Task-Based Effectiveness of Interactive Contiguous Area Cartograms

Ian K. Duncan, Shi Tingsheng, Simon T. Perrault, and Michael T. Gastner

Abstract—Cartograms are map-based data visualizations in which the area of each map region is proportional to an associated numeric data value (e.g., population or gross domestic product). A cartogram is called contiguous if it conforms to this area principle while also keeping neighboring regions connected. Because of their distorted appearance, contiguous cartograms have been criticized as difficult to read. Some authors have suggested that cartograms may be more legible if they are accompanied by interactive features (e.g., animations, linked brushing, or infotips). We conducted an experiment to evaluate this claim. Participants had to perform visual analysis tasks with interactive and noninteractive contiguous cartograms. The task types covered various aspects of cartogram readability, ranging from elementary lookup tasks to synoptic tasks (i.e., tasks in which participants had to summarize high-level differences between two cartograms). Elementary tasks were carried out equally well with and without interactivity. Synoptic tasks, by contrast, were more difficult without interactive features. With access to interactivity, however, most participants answered even synoptic questions correctly. In a subsequent survey, participants rated the interactive features as "easy to use" and "helpful." Our study suggests that interactivity has the potential to make contiguous cartograms accessible even for those readers who are unfamiliar with interactive computer graphics or do not have a prior affinity to working with maps. Among the interactive features, animations had the strongest positive effect, so we recommend them as a minimum of interactivity when contiguous cartograms are displayed on a computer screen.

Index Terms—Cartogram, geovisualization, interactive data exploration, quantitative evaluation

1 Introduction

The amount of geospatial information stored in digital databases and shared over the Internet is growing rapidly. To communicate information contained in geospatial data to a wide audience, we need effective visualization tools. Cartograms have emerged as an alternative to traditional thematic mapping techniques, such as choropleth maps, proportional symbol maps, and dot-density maps [1], [2]. The Worldmapper [3], [4] and Londonmapper projects [5], for example, make extensive use of cartograms. In a cartogram, each region is depicted by an area that is proportional to its corresponding value in the statistical data set used. In Fig. 1, we illustrate this idea using data for the gross domestic products (GDPs) of federal states in Germany. The left map in Fig. 1 is a conventional equal-area map, where Berlin (labeled as BE) occupies only 4.8 percent of the area of Saxony (SN). In the cartogram (the map on the right of Fig. 1), however, Berlin's area appears 12 percent larger than Saxony's so that the proportions of the states' GDPs can be correctly represented.

In this article, we concentrate on cartograms that are contiguous. That is, if map regions share a common border in a geographic space (i.e., on a conventional map) then they are also neighbors on the cartogram, and vice versa. The cartogram in Fig. 1, for example, is contiguous. Many other cartogram types exist (e.g., mosaic cartograms [6], Olson's noncontiguous [7], and Dorling's circular cartograms [8]). Previous surveys and experiments have compared the effectiveness of different cartogram types [9], [10]. Contiguous cartograms were consistently among the participants' preferred visualizations across a variety of contexts and tasks, even if other cartogram types might have been more suitable for specific applications or user groups.

Cartograms first gained popularity in the early 20th century [11] when they were hand-drawn and intended for inclusion in print media. Like most other forms of data visualization, currently cartograms are more frequently generated electronically and viewed on a computer screen than printed on paper. With the increasing presence of cartograms on the World Wide Web, news media have also adopted them to support their online content. Various newspapers have posted cartograms of election forecasts [12], [13], election results [14], [15], [16], and other demographic statistics [17], [18] on their websites. Despite the popularity of cartograms, they inevitably appear distorted when compared to conventional maps. For this reason, some authors have questioned whether cartograms are legible and comprehensible for the average reader [19], [20], [21], [22]. If a cartogram is displayed on a computer screen, it is possible to include interactive features to alleviate these concerns. Indeed, many of the cartograms posted by the news media include features such as a button or slider to switch between different cartograms and a conventional map [13], [17], [18], [23], [24], or a mouse-over effect to highlight regions [16], [25], or infotips [12], [13], [14], [15], [16], [18], [23], [24], [25]. However, it remains an open research question whether

- *Ian K. Duncan, Shi Tingsheng, and Michael T. Gastner are with the Yale-NUS College, Singapore 138527, Singapore. E-mail: {ian.duncan, tingsheng.shi}@u.yale-nus.edu.sg, mgastner@gmail.com.*
- *Simon T. Perrault is with the Singapore University of Technology and Design, Singapore 487372, Singapore. E-mail: simon_perrault@sutd.edu.sg.*

Manuscript received 17 Dec. 2019; revised 9 Oct. 2020; accepted 18 Nov. 2020. Date of publication 1 Dec. 2020; date of current version 28 Jan. 2021. (Corresponding author: Michael T. Gastner.) Recommended for acceptance by J. Dykes. Digital Object Identifier no. 10.1109/TVCG.2020.3041745

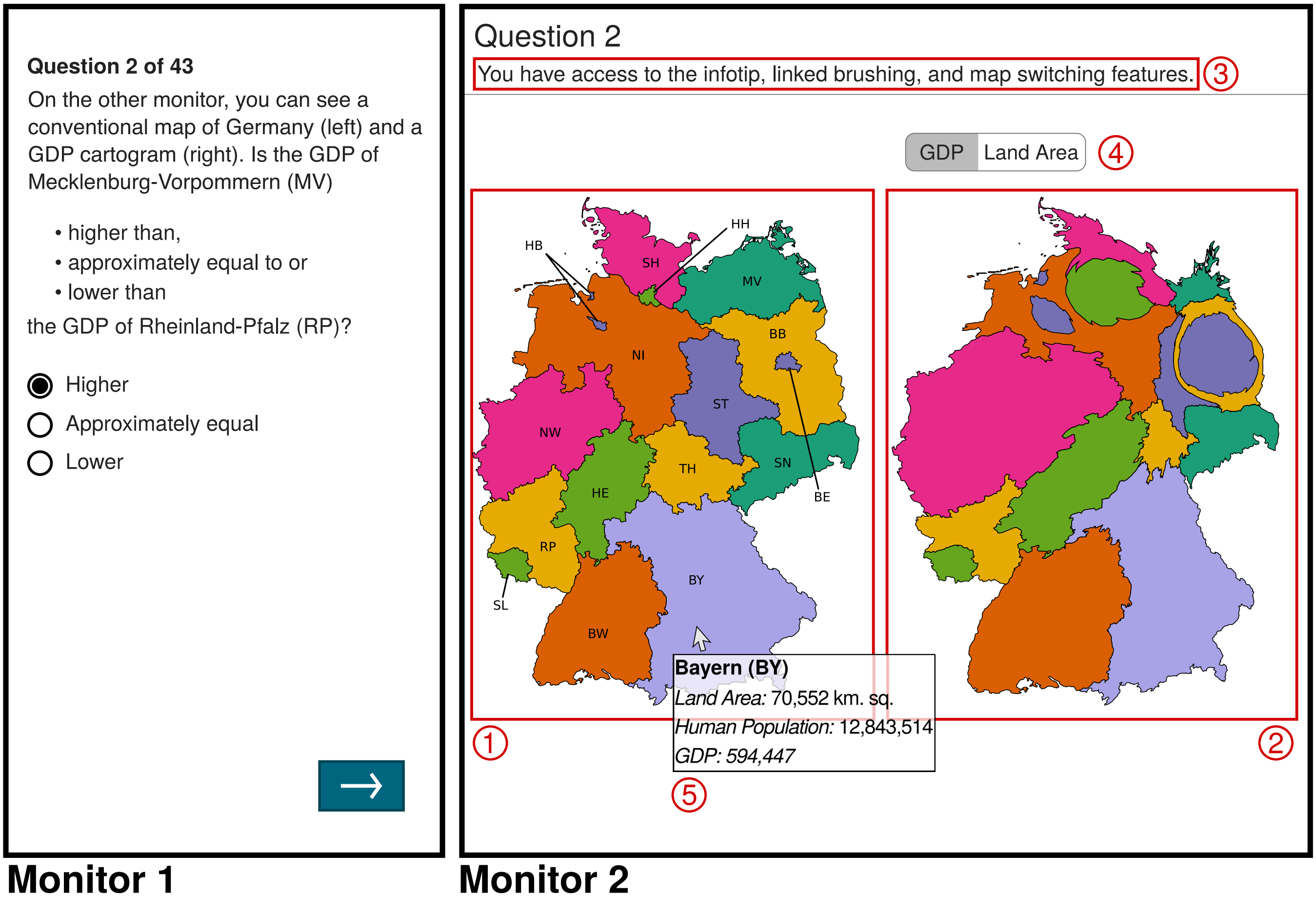

Fig. 1. Screenshot of the two-monitor setup used by participants to complete the map reading tasks. We rescaled the aspect ratios of the screens to fit the page width. Here we use a *Compare* task for Germany as an example (see Section 3.1 for the task description). On Monitor 1, participants read the current task and entered their answer. On Monitor 2, participants were presented with the cartogram viewing interface containing (1) a labeled conventional map and (2) a contiguous cartogram of the same country. Corresponding regions appear in the same color on the conventional map and the cartogram. Participants were informed of (3) which interactive features they could use to complete the current task. For some tasks, they could also switch the map displayed on the right to a cartogram based on another statistic using (4) the map switching selector. One of the buttons that could be selected was titled "Land Area." When this button was pressed, the map displayed on the right switched to an equal-area map, which can be considered as a cartogram where the statistic displayed is the land area of each administrative region. Hovering the mouse over a map region triggered (5) an infotip containing the region's name followed by the numeric data used to generate the displayed cartograms.

interactivity makes it easier to understand the information shown in a cartogram [26].

The purpose of this study is to evaluate whether cartograms can be used to communicate information more effectively if the viewer can interact with them using the following three interactive features proposed in the previous cartogram literature [27], [28], [29]:

- *Cartogram-switching animation*: The user can choose between different data sets by clicking a button on the screen (Fig. 1). As the user selects a new data set, the previously displayed cartogram morphs into a new cartogram. In our implementation, the transition from one cartogram to another is achieved by smoothly moving the polygon vertices from the start to end position in a one-second time interval.
- *Linked brushing:* The cartogram is displayed alongside a conventional map, as shown in Fig. 1. As the participant hovers the mouse over a region on the cartogram, the corresponding region is simultaneously highlighted on the conventional map and vice versa. In our experiment, we highlighted the region by increasing the brightness of its fill color, but other forms of highlighting (e.g., changing the border color or thickness) are also conceivable.
- *Infotip:* As the participant hovers the mouse over a map region, a pop-up appears at the location of the cursor (see Fig. 1). The text in the pop-up contains the region's name and the data (e.g., GDP) represented by the corresponding area in the cartogram.

We judge the effect of these interactive features on map-reading tasks, which range from elementary lookup tasks to high-level synoptic summaries of the cartograms, drawn from an objective-based task taxonomy [30]. Results from our experiment indicate that interactive features can help readers perform certain data analysis tasks more accurately than on cartograms without such features. While the inclusion of interactive features may not have an observable effect on simple tasks (with $< 20\%$ error rates), we did notice significant improvements for synoptic tasks, where cartogram-switching animations dramatically improved the accuracy of the participants' responses. Cartogram-switching animations may thus effectively remove concerns about the legibility and effectiveness of contiguous cartograms in the previous literature [21], [31].

At the end of the experiment, participants filled out an attitude survey. Overall, they expressed strongly positive opinions about all the interactive features tested in the experiment. In addition to the longstanding recommendations of presenting cartograms with a legend and alongside a conventional map [32], we therefore suggest that cartograms be presented with interactive features to improve reader comprehension.

2 Related Work

2.1 Previous Evaluations of Cartograms

The utility of cartograms has been debated for several decades. On one hand, Dent already reported in 1972 that students found cartograms intriguing and were eager to experiment with them [33]. On the other hand, he also criticized cartograms as bordering on the "surrealistic" [32] and primarily being used for their shock value [1]. In 1975, he conducted a series of experiments, which were among the earliest attempts to objectively evaluate cartogram effectiveness [32]. Participants were asked to estimate the population in the northeastern United States either from a cartogram or a proportional symbol map based on a conventional map projection. The accuracy of the participants' estimates was comparable for both map types. At the end of the test, participants characterized the cartogram as "innovative" and "interesting," but also pointed out that they found it difficult to read.

In a similar experiment in 1983, Griffin measured the accuracy and speed with which participants identified regions on a cartogram shown alongside a conventional map of Adelaide's electoral subdivisions [34]. One region was highlighted on one of the maps, and the task was to find the corresponding region on the other map. When the distortions were greater, the task became more challenging. Griffin hypothesized that participants struggled to establish the mental transformation between conventional maps and their cartograms. In a similar experiment, Kaspar *et al.* [35] noticed that the difficulty of the spatial inference task also depended on the shape of the polygon on a conventional map. Polygons with regular shapes (e.g., rectangles) appeared subjectively more distorted on the cartogram than polygons with irregular shapes.

At least for simple tasks, Kaspar *et al.* still concluded that cartograms could be as effective and efficient as traditional graduated circle maps. Most other cognitive experiments have echoed this result [9], [36], [37], whereas Gao *et al.* [38] argue that value-by-alpha maps [22] or proportional symbol maps are at least as effective, efficient, and popular with users as cartograms. Comparisons between different cartogram methods have shown that university students can cope well with the distortions inherent in contiguous cartograms [10]. Secondary school students, however, preferred the simpler geometries depicted in rectangular or mosaic cartograms [26], [39].

All the experiments mentioned in this section were done with static cartograms. Because many cartograms are currently shown on websites rather than in print, one may wonder whether interactivity (e.g., implemented with D3.js [40]) fundamentally alters the viewers' attitudes towards cartograms. As Goodchild noted about map displays in general, "it is unreasonable that a technology optimized under the narrow constraints of pen and paper would turn out to be indistinguishable from one optimized under the much broader constraints of digital technology" [41].

2.2 Interactivity in General Mapping Software

Interactive graphics have been included in general-purpose statistical software (e.g., for exploratory data analysis [42]) and geographic software (e.g., GIS and web-based mapping services [43]) for a long time. Zooming, for example, is a quintessential interactive feature in mapping software that encourages the viewer to focus on small portions of a map [44], [45]. According to Harrower and Sheesley [46], cartographers should consider incorporating zooming into interactive mapping because the feature allows for greater information density and is commonly understood by map readers. Apart from zooming, there are interactive features permitting comparisons between two or more maps that represent the same region, such as translucent overlays, a blending lens, and swiping [47]. Previous studies have exhibited examples of how these features are implemented on a computer display [48]. Several experiments concluded that interactivity has a positive impact on performance for map-related tasks [49], [50], [51], [52], although the type of interactivity and the type of tasks involved vary greatly between these experiments. Experiments by Keehner *et al.* [53] underline that the value of interactive features in spatial cognition tasks is not the interactivity in itself, but the easier access to informative views of the spatial objects.

It is surprising that the literature on cartograms, which are at the intersection of statistical and geographic visualization [54], has so far largely ignored the opportunities of interactivity. One reason for this neglect might be that the most discussed interactive features in geographic visualization are not directly useful when applied to cartograms. For example, while zooming is helpful in displaying the boundaries of the map in greater detail, the main purpose of a cartogram—as well as most other infographic maps [55]—is to show the relative importance of regions in a larger geographic context with a fixed spatial extent. For instance, zooming into the cartogram on the right of Fig. 1 would not help us relate Berlin's GDP to the total GDP of Germany. Other features are better suited for map comparison than zooming, but they often rely on the assumption that the two maps that are to be compared are based on the same underlying map projection. For contiguous cartograms, this assumption is invalid because different input statistics produce different cartogram projections. Therefore, many tools for map comparisons, such as translucent overlays, a blending lens, or swiping, would leave the viewer confused about the relation between the incongruent boundaries on the two cartogram layers. The only potentially promising map comparison technique that we found in the literature was map morphing, which does not require the same projection for both maps. Map morphing continuously interpolates between the projections of two incongruent maps with a short animation [56], [57], [58]. Reilly and Inkpen [59] reported that map morphing significantly improved accuracy for tasks that included the comparisons of region sizes on different projections. They noted that morphing also felt effective and enjoyable to the participants because it was easy to follow shifts in positions during the animation.

2.3 Interactivity in Cartograms

In addition to reviewing the general literature that covers the broad topic of interactivity in maps, we also thoroughly reviewed the more specific literature about cartograms, examined the available cartogram software, and inspected many cartograms posted on the World Wide Web. All interactive features that were mentioned or implemented fall into three categories: animations, linked brushing, and infotips. Therefore, we restricted our experiment to these three features for which we found a precedent in previous cartogram research.

The idea of combining animations with cartograms had been discussed since the mid-1980s [60], [61], [62], but it remained a theoretical prospect until Dorling produced an animated time series of UK election cartograms in 1992 [63]. Each constituency was represented by an arrow on a circular cartogram. The color of the arrow illustrated the vote composition, and the direction indicated the vote swing. The animation was played from a videotape so that viewers could not directly interact with the graphics. The first experiment in which participants could interact with a cartogram animation on a computer screen was conducted by Ware in 1998 [27]. The animation showed a morphing transition from a conventional map to a contiguous cartogram. Participants were assigned to different groups. Some groups could play, pause, or rewind the animation, whereas the control group could only view still images. All participants had to identify regions highlighted on a conventional map by clicking on the corresponding region on a cartogram. Participants with access to the animation needed more time to complete the tasks, but were more likely to give correct answers, especially when they were unfamiliar with the geography of the displayed country. Ware hypothesized that these participants spent some of their time rewinding the animation to ensure that their response was correct. Overall, she strongly advocated the use of animated presentations of cartograms, pointing out that user satisfaction is more important than a quick response time.

Instead of gradually deforming one map into another, it is also possible to compare maps by juxtaposition. Placing two maps alongside each other can be particularly effective in combination with linked brushing, whereby pointing at a position on one of the maps simultaneously highlights the corresponding position on the other map [47]. Inspired by implementations of scatterplot brushing in statistical software [64], [65] and its generalization to geographic data [66], [67], linked brushing has been proposed by several authors as a method of highlighting the correspondence between a conventional map and its cartogram [28], [68], [69]. Dykes commented that linked brushing "reduces the oft-quoted difficulties in relating cartogram symbols with the places that they represent." [70]. Linked cartogram brushing has been implemented by the GeoViz Toolkit [71] and Tableau [72]. However, as Tobler noted, the efficacy of linked cartogram brushing has so far not been evaluated [28].

Besides visually highlighting the region under the mouse pointer, Nusrat *et al.* [73] hypothesized that cartograms may also benefit from an infotip that reveals the exact value of the numeric data represented by the region's area. Infotips have been implemented in cartogram software such as MAPresso [29] and are included in many cartograms posted on the World Wide Web [12], [14], [18], [74], [75]. In the context of general map applications, experiments have found that users preferred infotips when tasked to retrieve background information about objects on a map even if other possibilities were available [76]. However, Brath and Banissi have remarked that an infotip may increase user response times without providing compensatory benefits in understanding and readability [77]. When unnecessary for completing a given data analysis task, an infotip may act as an intrusive visual stimulus that makes it more difficult for users to keep their attention on relevant aspects of a data visualization, potentially increasing the response time and decreasing the accuracy [78]. Despite these concerns, we are not aware of any previous experimental assessments of infotips in the context of cartograms.

2.4 Evaluating Interactivity With Task Taxonomies

While a general theory of interactive cartography is still in its infancy, there is broad agreement that we should assess the quality of a cartographic visualization with well-defined task taxonomies [79], [80], [81]. Roth distinguishes between objective-based taxonomies, which define tasks in terms of verbs that imply user intent (e.g., "identify," "compare"), and operator or operand-based taxonomies, which categorize tasks by the specific features or visualizations used [82]. From a card-sorting study with expert interactive map users, Roth identified five general objective primitives for interactive geovisualization: "identify," "compare," "rank," "associate," and "delineate" [83]. However, he admits that there was a large amount of variation in the participants' opinions, so it is unclear whether these objectives are optimal for user studies of all forms of geovisualization.

Here, we adopt the objective-based taxonomy that Nusrat and Kobourov developed specifically for cartograms [30]. Their suite of tasks is not explicitly designed to evaluate interactivity but can be easily modified to test the interactive features in our experiment. By adopting the objective-based perspective, we treat the interactive features as conditions under which the objectives can be reached, but participants are free to choose whether they use the implemented features during the experiment.

3 Experiment

3.1 Tasks

From the cartogram tasks proposed by Nusrat and Kobourov [30], we chose eight general categories that are relevant in the context of interactive cartograms. We list the task types together with example tasks in Table 1. In the typology of [84], the first seven tasks are "elementary." That is, they are basic statistical and map reading tasks that refer to individual subregions (e.g., states or provinces). By contrast, the last task in Table 1 ("Summarize") is "synoptic": it is a complex task that required participants to analyze and compare the whole set of regions on multiple cartograms that were jointly displayed on the same screen.

TABLE 1
Eight Task Types Used in Our Experiment With Example Tasks for Each Type

Cluster	
Task	For a given region, participants were required to select the region with the most similar area on the displayed cartogram from a set of four possible candidate regions. Here and in all other tasks, the task description was shown on monitor 1. The cartogram was displayed on monitor 2 (see Fig. 1).
Example	On monitor 2, you can see a conventional map of Brazil (left) and a cattle population cartogram (right). Out of the states listed below, which has a cattle population most similar to Mato Grosso do Sul (MS)?
Compare	
Task	For two given regions, R_1 and R_2, participants were required to decide if the area of R_1 was higher than, lower than, or approximately equal to the area of R_2 on the displayed cartogram.
Example	On monitor 2, you can see a conventional map of Germany (left) and a GDP cartogram (right). Is the GDP of Mecklenburg-Vorpommern (MV) higher, lower, or approximately equal to the GDP of Rheinland-Pfalz (RP)?
Detect Change	
Task	For a given region and two cartograms C_1 and C_2, participants were required to decide if its area in C_1 was higher than, lower than, or approximately equal to its area in C_2.
Example	On monitor 2, you can see a conventional map of Brazil (left), a cattle population cartogram, and a human population cartogram (right). Is the area of Amazonas (AM) in the cattle population cartogram higher, lower, or approximately equal to its area in the human population cartogram?
Filter	
Task	For a given region R, participants were required to select a map region with an area greater than that of R on the displayed cartogram.
Example	On monitor 2, you can see a conventional map of Germany (left) and a population cartogram (right). Out of the states listed below, which one(s) have a population higher than Baden-Württemberg (BW)? There may be more than one correct answer.
Find Adjacency	
Task	For a given region, which was highlighted on a cartogram displayed on monitor 1, participants were required to identify which regions are neighbors from a set of four candidates. Participants had to infer the answer from maps displayed on monitor 2.
Example	On monitor 2, you can see a conventional map of Brazil (left) and a population cartogram (right). Which state(s) are neighbors of the state highlighted in red in the population cartogram below (i.e., on monitor 1)? There may be more than one correct answer.
Find Top	
Task	Participants were required to identify the region with the largest area in the displayed cartogram.
Example	On monitor 2, you can see a conventional map of the USA (left) and a crop production cartogram (right). Which state has the highest crop production?
Recognize	
Task	For a given region, which was highlighted on a cartogram displayed on monitor 1, participants were required to identify its name based on maps shown on monitor 2.
Example	On monitor 2, you can see a conventional map of India (left) and a population cartogram (right). What is the name of the state highlighted in red in the population cartogram below (i.e., on monitor 1)?
Summarize	
Task	Regions were partitioned into three zones, which were colored yellow, purple, and pink. For two given cartograms, C_1 and C_2, participants were required to identify whether the yellow, purple, and pink zones increased in area, decreased in area, or showed no noticeable change in area between C_1 and C_2.
Example	On monitor 2, you can see a conventional map of Germany (left) and population cartograms for the years 1985 and 2015 (right). Three different regions are highlighted in yellow, purple, and pink. What can you say about the trend in population growth between 1985 and 2015?

3.2 Data Sets

For each task, we generated cartograms of five different parts of the world:

- Germany (all 16 Bundesländer),
- Brazil (all 26 states and the Federal District),
- the conterminous United States (48 states and Washington, D.C.),
- India (all 28 states and 7 union territories excluding Lakshadweep),

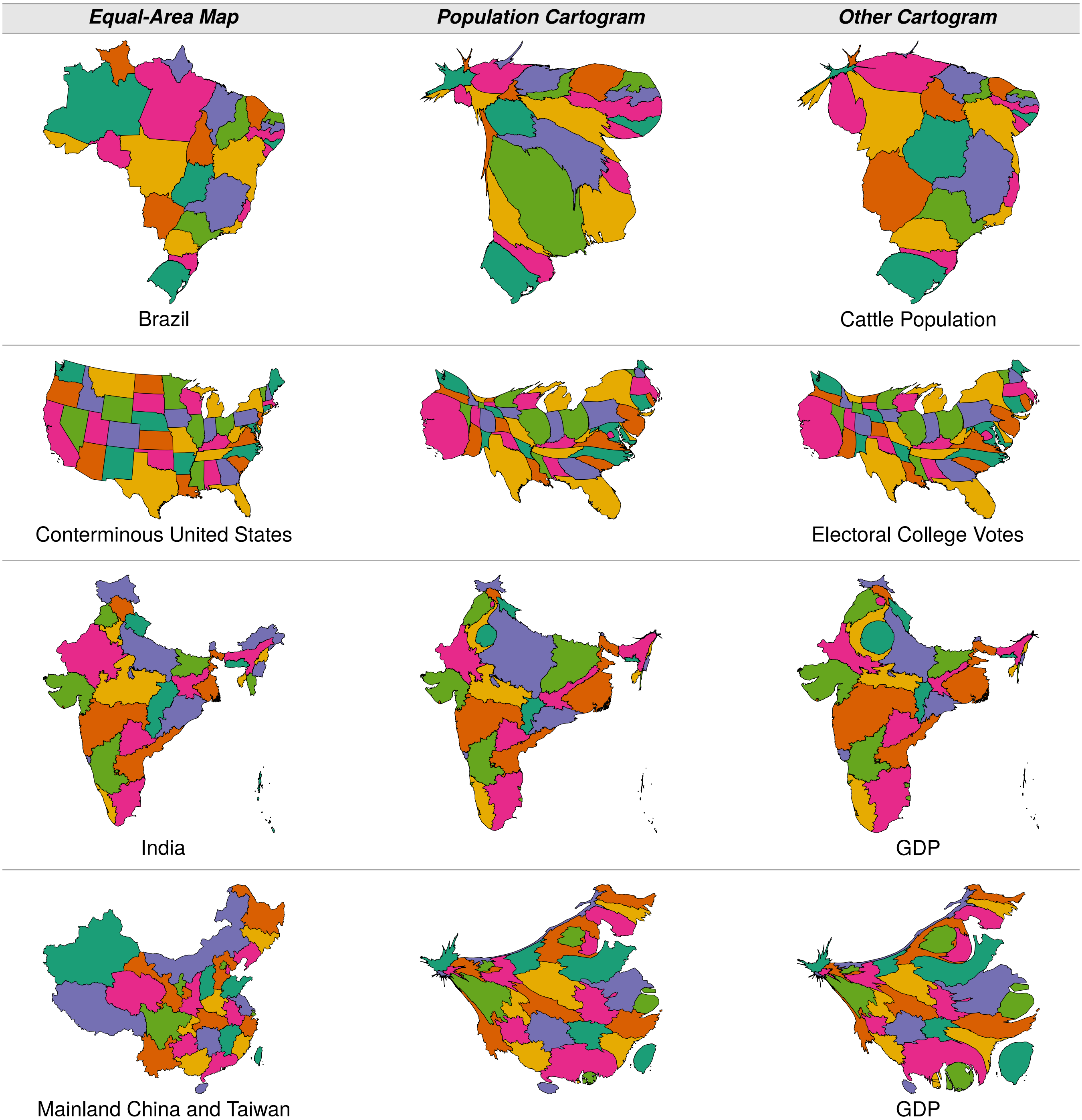

Fig. 2. Selection of cartograms shown during the experiment.

- mainland China, the Hong Kong and Macao Special Administrative Regions, and Taiwan (total of 34 administrative units).

Representative cartograms are shown in Fig. 1 (Germany) and Fig. 2 (Brazil, United States, India, and China). We opted for real countries with recognizable outer boundaries rather than artificial geometries or deliberately unidentifiable subsets of census areas [85] because we wanted the experimental tasks to resemble realistic use cases of cartograms. We controlled for possible previous knowledge bias by treating the regions as a blocking factor in the experimental setup (see Section 3.5).

The cartogram areas represented statistics from a variety of data sets. Apart from population and GDP data, we also visualized data that we expected few participants to be familiar with (e.g., agricultural production by state in the US, cattle production by state in Brazil). For *Summarize* tasks, we used actual or predicted population data from two different years. All the cartograms were produced with the fast flow-based method [86]. In each task, we presented a map made with a conventional equal-area projection alongside the cartogram (Fig. 1).

On the conventional map, we identified all regions with two-letter abbreviations—an example is shown in Fig. 1—to simplify the task for participants who were unfamiliar with the geography of the displayed country. Unlike the conventional maps, none of the cartograms showed the regions' abbreviations because those labels would have defeated the purpose of some task types (e.g., *Recognize*). As a visual hint to the participants, the colors of matching regions on the conventional maps and their corresponding cartograms were identical.

3.3 Participants

We recruited 55 participants. They were all students or staff of the National University of Singapore. Seventeen of the participants were female; 38 were male. The age range was from 18 to 52 years (mean 21.8, standard deviation 4.6). The participants received 10 SGD (around 7.05 USD) as compensation for their time. Because the experiment required the participants to distinguish between map regions highlighted with different colors, we used an Ishihara test to determine whether any participants were color blind. One participant exhibited potential color blindness. The performance of this participant did not differ significantly from that of the other participants. Hence, we included this participant's responses in our analysis.

Because almost all participants were university students, we acknowledge that our results may only apply to a younger, more educated group of people. Similar limitations apply to previous cartogram evaluations in the literature (e.g., [10], [19], [32], [35]). However, given the simplicity of the tasks, we believe that our results can be generalized to healthy adults and teenagers without any major vision impairment or reading disabilities.

3.4 Procedure

The participants were seated in front of two liquid-crystal display monitors, each with a resolution of 1920×1080. On monitor 1, the participants read the task descriptions and entered their answers with the mouse or keyboard. Monitor 2 displayed a graphical user interface that showed the conventional maps and cartograms. The user interface was a web-based software application that allowed the user to trigger an animation by clicking a selector button (Fig. 1). Linked brushing and infotips were enabled by hovering over a map region with the mouse. We used Qualtrics XM to display the experiment tasks to the participants and collect their answers.

The experiment consisted of four parts.

1) *Introduction:* At the beginning of the experiment, all participants signed a form consenting to participate in the experiment. Then, they watched a five-minute video containing an introduction to cartograms and a description of the experiment. The participants were allowed to pause and rewind the video as they wished. They could also ask the experiment supervisor for clarification at any time. Afterwards, the participants had an opportunity to practice how to use cartogram-switching animations, linked brushing, and infotips. For this practice run, we showed participants a conventional map of the conterminous United States alongside an interactive cartogram whose areas represented either electoral votes, population, or land area. Participants did not have to complete any specific map reading task, but we required participants to confirm that they understood how each interactive feature worked before they were permitted to continue to the next stage of the experiment.
2) *Preliminary questions:* We collected information about each participant's age, gender, and level of education. We also asked participants to judge their familiarity with maps, cartograms, and interactive computer graphics using a 5-point Likert scale. Finally, we conducted an Ishihara color blindness test.
3) *Cartogram tasks:* The participants answered 40 multiple choice questions that required them to analyze a conventional map and one to two cartograms. If answer options consisted of named regions (i.e., for the task types *Cluster*, *Filter*, *Find Adjacency*, *Find Top*, and *Recognize*), we selected distractors randomly, but we aimed to include at least one distractor that appeared plausible after a cursory glance. Participants were informed that their responses and the amount of time taken to complete each task was recorded, but we also told them that there was no time limit for their answers. Additionally, we recorded the computer screen so that we could analyze how participants used the interactive features provided.
4) *Attitude study:* Adapting the semantic differential technique used by Dent [32] and Nusrat *et al*. [10], we asked participants to rate the aesthetics and effectiveness of the three interactive features evaluated in the experiment. We selected pairs of opposite words (e.g., "conventional" versus "innovative," "hindering" versus "helpful"). For every pair of words, we asked participants to rate each interactive feature on a 5-point Likert scale. Following the example of [10], participants did not have to give separate responses for each task type. Hence, the participants responded three times to each word pair, once for each interactive feature.

The participants were supervised in person, in a one-on-one setting. All participants completed the experiment in around 50 minutes. The instructional video used in part (1) and the complete list of questions used in parts (2)–(4) are available as supplemental material for this article, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TVCG.2020.3041745.

3.5 Design

We used an 8×5 within-subject experimental design with two independent variables: the eight task types listed in Table 1 and five experimental conditions, which depended on the availability of interactive features during the task:

- no interactivity,
- only a cartogram-switching animation is available,
- only linked brushing is available,
- only the infotip feature is available,
- all three features (i.e., animation, linked brushing, and infotips) are available.

Every combination of a task type with one of these five conditions appeared exactly once during each session. Thus,

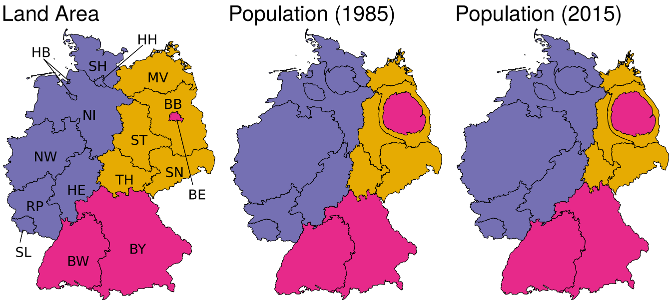

Fig. 3. Conventional map and cartograms presented to participants during the *Summarize* task for Germany. The map was divided into three zones: purple, pink, and yellow. Participants were required to identify whether each zone increased, decreased, or did not change in area between the first (1985 population) and second (2015 population) cartogram.

each participant completed exactly 40 trials during part (3) of the experiment. Because each participant encountered each combination of a task type and a feature-condition only once, we cannot infer whether participants became more efficient at using the features during the experiment. However, participants rated all features as "easy to use" at the end of the experiment (see Section 4.4), so we expect the learning curve to be almost flat.

For most task types, the participants had to select one answer out of four choices. However, for the *Filter* and *Find Adjacency* task types, participants could select multiple answers. For these questions, it was possible that more than one region matched the search criterion. All such regions had to be selected for the task to be completed correctly. *Summarize* tasks were split into three sub-tasks (one for each colored zone, see Fig. 3 for an example), but participants could only select one answer ("Growth," "Approximately no change," "Shrinking") for each zone. For this task type, we considered a participant's trial as a success only if the answers to all three sub-tasks were correct. We also deemed "Approximately no change" as an alternative correct answer if the change in area was no larger than 1 percent. In a pilot study, we noticed that a difference in area of below 1 percent is too subtle to be observed. We discuss the effect of setting different thresholds in Section 4.2.

The order of the tasks was the same for all participants. The order of the five interactive feature-conditions was counterbalanced using a Latin square. For the five tasks of the same type (e.g., the five tasks of the *Cluster* type), we used each of the five parts of the world listed in Section 3.2 (i.e., a different region for each task) to reduce the potentially confounding effect of different levels of familiarity with the displayed maps. All participants encountered the same random sequence of all 40 possible combinations of task types and parts of the world. The five interactive-feature conditions appeared in a random permutation for the five tasks of the same type. In Section 1 of the online supplemental text, available online, we include a table with the order in which participants encountered the combinations of map regions, task types, and interactive features.

3.6 Data Analysis

For each task type in Table 1, we wanted to compare the error rates and response times under the five different conditions listed in Section 3.5. To compare error rates, we treated the participants' responses as binary data (correct versus incorrect) and used Cochran's Q test with the null hypothesis that the interactive features had no influence on the probability of giving a correct response. The test statistic is χ^2-distributed with 4 degrees of freedom. For post-hoc analysis, we used pairwise McNemar tests. We applied the Bonferroni-Holm correction to adjust the p-values. We consider the adjusted p-values as significant if they are below 0.05. We caution against overinterpreting p-values [87] and, therefore, also state effect sizes and confidence intervals (CIs) for the post-hoc analysis. For the McNemar test, we measured the effect size with the odds ratio and determined CIs with the method developed by Fay [88]. The null hypothesis corresponds to an odds ratio equal to 1.

To compare the response times, we discarded all incorrect responses given by the participants. Consequently, our time observations are not necessarily paired across the five interactivity conditions. Even after discarding wrong responses, the distribution of response times is right-skewed. Outliers with long response times were presumably a consequence of our instruction to the participant that they could take as long as they needed to answer each question. With this instruction, we aimed to mimic a realistic scenario for reading cartograms in online news stories, where readers can look at cartograms without a fixed time limit. Because of the outliers, the response time distributions for some tasks fail the Shapiro-Wilk test of normality. Thus, we resorted to nonparametric Kruskal-Wallis tests to identify the main effects. The test statistic of the Kruskal-Wallis tests is χ^2-distributed with 4 degrees of freedom. For post-hoc analysis, we used pairwise Mann-Whitney U tests with the Bonferroni-Holm correction. We express the effect size of the Mann-Whitney U tests in terms of the pseudomedian difference [89], which we denote by Δ.

We have made the data and R scripts used for our statistical analysis publicly available at Zenodo [90].

3.7 Hypotheses

Prior to the experiment, we expected interactive features to make some tasks in Table 1 easier for the participants. Our hypotheses were as follows.

3.7.1 Cartogram-Switching Animations

Cartogram-switching animations are useful to compare the evolution of different quantities over time [63]. For this reason, we believed that animations may greatly help participants to perform the *Detect Change* and *Summarize* tasks. We foresaw a potentially positive effect on accuracy for these tasks, but the effect on the response time was unclear, as animations would also consume time.

- *H1*: Participants will make fewer errors with cartogram-switching animations in *Detect Change* and *Summarize* tasks.

3.7.2 Linked Brushing

Linked brushing is useful for participants because it highlights a selected region simultaneously on a regular map and a corresponding cartogram, allowing participants to quickly locate regions of interest on both representations. We believed that this feature could help participants to be more accurate when they execute *Find Adjacency*, *Find Top*, and *Recognize* tasks, because they can first identify the relevant region on the cartogram on monitor 2 by comparing it with the region's shape or size on monitor 1, then hover the mouse over this region on monitor 2, and finally read off the correct answer from the labeled conventional map. We also hypothesized that participants might be faster when using linked brushing.

- *H2-a*: Participants will make fewer errors with linked brushing in *Find Adjacency*, *Find Top*, and *Recognize* tasks.
- *H2-b*: Participants will need less time to perform these tasks using linked brushing.

3.7.3 Infotips

Infotips impart precise numeric information when the mouse hovers over a specific region, but reading the text in the infotip takes time. Interacting with infotips also demands from the users that they carefully control how the cursor moves between different parts of the maps. In addition, the infotip may occlude some parts of the maps, which could slow down map reading in general. Therefore, we expected that infotips would increase execution time. However, we believed that infotips could improve accuracy for all tasks. The only exception is the synoptic task *Summarize*, where information about small-scale individual regions is not directly relevant.

- *H3-a*: Participants will make fewer errors with infotips in all elementary tasks compared to the no-interactivity condition.
- *H3-b*: Participants will need *more* time when using infotips.

3.7.4 All Interactive Features

Using all interactive features should, theoretically, provide participants with more tools and information to perform the task. The three features do not visually interfere with each other, so we predicted that the all-features condition would lead to higher accuracy than all other conditions. However, when all three features are active, participants must process more information, so we expected the increased accuracy to come at the cost of increased execution time.

- *H4-a*: Unlike the no-interactivity condition, participants will make fewer errors for every task type when all interactive features are available.
- *H4-b*: Participants will generally need more time in the all-features condition than in the no-interactivity condition.

4 Results

4.1 Error Rates and Response Times

Participants made few errors overall. The distribution of the number of errors by participant is roughly symmetric and peaked around the mean (5.5 errors in 40 trials) with a range from 0 to 10 errors and a standard deviation of 2.3. (See Section 2 in the online supplemental text, available online, for more summary statistics.) Because of the narrow distribution, we can regard the group of participants as sufficiently homogeneous to include the responses of all participants for further statistical analysis.

In Fig. 4, we summarize the results of the data analysis outlined in Section 3.6. Tabular summaries of the error rates and average response times are in Sections 2 and 3 of the online supplemental text, available online. In the following list, we provide details about the results for each task type.

- *Cluster:* Judging by a mean error rate of 15.3 percent, this task type was of intermediate difficulty. Interactive features caused a significant effect on the accuracy [$\chi^2(4) = 16.33$, $p < 0.01$], but did not have a significant effect on response times [$\chi^2(4) = 2.69$, $p = 0.61$]. When participants had access to all interactive features, they made significantly fewer errors than when they were only allowed to use cartogram-switching animations (5.5 versus 30.9 percent, odds ratio 0.13, 95 percent CI $[0.02, 0.78]$, $p = 0.01$).
- *Compare:* The participants found this task type slightly easier than *Cluster* (mean error rate 9.5 percent). We observed a significant effect of the interactive features on error rates [$\chi^2(4) = 12.41$, $p = 0.01$] and response times [$\chi^2(4) = 13.54$, $p < 0.01$]. For the error rates, the post-hoc analysis reveals a pairwise difference between the all-features and linked-brushing-only conditions (1.8 versus 20.0 percent, odds ratio 0.09, 95 percent CI $[0.01, 0.96]$, $p = 0.04$). For the response times, we find a significant difference between the all-features (median $28.1\,\mathrm{s}$) and infotip-only conditions (median $36.9\,\mathrm{s}$, $\Delta = 7.3\,\mathrm{s}$, 95 percent CI $[0.9\,\mathrm{s}, 14.0\,\mathrm{s}]$, $p = 0.02$).
- *Detect Change:* For tasks of this type, the participants achieved a low mean error rate of 9.1 percent, similar to their performance for *Compare*. Differences between the interactive features did not seem to impact the accuracy; the p-value for the main effect is 0.76 [$\chi^2(4) = 1.86$]. However, we find a significant effect on response times [$\chi^2(4) = 13.85, p < 0.01$]. Specifically, participants were significantly faster (median $23.7\,\mathrm{s}$) under the linked-brushing-only condition than the all-features condition (median $32.1\,\mathrm{s}$, $\Delta = 7.0\,\mathrm{s}$, 95 percent CI $[0.4\,\mathrm{s}, 14.6\,\mathrm{s}]$, $p = 0.03$) or infotip-only condition (median $32.2\,\mathrm{s}$, $\Delta = 6.8\,\mathrm{s}$, 95 percent CI $[0.4\,\mathrm{s}, 14.2\,\mathrm{s}]$, $p = 0.03$).
- *Filter:* The mean error rate of *Filter* (16.4 percent) is close to that of *Cluster*. *Filter* is also similar to *Cluster* in having an effect on accuracy [$\chi^2(4) = 13.50$, $p < 0.01$], but not on response times [$\chi^2(4) = 6.50$, $p = 0.16$]. The participants' error rates were significantly lower under the all-features condition (7.3 percent) than under the linked-brushing-only condition (29.1 percent, odds ratio 0.20, 95 percent CI $[0.04, 0.99]$, $p = 0.05$).
- *Find Adjacency:* The participants found this task type easy (mean error rate 3.3 percent). The error rates are low for all conditions, ranging from 0 percent for linked-brushing-only to 5.5 percent if there is no interactivity. The response times also hardly deviate from the median ($21.5\,\mathrm{s}$). Therefore, the interactive

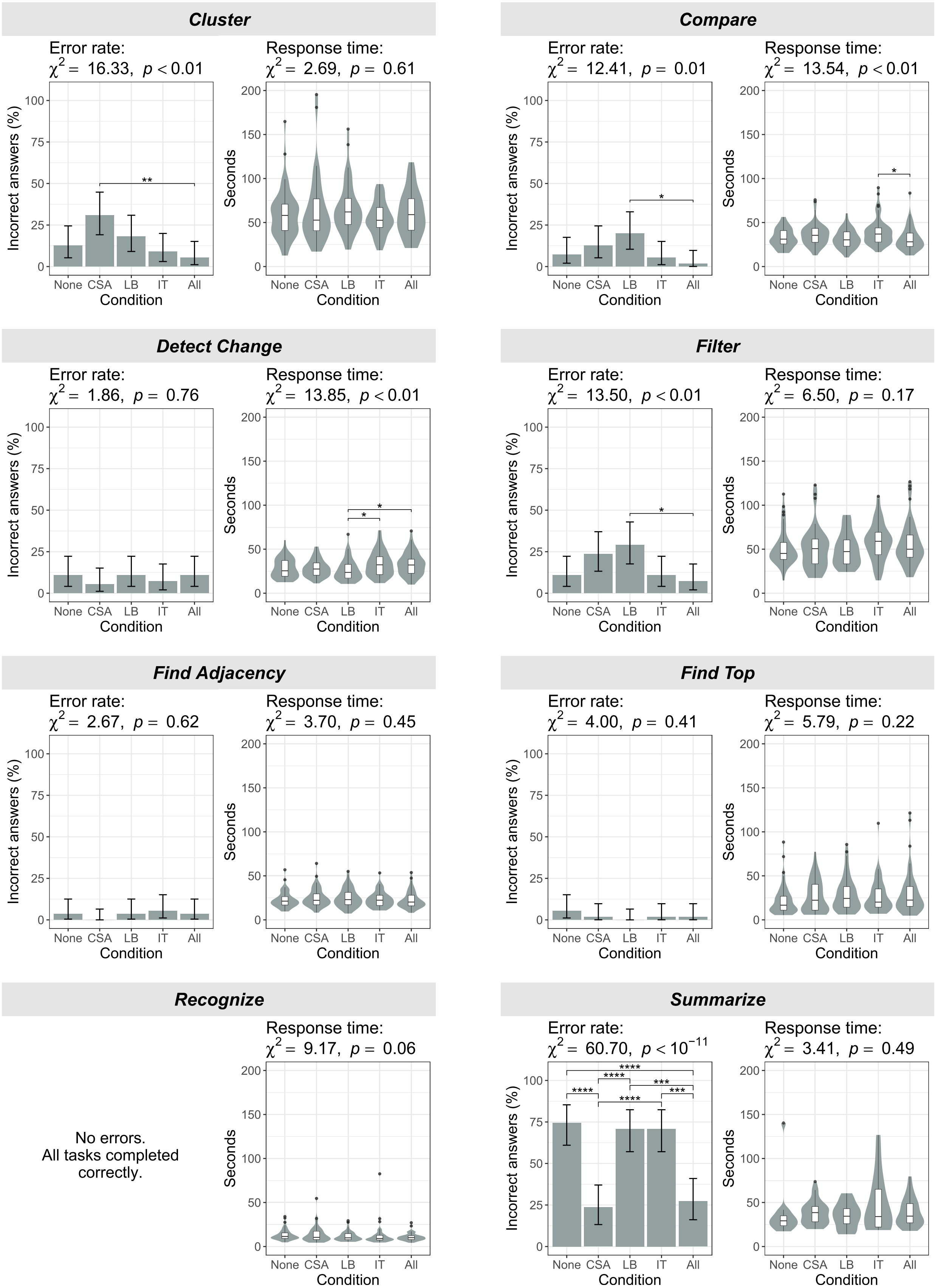

Fig. 4. Error rates and response time distributions for the cartogram task types in Table 1. We use the following abbreviations for the axis labels. CSA: cartogram-switching animation is the only available interactive feature. LB: linked brushing only. IT: infotip only. Brackets inside the panels indicate significant differences between pairs of conditions at a significance level of 0.05. Asterisks above the brackets indicate p-values. $*$: p-value ≤ 0.05, $**: \leq 0.01$, $***: \leq 0.001$, $****: \leq 0.0001$. Error bars represent 95 percent CIs.

features had no significant effect on error rates [$\chi^2(4) = 2.67, p = 0.62$] or on response times [$\chi^2(4) = 3.70, p = 0.45$].

- *Find Top:* The mean error rate for this task type is similarly low (2.2 percent) as that for *Find Adjacency* with little variability across different conditions (0 to 5.5 percent). We detect no significant effect on accuracy [$\chi^2(4) = 4.00, p = 0.41$] or on response times [$\chi^2(4) = 5.79, p = 0.22$].
- *Recognize:* All participants answered every *Recognize* task correctly. Therefore, no further analysis is needed for the error rates. This task type also had the shortest median response time (10.0 s) with no significant difference between the experimental conditions [$\chi^2(4) = 9.17, p = 0.06$].
- *Summarize:* This task type was clearly the most challenging in terms of the mean error rate (53.5 percent). Participants' performance depended strongly on the experimental condition. The error rates range from 23.6 percent for the cartogram-switching-animation-only condition to 74.5 percent when there was no interactivity. The p-value for a main effect is accordingly small [$\chi^2(4) = 60.70,\ p < 10^{-11}$]. The post-hoc analysis reveals that cartogram-switching animations—either as the only interactive feature or in combination with all other features—significantly improved performance compared to all other experimental conditions. For all pairwise McNemar tests that involve one condition with animations and another without animations, we find p-values below 0.01 and CIs that clearly exclude the null hypothesis of an odds ratio equal to 1. For example, the 95 percent CI for the comparison between cartogram-switching-animation-only and the no-interactivity treatment is $[0.00, 0.35]$. In contrast to the clear benefit of animations in terms of accuracy, we do not find a main effect of interactivity on response times [median$= 34.1$ s, $\chi^2(4) = 3.41, p = 0.49$].

4.2 Performance for Summarize

Because *Summarize* is the task type with the highest error rate, we conducted further data analysis. We first investigated whether the statistical conclusions depend on the numerical threshold for accepting "Approximately no change" as a correct answer. Then we examined whether different user groups benefit from animations to the same degree.

As we explain in Section 3.5, a pilot study revealed that the minimal detectable area change is $\theta \approx 1\%$. Repeating the analysis for thresholds equal to $\theta_{\text{low}} = 0.5\%$ and $\theta_{\text{high}} = 2\%$, we find that the mean error rate decreases as the threshold increases (from 68.4 percent for θ_{low} to 34.5 percent for θ_{high}). This tendency is expected, but it is noteworthy that, regardless of which threshold we choose, the Cochran Q test always rejects the null hypothesis that interactive features had no effect on the error rates. The numeric results can be found in Section 2 of the online supplemental text, available online. The post-hoc McNemar tests also always identify the same significant pairwise differences between interactive features. Hence, our conclusion that animations are a significant help when answering *Summarize* tasks is independent of the exact definition for the minimal detectable area change θ.

We also find evidence that this conclusion is valid for user groups with different levels of prior experience. We divided participants into two categories based on whether they considered themselves to be familiar with interactive computer graphics (≥ 4 on a 5-point Likert scale) during the preliminary questions. The group with greater familiarity contained 25.5 percent of the participants. We observed that interactivity greatly reduces the error rate in this group from 71.4 percent (no features) to 14.3 percent (all features). The error rates for the second group are higher for both conditions, but we find again a clear improvement from 75.6 percent (no features) to 31.7 percent (all features).

The same trend occurs when we divide the participants into two categories based on their general affinity with maps. To infer their attitude towards maps, we posed the following preliminary question: "When you encounter the names of unfamiliar locations (e.g., countries, islands, lakes), how frequently do you immediately look them up on a map to find out where they are?" We dichotomize the participants depending on whether their answer was ≥ 4 on a 5-point Likert scale. The group who declared a tendency towards reading maps (34.5 percent of the participants) benefited tremendously from the interactive features, decreasing their error rate from 68.4 percent (no features) to 26.3 percent (all features). The other group had slightly higher error rates, but we still observe a marked decrease associated with the interactive features (77.8 percent without any features, 27.8 percent with all features).

Because linked brushing or infotips in isolation do not result in reduced error rates for *Summarize* tasks (Fig. 4), animations seem to be the main reason for the improved performance in the all-features condition. In Section 2 of the online supplemental text, available online, we show that the positive effect of animations—either as the only available feature or in combination with the other two features—is independent of the participants' confidence using interactive computer graphics and independent of their personal inclination towards reading maps. Conversely, *Summarize* is the only task type in which animations were associated with significantly improved performance (see Section 2 of the online supplemental text, available online), but error rates for elementary tasks were generally so low that there would not have been much room for improvement.

4.3 Hypotheses

In terms of our hypotheses in Section 3.7, *H1 is partially supported* because we found that cartogram-switching animations improved accuracy in *Summarize* tasks but not in *Detect Change* tasks. For linked brushing, we observed neither any significant decrease in the error rate nor in the response time compared to the no-interactivity baseline. Thus, *H2-a and H2-b are rejected*. The results for infotips were also inconclusive, so *H3-a and H3-b are also rejected*. For seven out of eight task types, the all-features condition showed no significant improvement in accuracy. The noteworthy exception is *Summarize*. Hence, *H4-a is partially supported*, but the all-features performance in *Summarize* was only on par with the cartogram-switching animation, suggesting that the gain in accuracy is caused by that specific feature. We did not find any significant increase in response times under the all-features condition, so *H4-b is rejected*.

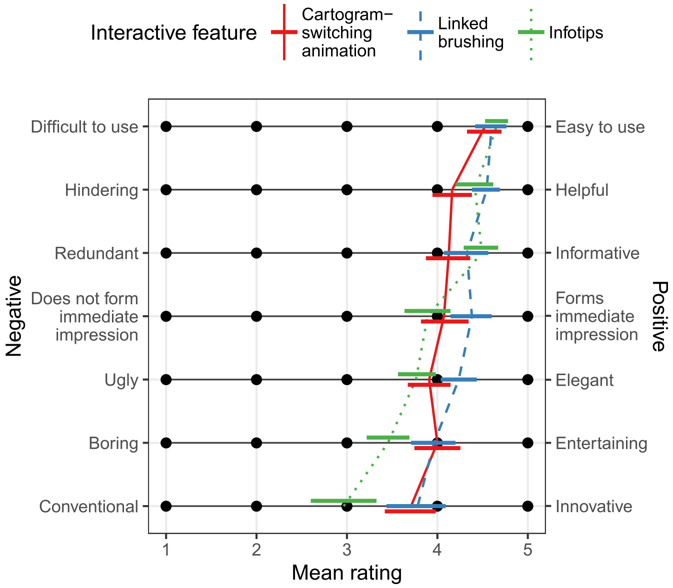

Fig. 5. Mean ratings in the attitude study conducted at the end of our experiment. Horizontal bars are bootstrap estimates of the 95 percent CIs. We sorted the phrase pairs along the vertical axis so that the pair with the overall strongest positive score ("Easy to use") is at the top.

4.4 User Preferences

In the final part of the experiment, we presented seven pairs of phrases to the participants. Each pair consisted of two phrases with opposite meaning:

- Difficult to use – Easy to use,
- Does not form immediate impression – Forms immediate impression,
- Conventional – Innovative,
- Redundant – Informative,
- Hindering – Helpful,
- Boring – Entertaining,
- Ugly – Elegant.

We asked the participants to rate each of the three interactive features (i.e., cartogram-switching animation, linked brushing, infotips) in terms of these phrases on a 5-point Likert scale. We show the mean rating for each combination of an interactive feature and a phrase pair in Fig. 5. More information about the distribution can be found in Sections 4 and 5 of the online supplemental text, available online.

The participants gave positive ratings (i.e., mean $>$ 3) for 20 out of 21 combinations of features and phrase pairs. Averaged over all interactive features, the positive phrase with the highest rating was "Easy to use" (mean rating 4.59). The positive phrase with the weakest agreement was "Innovative" (mean rating 3.48). Taking the average of all pairs of phrases, linked brushing achieved the highest approval (mean 4.26), closely followed by cartogram-switching animations (4.07) and infotips (3.95). In summary, participants gave strong subjective feedback about all three features even though only cartogram-switching animations led to objective improvements in performance.

5 Discussion

For most tasks in our experiment, we find that the average error rates are below 10 percent. Our results are in line with previous observations that most readers can easily extract information from cartograms even without interactivity [26], [32]. The most notable exception is the task type *Summarize* with an average error rate of 53.5 percent. This task was challenging because participants had to distinguish subtle area differences. In Fig. 3, for example, the purple zone in the northwest increases by only 1.4 percent from the 1985 cartogram to the 2015 cartogram. *Summarize* tasks, therefore, asked participants to assess differences between cartograms more carefully than any other task, resulting in lower overall accuracy. Our observation is consistent with results obtained by Kasper *et al.* [35], who also noted that the error rates in their experiment depended strongly on the complexity of the cartogram task.

5.1 Effect of Cartogram-Switching Animations on Performance

We found that a cartogram-switching animation was an effective way to improve the accuracy in *Summarize* tasks, dramatically reducing the error rate from above 70 percent to around 30 percent. Animations may make area changes much easier to detect because the viewer does not have to shift the gaze between two spatially separated cartograms. Moreover, the smooth transition of the regions' boundaries helps to detect the direction of movement so that it becomes clearer whether the enclosed area expands or contracts.

In contrast to the synoptic *Summarize* tasks, we found that for elementary task types, cartogram-switching animations were less effective, but were not associated with a practically significant decrease in performance. The error rate of *Cluster* tasks increased nominally if animations were the only available feature (Fig. 4). However, when we checked the screen recordings, we found that only 17 out of 54 participants actually used cartogram-switching animations when performing a *Cluster* task under the animation-only condition. (For one participant, we have no screen recording of this task.) The error rate for these 17 participants was almost the same as for the remaining participants (29.4 percent with and 29.7 percent without using animations). A two-sample proportion test yields a p-value close to 1 with a 95 percent CI of $[-26.2\%, 26.8\%]$ for the difference in proportion. Moreover, the animations were also available under the all-features condition, where the accuracy was higher than in the case of no interactivity. Therefore, there is no evidence that cartogram-switching worsened the performance for any task type. In Section 7 of the online supplemental text, available online, we give more details about the frequency with which participants chose to use animations for each task type.

Participants rated cartogram-switching animations as the most entertaining feature (4.00 out of 5). Especially during *Summarize* tasks, the animations piqued the participants' interest. In all *Summarize* trials in which an animation was available (i.e., under the cartogram-switching-animation-only and all-features conditions), the screen recordings show that all participants played the animation multiple times. In many cases, the participants may have repeated the animation to confirm that their answer was correct. Because the buttons for switching between different cartograms were placed next to each other (e.g., the box labeled "4" in Fig. 1 shows the neighboring buttons for GDP and land area), it only needed small hand movements and little motor control to play the animations multiple times. In addition, participants may have felt encouraged to replay the animation because we told them during the introduction that they could take as long as they needed to answer questions. Although each animation only lasted for one second, repeated use of this feature is the most likely reason we do not observe a significant speedup for *Summarize* tasks. Ware [27] reached a similar conclusion in her cartogram experiment. We agree that the higher accuracy and the positive subjective ratings for cartogram-switching animations more than outweigh the (statistically insignificant, in our experiment) increase in response times. Another strong argument in favor of cartogram-switching animations is that they improved the accuracy even for participants who were less confident users of computer graphics or less inclined to read maps.

5.2 Effect of Infotips on Performance

For all eight task types, the presence of infotips did not negatively impact the accuracy of the participants' responses in a statistically significant way. Although in some cases the error rate nominally increased when infotips were the only available feature, none of the pairwise post-hoc tests revealed a significant deterioration compared to the case of no interactivity. In the combination of all three features, infotips did not appear to be detrimental to accuracy either.

Infotips did not significantly affect participants' response times compared to the no-interactivity condition (Fig. 4). The only effect we observe is that the infotip-only condition was significantly slower than the all-features condition for *Compare* tasks. Judging from the screen recordings, it appears likely that most participants with access to an infotip performed the *Compare* tasks by reading the numbers in the pop-ups for the two regions mentioned in the question rather than by visually comparing cartogram areas. Retrieving information from text tends to be slower than from diagrams [91]. Furthermore, after reading the numbers in one pop-up, participants had to shift the mouse to a different region to view the numbers in another pop-up. Consequently, this strategy was slower than estimating the regions' areas by eye from the cartograms. An infotip might also have distracted from the task by obstructing parts of the cartograms, thus explaining the observed slowdown.

Consistent with this hypothesis, the participants rated infotips slightly lower in terms of "forming an immediate impression" compared to the other interactive features (Fig. 5). Infotips also received only intermediate ratings halfway between "conventional" and "innovative" (2.96 on a scale from 1 to 5), presumably because similar mouse-over effects are currently common on many websites. It is also conceivable that infotips were less popular because the participants had to place the mouse pointer directly on top of the region of interest, which might have felt tedious, especially if the task required activating infotips related to distant regions. Nevertheless, the overall ratings for infotips are positive, and they do not lead to a significant loss of accuracy compared to the no-interactivity condition for any of the task types (Fig. 4). We therefore still recommend including an infotip with interactive cartograms. In our experiment, the pop-up immediately displayed the regions' statistics as soon as the mouse hovered over the map. The pop-up disappeared only when participants moved the mouse off the map. A small change in the design of the infotip may make it less obtrusive: if participants can display and hide the pop-up with a mouse click, the infotip will not permanently obstruct space on the map.

5.3 Effect of Linked Brushing on Performance

Like infotips, linked brushing did not appear to be detrimental to accuracy. Furthermore, the ratings for linked brushing were positive across the board. Linked brushing is a subtle, unobtrusive feature that did not cause a significant increase in response time for any task type. Unlike infotips, linked brushing needs less hand motor control, because the highlighting becomes visible when the mouse pointer moves across the region of interest, but does not need to be placed precisely on top of it. Our hypotheses H2-a and H2-b, that linked brushing would improve performance in *Find Adjacency*, *Find Top*, and *Recognize* tasks, were not supported. However, the error rates and response times for these task types were generally low in our experiment, so there was not much room for improvement. For maps with a larger number of regions than those we used in our experiment, linked brushing may have greater benefits.

5.4 Effect of Map Presentation, Map Complexity, and Familiarity on Performance

On all conventional maps, we labeled every administrative unit with a two-letter or three-letter identifier (see Fig. 1). Space permitting, these labels were placed near the center of the region. Otherwise, labels were placed outside the map and connected to the corresponding regions by lead-out lines. Screen recordings indicate that a few participants may have been confused by the density of lead-out lines on some maps, so they may have misidentified some map regions. However, the videos suggest that most participants interpreted region labels and lead-out lines correctly. We conclude that the presentation of these map elements is unlikely to have had a measurable effect on participants' overall performance.

In Section 6 of the online supplemental text, available online, we present statistics about the participants' performance for cartograms of different countries. In general, the accuracy of the participants' responses did not seem to depend on the country shown on the map. However, we found clear statistical evidence that the country shown on the map had an influence on the response time. The responses for Brazil (median $33.9\,\mathrm{s}$) as well as for mainland China and Taiwan ($33.2\,\mathrm{s}$) were significantly slower than for the United States (median $23.2\,\mathrm{s}$). A regression analysis in Section 6 of the online supplemental text, available online, does not find evidence that the median response time increased with the number of administrative units, which was highest on the US map. Instead, we hypothesize that the response time was mainly influenced by prior familiarity with the conventional map: the US state map is a recognizable icon worldwide, but few Singaporean readers are regularly exposed to maps of Brazilian states. In this context, it is important to note that our experimental design controlled for country-dependent differences because the displayed countries were a blocking factor (see Section 3.5): for each task type, there were equally many participants assigned to using each combination of a country with an interactive-feature condition.

Although our experiment gave no indication that performance depended on the number of administrative units, we believe that the interactive features work best for the intermediate range we used in our experiment (from 16 for Germany to 49 for the US). If the units on the screen become too numerous and hence too small, it will become difficult to position the mouse pointer with enough precision to activate the correct infotip or highlight the target region with linked brushing. We hypothesize that the effectiveness of animations is less dependent on the number of administrative units because animations can always be triggered with the same degree of motor control. However, a larger number of polygons on the map may reduce an animation's effectiveness owing to the increased cognitive load. It will be an interesting task for future research to investigate this hypothesis.

5.5 Generalizability of the Results to Other Cartogram Types

The cartograms in our experiment were the result of a continuous map projection [86]. We believe that most of our results apply to other contiguous cartogram types too (e.g., rectilinear [92] or mosaic cartograms [6]) although the absence of a map projection may make it more difficult for viewers to mentally establish the underlying transformation. If the cartograms are noncontiguous (i.e., regions are displayed by disconnected polygons), the error rates of *Find Adjacency* tasks will presumably be higher than those we found in our experiment. A switching animation that morphs a noncontiguous cartogram into a conventional map would reveal information about adjacency that is not obvious from a still image. Therefore, animations may have a stronger positive effect when working with noncontiguous instead of contiguous cartograms.

5.6 Generalizability to Bivariate Cartograms

The cartograms in our experiment were univariate maps: we only represented one variable per cartogram (e.g., only population size or only GDP), and areas were the only visual variable that conveyed information about these numbers. On univariate cartograms, colors can be freely chosen to support readability. Here we selected fill colors from the 6-class ColorBrewer palette "Dark2" [93], which is designed to work well on liquid-crystal display screens such as those used in our experiment. We deliberately chose a dark palette so that we could reserve bright colors for the linked brushing effect. The combination of colors was unlikely to be misconstrued as a way to represent data. We chose different colors for all neighboring regions, but corresponding regions on the juxtaposed conventional maps and cartograms had the same color so that matching pairs were easy to spot even without linked brushing.

For bivariate cartograms, linked brushing may be more essential for readability. If colors represent categorical or quantitative data, it may be inevitable that neighboring regions are filled with the same color. For example, on the classic US presidential election cartogram [94], one variable (number of electors) is represented by area and the other variable (party affiliation of the electors) by a binary color scheme: red for Republican, blue for Democratic. On such cartograms, large contiguous swathes of states typically appear in the same color (e.g., the entire South is usually red, the Northeast blue). We hypothesize that even elementary cartogram reading tasks become substantially more challenging if the reader cannot take for granted that neighboring regions have distinct colors. Linked brushing may reduce the challenge posed by tasks of the types *Detect Change* and *Recognize*, which can be answered by finding matching pairs of regions on two juxtaposed maps. However, linked brushing must be implemented judiciously. If colors represent data values, we advise avoiding confusion and refraining from changing the color to highlight the region under the mouse pointer. For bivariate maps, increasing the border thickness promises to be a more effective form of linked brushing than changing the fill color.

Fewer changes are needed to implement the other two interactive features in our study (i.e., infotips and animations) for bivariate cartograms. As suggested by Nusrat *et al.* [95], the text in the infotip can simply reveal the data for both variables simultaneously. If colors are used to represent one of the thematic mapping variables, a cartogram-switching animation would have to depict area and color changes simultaneously. Judging from previous experiments [96], there is no simple rule whether colors should change abruptly, or whether one should apply tweening to achieve a smooth color transition. In some cases, animations may in fact make bivariate cartograms entirely superfluous. As an alternative to representing two mapping variables with two different visual variables (area and, for example, color), we can make two univariate cartograms and allow users to discover the differences between the cartograms with an animation.[1] Animated univariate cartograms allow viewers to concentrate on a single visual variable (area). We hypothesize that dynamic changes in a single visual variable are easier to detect synoptically than associations between two different visual variables, especially if the animation between two univariate cartograms can be played repeatedly in both directions. The present study does not allow us to compare animations to bivariate cartograms, but it encourages future research in this direction.

5.7 Guidelines for Adding Interactivity to Cartograms

Our experimental results suggest that the need for interactivity depends on the complexity of the task that the cartogram designer has in mind (i.e., whether the task is "elementary" or "synoptic" according to the typology by Andrienko and Andrienko [84]).

- *Elementary tasks:* For all seven elementary task types in our experiment (*Cluster*, *Compare*, *Detect Change*, *Filter*, *Find Adjacency*, *Find Top*, and *Recognize*), error rates without any interactivity were very low—below 13 percent. Response times for elementary tasks did

1. If both mapping variables are measured in the same units and add up to a meaningful total, Nusrat *et al.*'s [95] bivariate pie-chart cartograms make it possible to use color as a single visual variable for both mapping variables. However, additive bivariate data is an exception rather than the rule.

not depend significantly on the availability of interactive features. We conclude that most readers can effectively decode the essential information presented by contiguous cartograms even if the display is static. However, we recommend following Dent's advice [32] to always show a cartogram together with a conventional map, which replicates our experimental setting.

- *Synoptic tasks:* For the only synoptic task in our experiment (*Summarize*), animations clearly had a positive effect. The error rates were more than halved when an animation was available. There was no significant effect on the response time, which is partly explained by our observation that animations were always played repeatedly. If interactive cartograms become less of a rarity in the future, we may find that more experienced users play the animation less often, thus improving efficiency in the long run. In addition to answering questions about the shown cartograms more accurately, participants also expressed a subjective preference for animations, so we recommend including them regularly if the intended goal is to perform a synoptic task. The other two interactive features (linked brushing and infotips) neither improved nor worsened the performance. They can be added as an option but are not strictly necessary.

Because we applied the task taxonomy of Nusrat *et al.* [30], our experiment mainly focused on elementary tasks. However, one can argue that the primary purpose of cartograms is to perform synoptic tasks. After all, cartograms of any kind make geometric compromises, either in the shapes or in the contiguity. If all we want to accomplish is an elementary task, then other map types have the advantage of familiarity to most users. It would be worth expanding the existing cartogram task taxonomy in the future with a greater variety of synoptic tasks, where cartograms can play to their strengths (e.g., detecting correlations between different variables). Still, even with only one synoptic task type included in our experiment, we have already found clear evidence in favor of cartogram-switching animations and recommend this feature as a minimum of interactivity.

We have developed a web application (https://go-cart.io/) that demonstrates our implementations of the interactive features described in this paper [97]. Apart from our minimal recommendation of including animations, we combine them on go-cart.io with linked brushing and infotips. These additional features did not lead to measurable improvements in our experiment, but they were not harmful either. Because of the generally favorable subjective evaluations by participants in our experiments, our overall recommendation is the full suite of features as implemented by go-cart.io.

6 Conclusion

Despite their inherent distortion, cartograms can be an effective tool in the cartographer's toolbox for displaying geospatial statistics. Previous research has already recommended best practices for displaying printed cartograms [28], [32], [98], such as including a conventional map and a legend as a reference for the reader. Based on the results of our experiment, we add another recommendation: if displayed electronically, cartograms should be presented with interactivity. We found that readers performed synoptic *Summarize* tasks much more accurately when they had access to cartogram-switching animations. Participants also expressed strongly positive opinions about the other two interactive features that we tested in this experiment (i.e., linked brushing and infotips). They characterized them as easy to use, informative, and helpful. Therefore, we recommend that cartograms that are displayed electronically—such as those on https://go-cart.io—should include all three features. We hope that this website will simplify reading and drawing interactive cartograms, similar to the way in which technologies such as Observable [99] and Vega-Lite [100] have simplified the creation of other types of interactive graphics.

Acknowledgment

This work was supported by the Singapore Ministry of Education under Grant R-607-000-401-114 and a Yale-NUS College start-up under Grant R-607-263-043-121. The authors are grateful to Barry J. Kronenfeld for his comments on the manuscript. They would also like to thank Editage (www.editage.com) for English language editing. They are indebted to the Associate Editor and three anonymous referees for their insightful comments about earlier drafts of this manuscript.

References

[1] B. D. Dent, J. Torguson, and T. W. Hodler, *Cartography: Thematic Map Design*, 6th ed. New York, NY, USA: McGraw-Hill Education, 2008.

[2] B. J. Kronenfeld and D. W. S. Wong, "Visualizing statistical significance of disease clusters using cartograms," *Int. J. Health Geograph.*, vol. 16, no. 1, 2017, Art. no. 19.

[3] T. Gotthard and B. D. Hennig, "Worldmapper: Rediscover the world as you've never seen it before.," 2005. Accessed: Oct. 04, 2019. [Online]. Available: https://worldmapper.org/

[4] B. D. Hennig, J. Pritchard, M. Ramsden, and D. Dorling, "ReMapping the world's population," *ArcUser*, vol. 2010, no. 1, pp. 66–69, 2010.

[5] D. Dorling, "London Mapper aims to show existing data in new light," *The Guardian*, 2014. Accessed: Dec. 8, 2020. [Online]. Available: https://www.theguardian.com/commentisfree/2014/may/11/london-mapper-data-new-light

[6] R. G. Cano, K. Buchin, T. Castermans, A. Pieterse, W. Sonke, and B. Speckmann, "Mosaic drawings and cartograms," *Comput. Graph. Forum*, vol. 34, no. 3, pp. 361–370, 2015.

[7] J. M. Olson, "Noncontiguous area cartograms," *Prof. Geogr.*, vol. 28, no. 4, pp. 371–380, 1976.

[8] D. Dorling, *Area Cartograms: Their Use and Creation*. Concepts and Techniques in Modern Geography series no. 59, University of East Anglia: Environmental Publications, 1996.

[9] H. Sun and Z. Li, "Effectiveness of cartogram for the representation of spatial data," *Cartograph. J.*, vol. 47, no. 1, pp. 12–21, 2010.

[10] S. Nusrat, M. J. Alam, and S. Kobourov, "Evaluating cartogram effectiveness," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 2, pp. 1077–1090, Feb. 2018.

[11] E. Raisz, "The rectangular statistical cartogram," *Geographical Rev.*, vol. 24, no. 2, pp. 292–296, 1934.

[12] The New York Times, "The electoral map: Building a path to victory," 2012. Accessed: Jul. 02, 2019. [Online]. Available: https://www.nytimes.com/elections/2012/ratings/electoral-map.html

[13] A. Zitner, R. Yeip, and J. Wolfe, "Draw the 2016 electoral college map," 2016. Accessed: Jul. 28, 2019. [Online]. Available: https://graphics.wsj.com/elections/2016/2016-electoral-college-map-predictions/

[14] Los Angeles Times, "U.S. election results," 2012. Accessed: Jul. 02, 2019. [Online]. Available: http://graphics.latimes.com/2012-election-results-national-map/

[15] The Telegraph, "Election 2015 results mapped: 2015 full list," 2015. Accessed: Jul. 02, 2019. [Online]. Available: https://www.telegraph.co.uk/news/general-election-2015/11584325/full-results-map-uk-2015.html
[16] W. Franklin, J. Holder, M. Osborn, S. Clarke, and N. Kommenda, "EU referendum full results: Find out how your area voted," 2016. Accessed: Jul. 03, 2019. [Online]. Available: http://www.theguardian.com/politics/ng-interactive/2016/jun/23/eu-referendum-live-results-and-analysis
[17] H. Fairfield, "The changing face of American Catholics," 2018. Accessed: Jul. 02, 2019. [Online]. Available: https://archive.nytimes.com/www.nytimes.com/interactive/2008/04/13/us/20080412_CATHOLIC_GRAPHIC.html
[18] L. Byron, A. Cox, and M. Ericson, "A map of Olympic medals," 2018. Accessed: Jul. 02, 2019. [Online]. Available: https://archive.nytimes.com/www.nytimes.com/interactive/2008/08/04/sports/olympics/20080804_MEDALCOUNT_MAP.html?_r=0
[19] K. A. Rittschof, W. A. Stock, R. W. Kulhavy, M. P. Verdi, and J. T. Johnson, "Learning from cartograms: The effects of region familiarity," *J. Geography*, vol. 95, no. 2, pp. 50–58, 1996.
[20] A. Fotheringham, C. Brunsdon, and M. Charlton, *Quantitative Geography*. London, U.K.: Sage Publications, 2007.
[21] D. Eckert, J. Fontanabona, L. Jégou, S. Robert, and G. Sillère, "Le site worldmapper. Commentaire sur l'article d'Anna Barford," *Mappemonde*, vol. 89, no. 1, 2008, Art. no. 08106.
[22] R. E. Roth, A. W. Woodruff, and Z. F. Johnson, "Value-by-alpha maps: An alternative technique to the cartogram," *Cartograph. J.*, vol. 47, no. 2, pp. 130–140, 2010.
[23] D. Clark and R. Houston, "The carbon map.," 2012. Accessed: Apr. 11, 2020. [Online]. Available: https://www.carbonmap.org/
[24] BBC News, "Results of the 2019 General Election," 2019. Accessed: Apr. 11, 2020. [Online]. Available: https://www.bbc.com/news/election/2019/results
[25] N. Kommenda, S. Clarke, and F. Hulley-Jones, "Election results 2019: Boris Johnson returned as PM with all constituencies declared," 2019. Accessed: Apr. 11, 2020. [Online]. Available: https://www.theguardian.com/politics/ng-interactive/2019/dec/12/uk-general-election-2019-full-results-live-labour-conservatives-tories
[26] J. J. R. Nuñez and B. Juhász, "Hungarian survey on the use of cartograms in school cartography," *Int. J. Cartography*, vol. 1, no. 1, pp. 5–17, 2015.
[27] J. A. Ware, "Using animation to improve the communicative aspect of cartograms," Master's thesis, Dept. Geography, Michigan State Univ., East Lansing, MI, USA, 1998.
[28] W. Tobler, "Thirty five years of computer cartograms," *Ann. Amer. Assoc. Geographers*, vol. 94, no. 1, pp. 58–73, 2004.
[29] A. Herzog, "Developing cartographic applets for the Internet," in *Maps and the Internet*, M. Peterson, Ed. Amsterdam, The Netherlands: Elsevier, 2005.
[30] S. Nusrat and S. Kobourov, "Task taxonomy for cartograms," in *Proc. Eurographics Conf. Vis.*, 2015. Accessed: Dec. 8, 2020. [Online]. Available: https://www2.cs.arizona.edu/~kobourov/cartogram_taxonomy.pdf
[31] A. Woodruff, "I hate your favorite election map.," 2008. Accessed: Sep. 05, 2019. [Online]. Available: http://andywoodruff.com/blog/i-hate-your-favorite-election-map/
[32] B. D. Dent, "Communication aspects of value-by-area cartograms," *Amer. Cartographer*, vol. 2, no. 2, pp. 154–168, 1975.
[33] B. D. Dent, "A note on the importance of shape in cartogram communication," *J. Geography*, vol. 71, no. 7, pp. 393–401, 1972.
[34] T. L. C. Griffin, "Recognition of areal units on topological cartograms," *Amer. Cartographer*, vol. 10, no. 1, pp. 17–29, 1983.
[35] S. Kaspar, S. I. Fabrikant, and P. Freckmann, "Empirical study of cartograms," in *Proc. 25th Int. Cartograph. Conf.*, 2011, Art. no. CO-112.
[36] C. Aschwanden, "Kognitionsstudien mit mengentreuen Flächenkartogrammen," *Kartographische Nachrichten*, vol. 48, no. 6, pp. 221–228, 1998.
[37] R. Han, Z. Li, P. Ti, and Z. Xu, "Experimental evaluation of the usability of cartogram for representation of GlobeLand30 data," *ISPRS Int. J. Geoinf.*, vol. 6, no. 6, 2017, Art. no. 180.
[38] P. Gao, Z. Li, and Z. Qin, "Usability of value-by-alpha maps compared to area cartograms and proportional symbol maps," *J. Spatial Sci.*, vol. 64, no. 2, pp. 239–255, 2018.
[39] A. Markowska, "Using area cartograms in teaching geography in secondary schools," *Abstr. Int. Cartograph. Assoc.*, vol. 1, pp. 1–2, 2019.
[40] M. Bostock, V. Ogievetsky, and J. Heer, "D^3 data-driven documents," *IEEE Trans. Vis. Comput. Graphics*, vol. 17, no. 12, pp. 2301–2309, Dec. 2011.
[41] M. F. Goodchild, "Stepping over the line: Technological constraints and the new cartography," *Amer. Cartographer*, vol. 15, no. 3, pp. 311–319, 1988.
[42] W. S. Cleveland and M. E. McGill, *Dynamic Graphics for Statistics*. Belmont, CA, USA: Wadsworth & Brooks/Cole, 1988.
[43] M. P. Peterson, *Interactive and Animated Cartography*. Englewood Cliffs, NJ, USA: Prentice Hall, 1995.
[44] J. W. Crampton, "Interactivity types in geographic visualization," *Cartogr. Geogr. Inf. Sci.*, vol. 29, no. 2, pp. 85–98, 2002.
[45] D. Persson, G. Gartner, and M. Buchroithner, "Towards a typology of interactivity functions for visual map exploration," in *Geographic Hypermedia: Concepts and Systems*, E. Stefanakis, M. P. Peterson, C. Armenakis, and V. Delis, Eds. Berlin, Germany: Springer, 2006, pp. 275–292.
[46] M. Harrower and B. Sheesley, "Designing better map interfaces: A framework for panning and zooming," *Trans. GIS*, vol. 9, no. 2, pp. 77–89, 2005.
[47] M.-J. Lobo, E. Pietriga, and C. Appert, "An evaluation of interactive map comparison techniques," in *Proc. 33rd ACM Conf. Hum. Factors Comput. Syst.*, 2015, pp. 3573–3582.
[48] M. L. Gough *et al.*, "Method and apparatus for providing translucent images on a computer display," U.S. Patent 6,072,489, 2000.
[49] S. Oviatt, "Multimodal interactive maps: Designing for human performance," *Hum.-Comput. Interact.*, vol. 12, no. 1, pp. 93–129, 1997.
[50] A. M. Brock, P. Truillet, B. Oriola, D. Picard, and C. Jouffrais, "Interactivity improves usability of geographic maps for visually impaired people," *Hum.-Comput. Interact.*, vol. 30, no. 2, pp. 156–194, 2015.
[51] M. J. Knight and M. Tlauka, "Interactivity in map learning: The effect of cognitive load," *Spatial Cogn. Comput.*, vol. 17, no. 3, pp. 185–198, 2017.
[52] A. L. Alencar de Mendonça and C. C. Lyra, "Does interactivity increase map user's reasoning? A study with environmental map users," *Abstr. Int. Cartograph. Assoc.*, vol. 1, pp. 1–2, 2019.
[53] M. Keehner, M. Hegarty, C. Cohen, P. Khooshabeh, and D. R. Montello, "Spatial reasoning with external visualizations: What matters is what you see, not whether you interact," *Cogn. Sci.*, vol. 32, no. 7, pp. 1099–1132, 2008.
[54] E. Raisz, *General Cartography*. New York, NY, USA: McGraw-Hill, 1938.
[55] D. A. Smith, "Online interactive thematic mapping: Applications and techniques for socio-economic research," *Comput. Environ. Urban Syst.*, vol. 57, pp. 106–117, 2016.
[56] Z. Li and M. Wong, "Animating basic operations for digital map generalization with morphing techniques," in *Proc. Int. Arch. Photogrammetry Remote Sens. Spatial Inf. Sci.*, 2008, pp. 637–642.
[57] D. N. Pantazis, B. Karathanasis, M. Kassoli, and A. Koukofikis, "Are the morphing techniques useful for cartographic generalization?," in *Urban and Regional Data Management*, A. Krek, M. Rumor, S. Zlatanova, and E. M. Fendel, Eds. London, U.K.: CRC Press, 2009, pp. 207–216.
[58] M.-J. Lobo, C. Appert, and E. Pietriga, "Animation plans for before-and-after satellite images," *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 2, pp. 1347–1360, Feb. 2019.
[59] D. F. Reilly and K. M. Inkpen, "Map morphing: Making sense of incongruent maps," in *Proc. Graph. Interface*, 2004, pp. 231–238.
[60] M. S. Monmonier, *Technological Transition in Cartography*. Madison, WI, USA: Univ. of Wisconsin Press, 1985.
[61] C. S. Campbell and S. L. Egbert, "Animated cartography: Thirty years of scratching the surface," *Cartographica*, vol. 27, no. 2, pp. 24–46, 1990.
[62] D. DiBiase, A. M. MacEachren, J. B. Krygier, and C. Reeves, "Animation and the role of map design in scientific visualization," *Cartogr. Geogr. Inf. Syst.*, vol. 19, no. 4, pp. 201–214, 1992.
[63] D. Dorling, "Stretching space and splicing time: From cartographic animation to interactive visualization," *Cartogr. Geogr. Inf. Syst.*, vol. 19, no. 4, pp. 215–227, 1992.
[64] C. M. Newton, "Graphics: From alpha to omega in data analysis," in *Graphical Representation of Multivariate Data*, P. C. C. Wang, Ed. New York, NY, USA: Academic, 1978, pp. 59–92.
[65] R. A. Becker and W. S. Cleveland, "Brushing scatterplots," *Technometrics*, vol. 29, no. 2, pp. 127–142, 1987.
[66] M. Monmonier, "Geographic brushing: Enhancing exploratory analysis of the scatterplot matrix," *Geographical Anal.*, vol. 21, no. 1, pp. 81–84, 1989.

[67] G. L. Andrienko and N. V. Andrienko, "Interactive maps for visual data exploration," *Int. J. Geographical Inf. Sci.*, vol. 13, no. 4, pp. 355–374, 1999.
[68] J. A. Dykes, "Exploring spatial data representation with dynamic graphics," *Comput. Geosci.*, vol. 23, no. 4, pp. 345–370, 1997.
[69] R. Haining, *Spatial Data Analysis: Theory and Practice*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 2003.
[70] J. Dykes, "Cartographic visualization: Exploratory spatial data analysis with local indicators of spatial association using TcI/Tk and cdv," *J. Roy. Statist. Soc. Ser. Statistician*, vol. 47, no. 3, pp. 485–497, 1998.
[71] F. Hardisty and A. C. Robinson, "The GeoViz toolkit: Using component-oriented coordination methods for geographic visualization and analysis," *Int. J. Geographical Inf. Sci.*, vol. 25, no. 2, pp. 191–210, 2011.
[72] R. Mundigl, "Cartograms in Tableau," 2015. Accessed: Jul. 26, 2019. [Online]. Available: https://www.clearlyandsimply.com/clearly_and_simply/2015/04/cartograms-in-tableau.html
[73] S. Nusrat and S. Kobourov, "The state of the art in cartograms," *Comput. Graph. Forum*, vol. 35, no. 3, pp. 619–642, 2016.
[74] Office for National Statistics, "Visualising your constituency," 2015. Accessed: Jul. 28, 2019. [Online]. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles/visualisingyourconstituency/2015-03-26
[75] D. Spruijt, "BouncyMaps," 2017. Accessed: Jul. 28, 2019. [Online]. Available: https://www.bouncymaps.com/#!/bouncymaps/WORLD/-2102779806
[76] F. Heidmann, F. Herrmann, and M. Peissner, "Interactive maps on mobile, location-based systems: Design solutions and usability testing," in *Proc. 21st Int. Cartograph. Conf.*, 2003, pp. 1299–1305.
[77] R. Brath and E. Banissi, "Bertin's forgotten typographic variables and new typographic visualization," *Cartogr. Geogr. Inf. Sci.*, vol. 46, pp. 119–139, 2019.
[78] R. W. Proctor and K.-P. L. Vu, "Human information processing: An overview for human-computer interaction," in *The Human-Computer Interaction Handbook*, J. A. Jacko and A. Sears, Eds. Hillsdale, MI, USA: L. Erlbaum, 2003, pp. 35–51.
[79] S. Wehrend, "Appendix B: Taxonomy of visualization goals," in *Visual Cues: Practical Data Visualization*, P. R. Keller and M. M. Keller, Eds. Los Alamitos, CA, USA: IEEE Comput. Soc. Press, 1993, pp. 187–199.
[80] M. X. Zhou and S. K. Feiner, "Visual task characterization for automated visual discourse synthesis," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, 1998, pp. 392–399.
[81] R. Amar, J. Eagan, and J. Stasko, "Low-level components of analytic activity in information visualization," in *Proc. IEEE Symp. Inf. Vis.*, 2005, pp. 111–117.
[82] R. E. Roth, "Cartographic interaction primitives: Framework and synthesis," *Cartograph. J.*, vol. 49, no. 4, pp. 376–395, 2012.
[83] R. E. Roth, "An empirically-derived taxonomy of interaction primitives for interactive cartography and geovisualization," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2356–2365, Dec. 2013.
[84] N. Andrienko and G. Andrienko, *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Berlin, Germany: Springer-Verlag, 2006.
[85] R. Beecham, J. Dykes, W. Meulemans, A. Slingsby, C. Turkay, and J. Wood, "Map LineUps: Effects of spatial structure on graphical inference," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 391–400, Jan. 2017.
[86] M. T. Gastner, V. Seguy, and P. More, "Fast flow-based algorithm for creating density-equalizing map projections," *Proc. Nat. Acad. Sci. USA*, vol. 115, no. 10, pp. E2156–E2164, 2018.
[87] G. Cumming, *Understanding the New Statistics*. New York, NY, USA: Routledge, 2011.
[88] M. P. Fay, "Two-sided exact tests and matching confidence intervals for discrete data," *R J.*, vol. 2, no. 1, pp. 53–58, 2010.
[89] M. Hollander, D. A. Wolfe, and E. Chicken, *Nonparametric Statistical Methods*, 3rd ed. New York, NY, USA: Wiley, 2013.
[90] I. K. Duncan, S. Tingsheng, S. T. Perrault, and M. T. Gastner, "Mgastner/effectiveness_of_interactive_cartograms: Data and R scripts," May 2020. [Online]. Available: https://zenodo.org/record/4064663
[91] G. V. R. J. S. Prasad and A. Ojha, "Text, table and graph: Which is faster and more accurate to understand?," in *Proc. IEEE 4th Int. Conf. Technol. Educ.*, 2012, pp. 126–131.
[92] M. De Berg, E. Mumford, and B. Speckmann, "Optimal BSPs and rectilinear cartograms," *Int. J. Comput. Geometry Appl.*, vol. 20, no. 02, pp. 203–222, 2010.
[93] C. A. Brewer, "ColorBrewer: Color advice for maps," 2002. Accessed: Apr. 25, 2020. [Online]. Available: https://colorbrewer2.org/#
[94] M. T. Gastner, C. R. Shalizi, and M. E. J. Newman, "Maps and cartograms of the 2004 US presidential election results," *Advances Complex Syst.*, vol. 8, no. 1, pp. 117–123, 2005.
[95] S. Nusrat, M. J. Alam, C. Scheidegger, and S. Kobourov, "Cartogram visualization for bivariate geo-statistical data," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 10, pp. 2675–2688, Oct. 2018.
[96] C. Fish, K. P. Goldsberry, and S. Battersby, "Change blindness in animated choropleth maps: An empirical study," *Cartogr. Geogr. Inf. Sci.*, vol. 38, no. 4, pp. 350–362, 2011.
[97] S. Tingsheng, I. K. Duncan, and M. T. Gastner, "go-cart.io: A web application for generating contiguous cartograms," *Abstr. Int. Cartograph. Assoc.*, vol. 1, 2019, Art. no. 333.
[98] S. Tingsheng, I. K. Duncan, Y.-N. Chang, and M. T. Gastner, "Motivating good practices for the creation of contiguous area cartograms," in *Proc. 8th Int. Conf. Cartography GIS*, 2020, pp. 589–598.
[99] Observable, Inc., "The magic notebook for exploring data / Observable.," 2017. Accessed: May 15, 2020. [Online]. Available: https://observablehq.com/
[100] A. Satyanarayan, D. Moritz, K. Wongsuphasawat, and J. Heer, "Vega-Lite: A Grammar of interactive graphics," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 341–350, Jan. 2017.

Ian K. Duncan is currently working toward the BSc degree in mathematical, computational, and statistical sciences with a minor in philosophy at Yale-NUS College, Singapore.

Shi Tingsheng received the BSc degree in mathematical, computational, and statistical sciences from Yale-NUS College, Singapore, in 2020. He is currently a software engineer with ByteDance Ltd.

Simon T. Perrault received the PhD degree in computer science from Telecom ParisTech, Paris, France, in 2013. He is currently an assistant professor in information systems technology and design with the Singapore University of Technology and Design (SUTD). Previously, he was a visiting professor with the Korean Advanced Institute of Science and Technology, Korea, and assistant professor with Yale-NUS College, Singapore.

Michael T. Gastner received the PhD degree in physics from the University of Michigan, Ann Arbor, Michigan, in 2005. He held postdoctoral positions with the Santa Fe Institute, University of Oldenburg, and Imperial College London. He is currently an assistant professor of mathematics, computational, and statistical sciences with Yale-NUS College, Singapore. Before moving to Singapore, he was a faculty member with the University of Bristol and the Hungarian Academy of Sciences.