 203–212. IEEE Computer Society Press, 1993.
[36] R. Woolson. Wilcoxon signed-rank test. *Wiley encyclopedia of clinical trials*, 2008.