

Article

# Using Eye Tracking to Evaluate the Usability of Flow Maps

Weihua Dong <sup>1,\*</sup>, Shengkai Wang <sup>1,\*</sup> , Yizhou Chen <sup>1</sup>  and Liqui Meng <sup>2</sup>

<sup>1</sup> State Key Laboratory of Remote Sensing Science, Beijing Key Laboratory for Remote Sensing of Environment and Digital Cities, Research Center of Geospatial Cognition and Visual Analytics, and Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China; yizhouch19@hotmail.com

<sup>2</sup> Chair of Cartography, Technical University of Munich, 80333 Munich, Germany; liqui.meng@tum.de

\* Correspondence: dongweihua@bnu.edu.cn (W.D.); wangsk@mail.bnu.edu.cn (S.W.); Tel.: +86-10-5880-9246 (W.D.)

Received: 26 May 2018; Accepted: 19 July 2018; Published: 21 July 2018



**Abstract:** Flow maps allow users to perceive not only the location where interactions take place, but also the direction and volume of events. Previous studies have proposed numerous methods to produce flow maps. However, how to evaluate the usability of flow maps has not been well documented. In this study, we combined eye-tracking and questionnaire methods to evaluate the usability of flow maps through comparisons between (a) straight lines and curves and (b) line thicknesses and color gradients. The results show that curved flows are more effective than straight flows. Maps with curved flows have more correct answers, fixations, and percentages of fixations in areas of interest. Furthermore, we find that the curved flows require longer finish times but exhibit smaller times to first fixation than straight flows. In addition, we find that using color gradients to indicate the flow volume is significantly more effective than the application of different line thicknesses, which is mainly reflected by the presence of more correct answers in the color-gradient group. These empirical studies could help improve the usability of flow maps employed to visualize geo-data.

**Keywords:** flow map; map usability; visual variables; visualization; eye tracking

## 1. Introduction

Flow maps are effective tools to represent connections and interactions volume between geographical regions [1], and are widely applied to research related to spatial trajectory and interactions, such as migration [2], transportation [3], population movement [4], disease spread [5], and social communication flows [6]. Researchers use various methods to generate flow, such as force-directed edges [6,7], bundled edges [8], and 3D curves [9]. Current theories and techniques related to geographical information systems (GIS) to produce flow maps are still immature [10]. Researchers have been attempting to solve the cartographical problems that influence human cognition, such as overlapping or intersecting symbols and inappropriate color and size of symbols [7,11]. Therefore, it is important to evaluate and improve the usability of flow maps, which can allow users to perform certain tasks more accurately and rapidly.

The usability of flow maps influences the effectiveness and efficiency of map users processing and interpreting cartographical flow information. Shape, color, and size are three important visual variables in cartography [12,13]. Previous cartographical studies have demonstrated the influence on effectiveness and efficiency in maps caused by shape, color, and size [14–16]. Researchers have also conducted experiments to evaluate the representation of these three visual variables in flow

maps [1,11,17,18]. We believe that these three visual variables have significant influences on the usability of flow maps. However, these studies did not reach consistent conclusions on the effectiveness and efficiency of flow shapes. In addition, the differences in usability between using line thicknesses and color gradients to represent flow volume remain to be studied.

In this study, we used eye-tracking and questionnaire methods to evaluate the usability of flow maps. We mainly focused on comparisons between (a) straight lines and curves to indicate flows and (b) line thicknesses and color gradients to indicate the flow volume. We also explored how users react to flow maps with different combinations of visual variables (a) and (b). This study will examine flow maps as research objects from the perspective of user cognition for the map visual variables, and it will probe into the flow map design principles through the process of visualization via an eye tracking experiment. This study seeks to answer questions regarding how visual variables for flow maps affect people's understanding of geographic information and how the usability of flow maps should be evaluated. Specifically, we address the following questions:

- To indicate flows, do straight lines and curves influence the usability of a flow map?
- To represent flow volume, do line thicknesses and color gradients have different impacts on the usability of a flow map?

The remainder of this article mainly includes five sections. In Section 2, the research background of flow maps and usability studies using eye tracking are reviewed. In Section 3, we describe the design of the experiment. Section 4 presents the analysis indices and results of our experiment. In Section 5, the experimental results are analyzed and discussed regarding the influences of line shape, line thickness, and line color gradient on the usability of flow maps. Finally, in Section 6, we make a brief conclusion and provide suggestions for the future.

## 2. Related Work

Imhof [19] suggested the use of multiple parallel lines or icons in addition to lines to represent the data type, volume, and flow speed. Dent, Torguson and Hodler [20] noted that the design principles used to create flow maps include the placement of small flows over larger flows if overlapping is inevitable, the use of arrows to indicate the flow direction, the application of varying line thicknesses to show different data quantities, and the drawing of arrows proportionally to the line thicknesses. The two aforementioned studies both employed curves in the map to represent the data flow, but they did not prove that a curve is more effective than a straight line. Scholars have produced different findings with regard to this issue. Xu et al. [21] conducted a user study of curved edges in graph visualization. They asked participants to complete network tasks, including the determination of connectivity, shortest path, node degree, and common neighbors, and to provide subjective ratings of the aesthetics of different edge types. They found that users perform more efficiently and accurately with maps containing straight lines than with maps possessing curves. Meanwhile, Purchase et al. [22] also presented tasks, including the shortest path, vertex degree, and common neighbor. They discovered that users performed better with straight lines, although they preferred curved graphs. However, Jenny et al. [1] found that curved flows are more effective than straight flows, arrows indicate direction more effectively than tapered line thicknesses, and flows between nodes are more effective than flows between areas. Furthermore, through a literature review, experimental data analysis, and questionnaire survey methods, they derived several design principles about flow maps, including reducing curve overlap, avoiding excessive curvature, and avoiding curves of unconnected nodes. Studies on the usability of flow maps have mostly utilized indices of the accuracy, task time, or user preference; however, few studies have involved the users' visual perception while reading a flow map.

Eye tracking provides a great deal of assistance for cognitive research on cartography [23] by recording real-time fixation, saccade, and duration data, and analyzing eye movement behavior to research visual information processing and to guide map design endeavors [24,25]. Based on the eye tracking method, many scholars have conducted thorough research on map interface design [26,27],

in addition to visual query strategies and map interpretation [28,29]. Dong [30] found that enhanced remote sensing images have a higher usability than the original images based on eye tracking data and that the effects of size, color, and frequency are related to the display resolution in dynamic maps. Popelka and Brychtova [31] compared the human perception of contour lines and 3D terrain, and they found that users have different strategies of cognition according to scan path. H Liao et al. [32] studied the influence of label density on maps, and they found that both response time for visual search tasks and visual complexity are positively correlated with label density. Therefore, exploring flow maps of geo-data and the differences among the visual variables based on eye-tracking technology is feasible, which may induce cognition patterns in users reading flow maps. Recently, eye tracking has been adopted to evaluate the usability of different types of maps, such as WebMaps [33], interactive maps [27], and animated maps [34], which can provide more direct cues to users' visual cognition of maps and has potential for evaluating the usability of flow maps.

However, few studies have employed eye-tracking to evaluate the usability of flow maps. Hsin-Yang Ho, et al. [35] conducted an eye tracking experiment to explore the influences of five different two-dimensional flow visualization methods on the visual perception of the user. They obtained several new findings that were not observed in previous studies. For example, they found that users are more likely to focus on key points or on highly variable regions and that the average gaze distance can be utilized to help find differences in the performances among different flow visualization methods. Blascheck et al. [36] proposed an approach to explore eye movements in node-link graphs with different layouts, and to generate and adjust node-link graphs with eye movements data. Netzel et al. [37] used eye tracking to compare the visualization of node-link trajectories between different ways to represent link types, nodes, and depth sorting for overlaps. However, they did not test the influence of the shape, color, and size of flows.

### 3. Methods

#### 3.1. Experimental Design

We performed an indoor eye tracking experiment using a desktop eye tracker to record the eye movement data of the participants. During the experiment, the participants were placed in a silent, distraction-free environment. They were required to read the provided flow map that appeared on the screen and subsequently answer the preset questions. To ensure that the participants maintained a relaxed state in consideration of the task duration, we paused the process between every two maps to allow the participants to rest, during which time they were able to record their answers. In addition, to minimize the impact of nontask-related actions on the eye movements of the participants, a simple human-computer interaction process consisting of a single mouse click for switching the map was designed. Finally, we evaluated the impacts of shape, size (line thickness), and color (color gradient) on visualizing the flow on user cognition through the results of the tasks and eye movement data.

In addition to the eye tracking experiment, we conducted a test of participants using questionnaires with identical flow maps. We required the participants to give scores for two parts: (a) the visual quality of the connections between nodes in maps with straight lines and curves (b) and the visual quality of the flow volume between nodes in maps with different line thicknesses or color gradients. A five-grade marking system was used: 1 = very unclear, 2 = unclear, 3 = normal, 4 = clear, and 5 = very clear. In part 1, participants were asked to evaluate from several perspectives, including the simplicity of judging the flow directions, the number of overlapping or intersecting features, and the ability to express details. In part 2, they were asked to evaluate the ability to discern the difference in flow quantities.

#### 3.2. Participants

In total, the experiment involved 40 participants and the questionnaire test involved another 42 participants. All the participants were between 20 and 23 years old, and they were undergraduates

in the Faculty of Geographical Science, Beijing Normal University. They were assigned to two groups to accomplish different tasks. Each group size ( $N = 20$ ) was consistent with those of other eye tracking cartography studies [26,27,30]. None of the participants possessed an eye disease or color vision deficiency. It is noted that the participants have homogeneity due to the limitation of available sources of participants. The homogeneity may influence universality of experimental results, but it will also increase the statistical effectiveness. The selection of participants is considered to be acceptable because our study focuses more on how different visual variables in flow maps influence users' performance, rather than how different users perform when reading flow maps.

### 3.3. Apparatus

This experiment used a Tobii T120 Eye Tracker ([www.tobii.com](http://www.tobii.com)) with a sampling frequency of 60 Hz. The eye tracker was connected to a 17-inch monitor with a screen resolution of  $1280 \times 1024$ . Tobii Studio 3.2, the corresponding software, was installed on a Lenovo PC.

### 3.4. Materials

In this paper, we used four mobile communication datasets collected from Northeast China as the origin data for the flow maps. The origin data were provided from a communication operator, and all data that may have included private information were encrypted. We used straight lines and arc curves to represent the flow of data. In the material flow maps, different line thicknesses and color gradients were used to indicate the flow volume, and their effects on the usability of the maps were compared. Line thicknesses and color gradients were both divided into five levels to represent the flow volume. We chose a red gradient for the color aspect and divided it into five levels. It should be noted that the base map had a gray tone; therefore, to avoid color confusion, we chose a red gradient rather than a gray gradient. Linear features were generated based on the abovementioned method and were then symbolized using line thicknesses and color gradients, after which we were finally able to complete the visualization of the flow of geo-data. Figure 1 shows samples of the experimental materials.

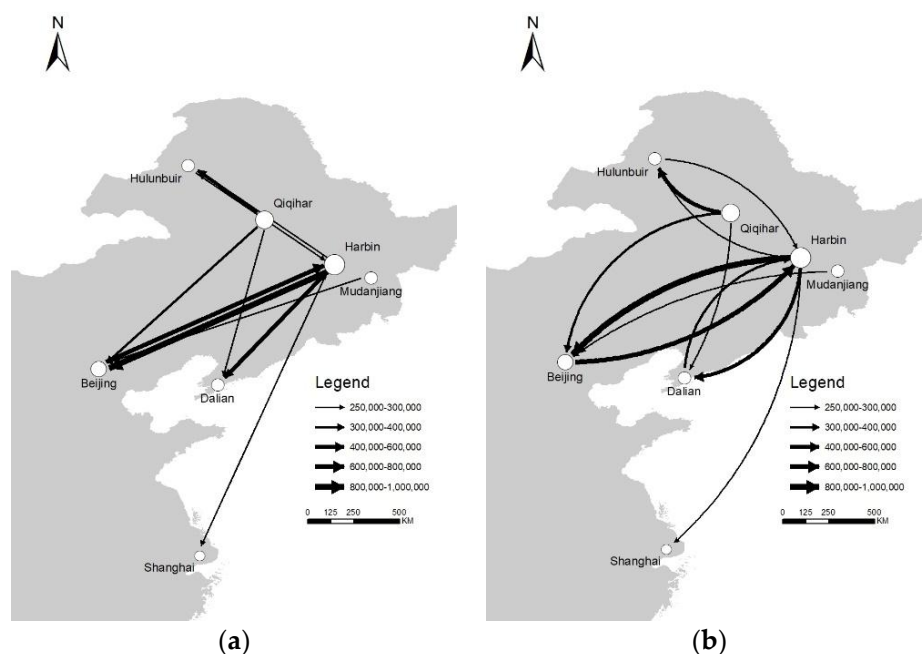
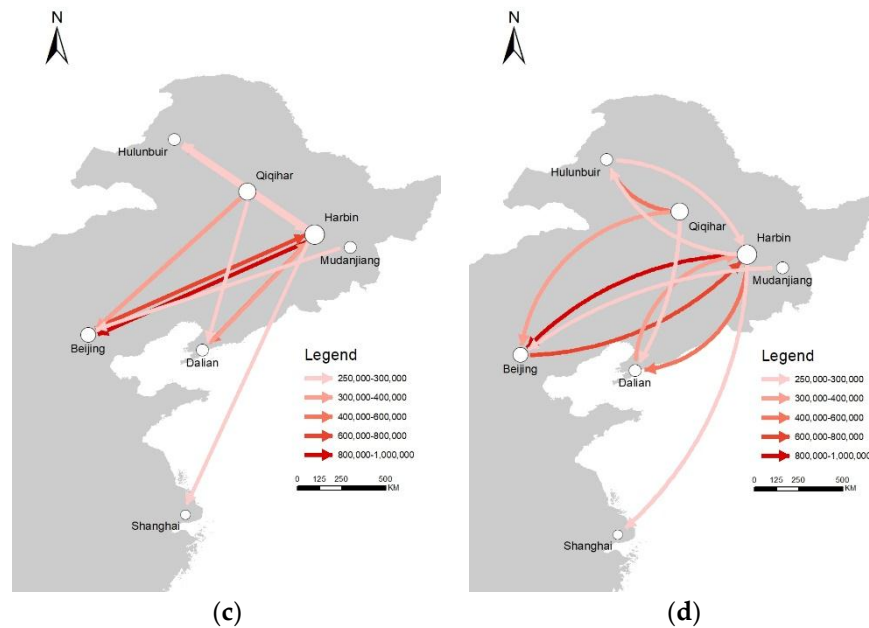


Figure 1. Cont.



**Figure 1.** Samples of the experimental materials generated with dataset 1 (Table 1). The cities shown in the maps include Hulunbuir, Qiqihar, Harbin, Mudanjiang, Beijing, Dalian, and Shanghai (not all cities that appear in the four datasets are displayed in the sample maps). *SL* = straight line; *CV* = curve; *LT* = line thickness; and *CG* = color gradient. (a) *SL* + *LT*; (b) *CV* + *LT*; (c) *SL* + *CG*; and (d) *CV* + *CG*.

Normally, point size represents quantized attributes of locations in flow maps [1,26,27,30]. In material maps, point features that represent cities do not have the same size because the size variable is used to represent the total volume of flow in each city. In our experiment, we do not focus on the effects on the usability of flow maps caused by point size. To reduce the point size influence on the results of the experiment, the meaning of point size is not displayed in the legends and was presented to participants before the experiment, and the difference in point size did not help participants to finish preset tasks.

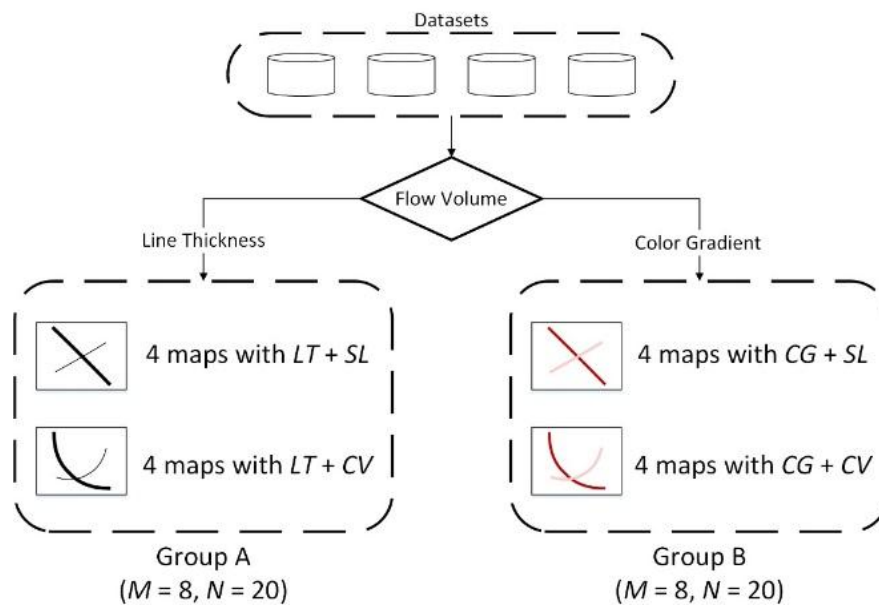
### 3.5. Procedure

As shown in Figure 2, the participants were equally divided into two groups. For group A, line thicknesses were used to indicate the flow volume of maps. For group B, color gradients were used to indicate the flow volume of maps. We used four basic flow datasets to generate material flow maps. Each group had four maps using straight lines and four maps using curves.

First, we introduced the instructions and procedure of the experiment to the participants, who were then given 1 to 3 min to familiarize themselves with the functions and operations of the interface.

Then, the participants were asked to read the maps shown on the screen and complete the given tasks. The main tasks in the experiment were to search for a target feature and to compare the flow volume. Participants were asked one question before each map was displayed, and they needed to give the answer after reading each map. Both group A and B used the same order of maps and the same questions for each map. To offset the learning effect of the participants, we allowed half the participants to read the maps in a normal order and the other half to read the maps in reverse order in each group. All questions were based on the three questions shown below. The order of maps and settings of relevant questions are shown in Table 1.

- Q1: Among the outflows from city A, to which city does the flow have the largest (smallest) volume?
- Q2: Among the inflows to city A, from which city does the flow have the largest (smallest) volume?
- Q3: The largest (smallest) data flow occurs between which two cities?



**Figure 2.** Group setting and material quantities in the experiment. *SL* = straight line; *CV* = curve; *LT* = line thickness; and *CG* = color gradient; *M* is the number of material flow maps; *N* is the number of participants.

**Table 1.** Map order and question setting for materials in both groups.

	Line Shape	Material (Order)	Dataset	Question	Variable 1	Variable 2
Group A & B	Straight line	Map 1	D1	Q2	Dalian	Smallest
		Map 2	D2	Q1	Harbin	Smallest
		Map 3	D3	Q2	Harbin	Largest
		Map 4	D4	Q3	-	Smallest
	Curve	Map 5	D1	Q2	Beijing	Smallest
		Map 6	D2	Q2	Tianjin	Largest
		Map 7	D3	Q1	Suihua	Largest
		Map 8	D4	Q1	Dalian	Smallest

Variable 1 represents which city was chosen to be city A; Variable 2 represents whether the largest or smallest flow needed to be found.

Notably, using straight lines to represent flows in maps will cause overlaps of flow and make it difficult for users to distinguish flows. This problem also existed in our material flow maps. To ensure that participants were able to complete preset tasks, we purposely adjusted the questions to avoid these overlapping flows as correct answers and tried to maintain the difficulty of the questions. For example, we selected a proper city as city A, such as Dalian in Figure 1, and chose a more visually salient flow from the largest flow and the smallest flow as the answer to the question.

The eye movement data and response accuracies were recorded for further analysis. The experiment duration for each participant was 10 to 15 min. A participant sampling rate exceeding 80 percent was considered acceptable for data analysis. After the eye tracking experiment, we conducted a test using questionnaires and asked the participants to score the material flow maps.

### 3.6. Analysis Indices

The analysis of the eye tracking data included effectiveness and efficiency indices. The effectiveness indices included the fixation count, percentage of fixations in area of interests (AOIs), and accuracy, while the efficiency indices included the finish time and time to first fixation. Their interpretations are listed in Table 2.



**Table 2.** Analysis indices and interpretations.

	Indices	Description
Task indices	Accuracy (AC)	Number of task questions answered correctly
	Finish time (FT, s)	Average time used to complete each task (seconds)
Statistical eye tracking indices	Fixation count (FC)	Average fixation counts during each task
	Percentage of fixations in AOIs (PFA)	Number of fixations located within AOIs divided by all fixations
	Time to first fixation (TtFF, s)	Duration from the beginning of the task to the first fixation located within the AOIs (seconds)

Five indices (Table 2) were selected to analyze the efficiency and effectiveness of users' performance when using flow maps to finish preset tasks. Two task indices (Table 2) directly measured users' performance: (1) accuracy (correctness), which reflects the effectiveness of users' performance, where a high accuracy indicates a high effectiveness; and (2) finish time, which reflects the efficiency of users' performance, where a short time to finish all tasks indicates a high efficiency.

Three statistical eye-tracking indices (Table 2) were selected to measure user gaze behaviors statistically: (1) The first was fixation count (number of fixations), which indicates users' efforts in processing information during reading maps. More fixations means users are distracted with nontarget areas when searching for targets [38]. A poorly designed flow map may hamper users' searching efficiency and lead to more fixations; (2) the second was the percentage of fixations in AOIs (the ratio of on target), which reflects users' attention on target AOIs and indicates the searching efficiency for targets [38]. Fixation count is correlative with the number of components that the user is required to process. If the numbers of components in two maps are quite different, then the percentage of fixations in AOIs is the better metric for indicating searching efficiency. A low percentage of fixations in AOIs indicates a low searching efficiency; (3) the third index was the time to first fixation, which indicates how long the users need to identify targets from a complex map [39]. A well designed flow map should quickly guide users' visual attention.

We used independent-sample Mann-Whitney U tests to analyze the differences between the abovementioned five indices of the effectiveness and efficiency between straight lines and curves and between line thicknesses and color gradients. To process the percentage of fixations in AOIs and time to first fixation, we defined AOIs as flows that match the correct answers for each map's question and constructed them in Tobii Studio. Each AOI was constructed using a buffer area of original flow due to the inaccuracy of the eye tracker. The maximum deviation from a gaze point that eye trackers can locate is  $0.5^\circ$ , which is approximately equivalent to 25 pixels on the screen of the eye trackers. Thus, we chose 25 pixels as the buffer distance. Figure A1 shows samples of the experimental materials covered by AOIs. We used gaze opacity maps to show where and for how long users fixated. These gaze opacity maps were generated in Tobii Studio from the fixation count of all participants with a radius of 50 pixels. In addition, we used Pearson chi-squared tests to analyze the data of the questionnaires, to test the difference between straight lines and curves, and between line thicknesses and color gradients.

## 4. Results

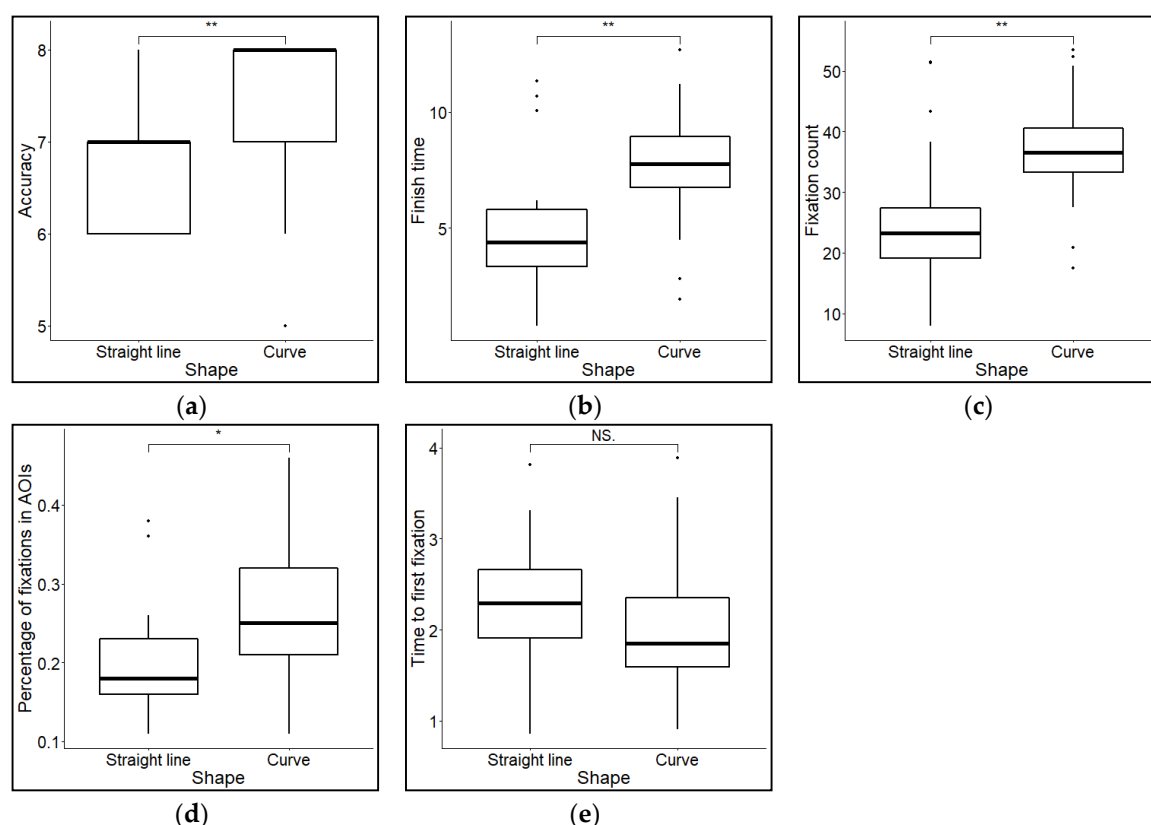
### 4.1. Straight Lines and Curves

With regard to the task indices, as shown in Table 3 and Figure 3, the accuracy for the maps with straight lines ( $M = 6.76$ ,  $SD = 0.62$ ) was significantly smaller than that for the maps with curves ( $M = 7.43$ ,  $SD = 0.81$ ,  $U = -3.085$ ,  $p = 0.002 < 0.01$ ). The finish time for the maps with straight lines ( $M = 5.10$  s,  $SD = 2.68$ ) was significantly shorter than that for the maps with curves ( $M = 7.58$  s,  $SD = 2.56$ ,  $U = -2.981$ ,  $p = 0.003 < 0.01$ ).

**Table 3.** Descriptive statistics for the different linear features.

	Descriptive		Inferential	
	Straight Line	Curve	Mann-Whitney U Test	
	$M \pm SD$	$M \pm SD$	$U$	$p$
Accuracy	$6.76 \pm 0.62$	$7.43 \pm 0.81$	−3.085	0.002 **
Finish time	$5.10 \pm 2.68$	$7.58 \pm 2.56$	−2.981	0.003 **
Fixation count	$25.99 \pm 11.31$	$37.47 \pm 9.36$	−3.283	0.001 **
Percentage of fixations in AOIs	$0.20 \pm 0.07$	$0.26 \pm 0.09$	−2.428	0.015 *
Time to first fixation	$2.27 \pm 0.69$	$2.07 \pm 0.84$	1.094	0.274

$M$  = mean,  $SD$  = standard deviation, \*  $p < 0.05$  and \*\*  $p < 0.01$ .



**Figure 3.** Statistics for the different line shapes and Mann-Whitney U test results (\*  $p < 0.05$ , \*\*  $p < 0.01$  and NS. = not significant). (a) Accuracy; (b) finish time; (c) fixation count; (d) percentage of fixations in AOIs; and (e) time to first fixation.

With regard to the statistical eye tracking indices, the fixation count for the maps with straight lines ( $M = 25.99$ ,  $SD = 11.31$ ) was significantly less than that for the maps with curves ( $M = 37.47$ ,  $SD = 9.36$ ,  $U = -3.283$ ,  $p = 0.001 < 0.01$ ). The percentage of fixations in AOIs for the maps with straight lines ( $M = 0.20$ ,  $SD = 0.07$ ) was significantly smaller than that for the maps with curves ( $M = 0.26$ ,  $SD = 0.09$ ,  $U = -2.428$ ,  $p = 0.015 < 0.05$ ). The time to first fixation for the maps with straight lines ( $M = 2.27$  s,  $SD = 0.69$ ) was insignificantly longer than that for the maps with curves ( $M = 2.07$  s,  $SD = 0.84$ ,  $U = 1.094$ ,  $p = 0.274$ ).

To conclude, using curved features to indicate the flow of data in maps was more effective than using straight lines, and thus, the participants could focus more on an effective area within the map to acquire accurate information. In terms of the efficiency, the finish times for the straight-line maps were shorter than those for the curve maps, although the times to first fixation were longer.



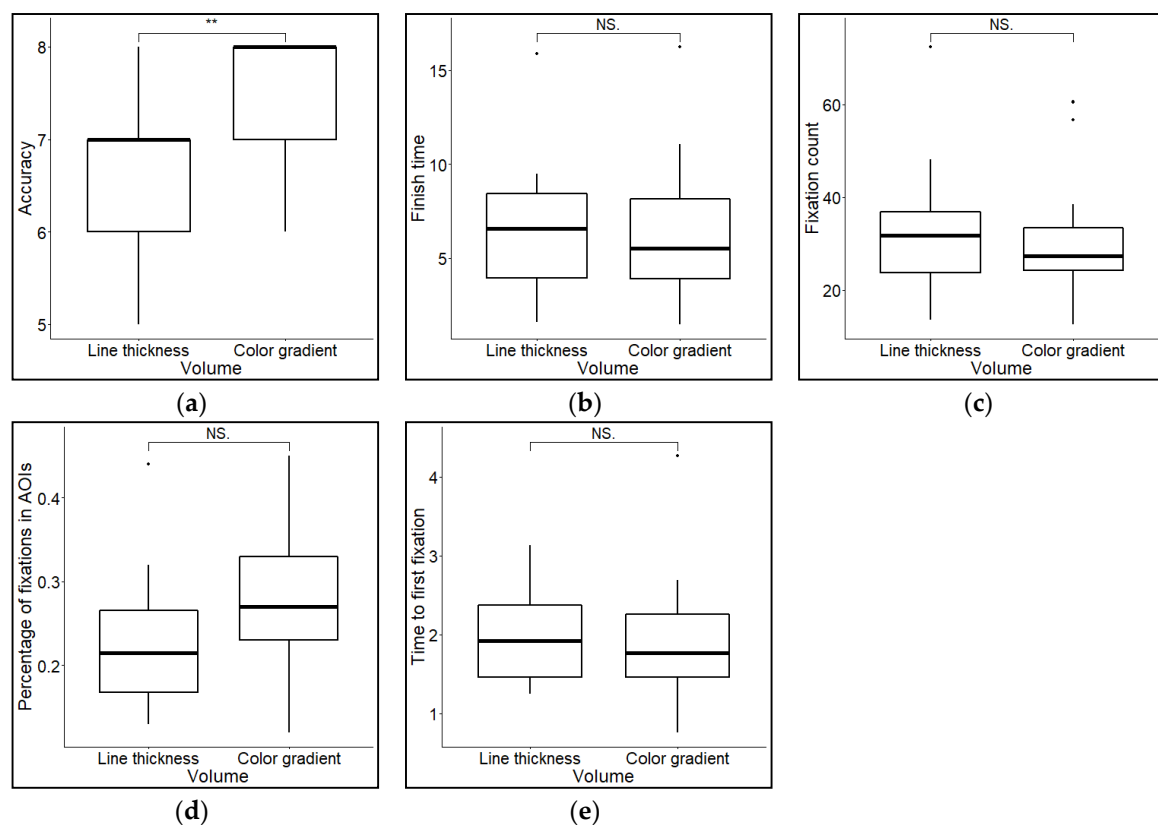
#### 4.2. Line Thicknesses and Color Gradients

The experimental results (Table 4 and Figure 4) show that with regard to task indices, the accuracy for the line-thickness maps ( $M = 6.75$ ,  $SD = 0.71$ ) was significantly smaller than that for the color-gradient maps ( $M = 7.52$ ,  $SD = 0.60$ ,  $U = -3.341$ ,  $p = 0.001 < 0.05$ ). The difference in the finish times between the maps with different line thicknesses ( $M = 6.50$  s,  $SD = 3.30$ ) and the color-gradient maps ( $M = 6.35$  s,  $SD = 3.46$ ,  $U = 0.183$ ,  $p = 0.855$ ) was insignificant.

**Table 4.** Descriptive and inferential statistics for line thicknesses and color gradients.

	Descriptive		Inferential	
	Line Thickness	Color Gradient	Mann-Whitney U Test	
	$M \pm SD$	$M \pm SD$	$U$	$p$
Accuracy	$6.75 \pm 0.71$	$7.52 \pm 0.60$	$-3.341$	$0.001^{**}$
Finish time	$6.50 \pm 3.30$	$6.35 \pm 3.46$	$0.183$	$0.855$
Fixation count	$32.74 \pm 12.89$	$31.28 \pm 13.37$	$0.574$	$0.566$
Percentage of fixations in AOIs	$0.23 \pm 0.07$	$0.28 \pm 0.08$	$-1.956$	$0.050$
Time to first fixation	$1.99 \pm 0.56$	$1.96 \pm 0.73$	$0.365$	$0.715$

$M$  = mean,  $SD$  = standard deviation, and  $^{**} p < 0.01$ .



**Figure 4.** Statistics for the two different representations of the flow volume and Mann-Whitney U test results ( $^{**} p < 0.01$  and NS = not significant). (a) Accuracy; (b) finish time; (c) fixation count; (d) percentage of fixations in AOIs; and (e) time to first fixation.

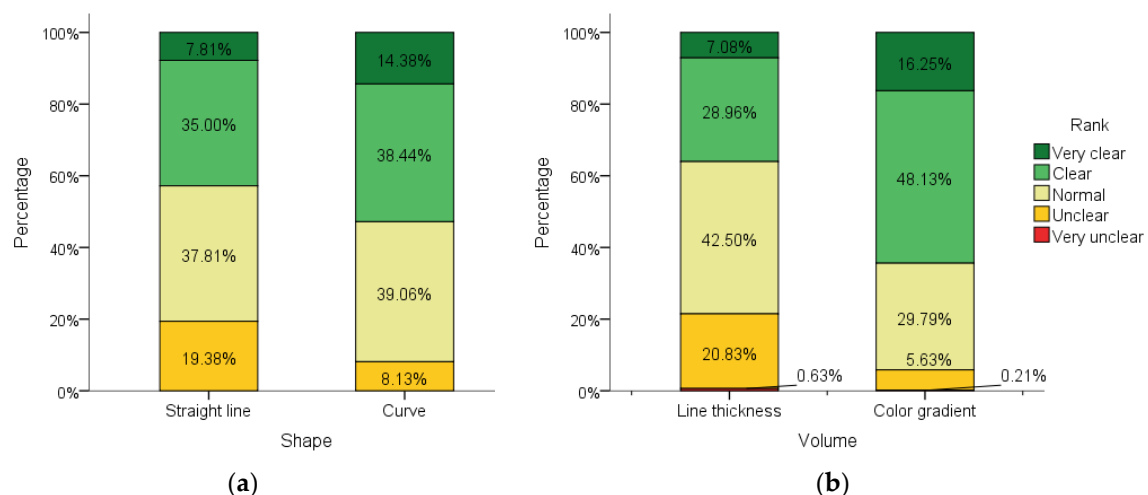
With regard to the statistical eye-tracking indices, the difference in the fixation counts between the maps using different line thicknesses ( $M = 32.74$ ,  $SD = 12.89$ ) and the maps using color gradients ( $M = 31.28$ ,  $SD = 13.37$ ,  $U = 0.574$ ,  $p = 0.566$ ) was insignificant. The percentage of fixations in AOIs for the maps using different line thicknesses ( $M = 0.23$ ,  $SD = 0.07$ ) was lower than that for the maps using

color gradients ( $M = 0.28$ ,  $SD = 0.08$ ,  $U = -1.956$ ,  $p = 0.050$ ), although the difference was insignificant. Similarly, with regard to the time to first fixation, the difference between the line-thickness maps ( $M = 1.99$  s,  $SD = 0.56$ ) and the color-gradient maps ( $M = 1.96$  s,  $SD = 0.73$ ,  $U = 0.365$ ,  $p = 0.715$ ) was also insignificant.

To conclude, the difference between the use of line thicknesses and color gradients is mainly embodied by a change in the map effectiveness. When using color gradients to indicate the flow volume, the percentage of fixations in AOIs and the number of correct answers were significantly larger, indicating that the participants were able to focus more efficiently on target areas with valid information in the maps and acquire correct answers more easily. However, there was no significant difference in the efficiency for map reading and searching tasks between these two mechanisms for indicating the flow volume.

#### 4.3. Questionnaires

The questionnaire results are shown in Figure 5, and the Pearson chi-squared test results are shown in Table 5. In terms of different line shapes (Figure 5a), the difference between maps using straight lines and curves was significant ( $p = 0.000 < 0.01$ ). For straight lines, 19.4% of the participants indicated that the maps were unclear because there were too many overlapping and intersecting lines that made them barely distinguishable. For curves, only 8.1% of participants rated the map as unclear. Neither maps with straight lines nor those with curves were rated as being very unclear. Curves were generally thought to be more complex due to excessive curvature. In addition, maps using the two types of line features received a similar percentage of normal evaluation (straight line = 37.8% and curves = 39.1%). The participants rated these maps as being nearly usable, but with some problems mentioned above. Furthermore, for straight lines, 35.0% and 7.8% rated them as clear and very clear, respectively (42.8% in all). For curves, the percentage was much higher—38.4% and 14.4% rated them as clear and very clear, respectively (52.8% in all). Maps using curves were rated as being clearer than those with straight lines. To conclude, compared to straight lines, curves can effectively avoid overlapping and intersecting and are less likely to make the maps misunderstood.



**Figure 5.** Statistics for questionnaires (SL = straight line, CV = curve, LT = line thickness, and CG = color gradient). (a) Line shape and (b) flow volume.

The difference between maps using line thicknesses and color gradients was also significant ( $p = 0.000 < 0.01$ ). In terms of different methods to visualize flow volumes (Figure 5b), for different line thicknesses, 29.0% and 7.1% rated the maps as clear and very clear (36.1% in all), while for color gradients, 48.1% and 16.3% rated them as clear and very clear (64.4% in all). Most participants indicated that it was much harder to judge the flow volume from maps using different line thicknesses than it

was from maps using color gradients. Thus, using color gradients is a better choice than using different line thicknesses to visualize flow quantities.

**Table 5.** Results of Pearson chi-squared test applied to questionnaires.

Comparison	Value	df	p
Straight line Curve	21.518 <sup>a</sup>	3	0.000 **
Line thickness Color gradient	93.845 <sup>b</sup>	4	0.000 **

<sup>a</sup> 0 cells (0.0%) have an expected count of less than 5. The minimum expected count is 35.50. <sup>b</sup> Two cells (20.0%, line thickness: very unclear and color gradient: very unclear) have an expected count of less than 5. The minimum expected count is 2.00. df = degree of freedom, \*\*  $p < 0.01$ .

## 5. Discussion

### 5.1. Comparison between Straight Lines and Curves

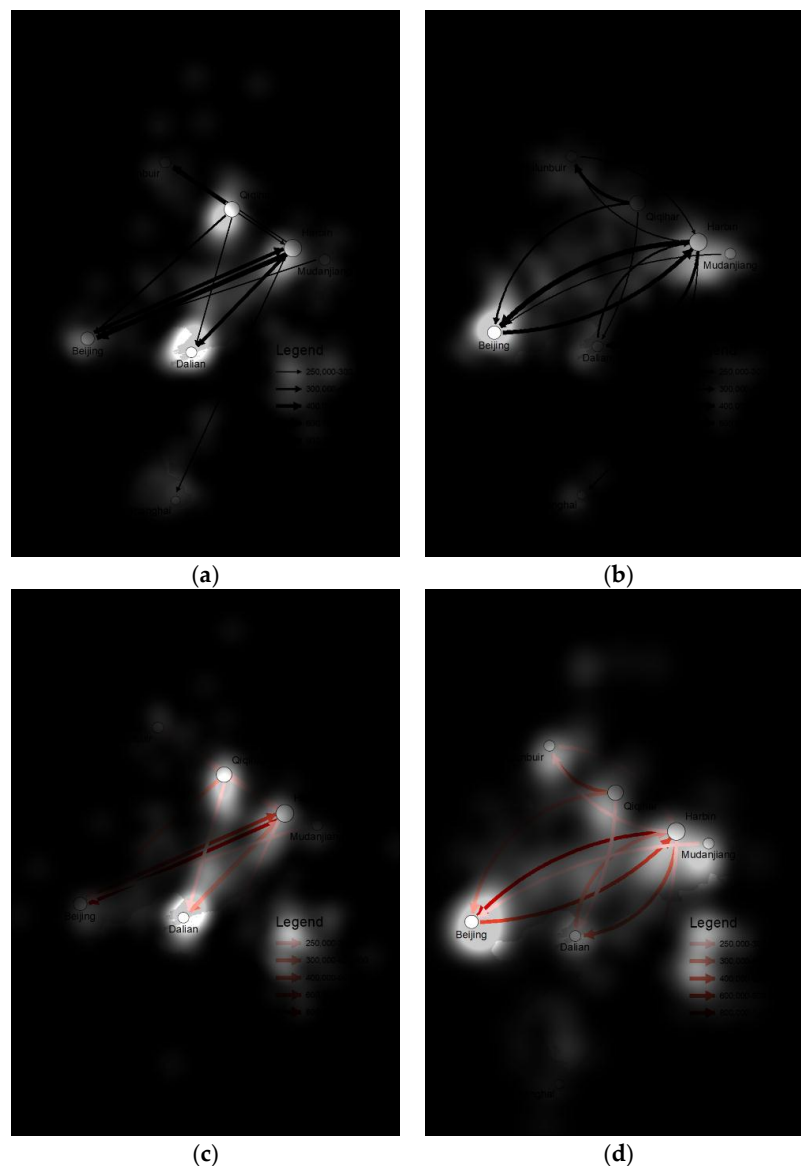
As the results of finish time show (Table 3 and Figure 3), the maps using curved features to represent the flow direction were significantly less efficient, which is a result similar to the findings of previous studies [21,22]. In our experiment, finish time reflects the average time that participants used in determining the directions of flows, comparing the volumes of flows, and judging the correct answers in one task. The participants needed to spend more time and efforts on maps using curves than on maps using straight lines.

The curve group shows that there was a higher number of fixations than that in the group with straight lines (Table 3 and Figure 3). Undoubtedly, fixation appeared more because of the greater area containing information in the maps using curves. According to the results of percentage of fixations in AOIs (Table 3 and Figure 3), the straight-line group showed a lower percentage of fixations in AOIs than the curve group. The participants spent less effort on searching flows corresponding to correct answers in maps using straight lines. This result means that, compared with straight lines, curves will decrease the searching efficiency and cost more time in nonsearching behaviors. According to the gaze opacity maps (Figure 6), in maps using straight lines, the opacities were very low at target points (starting point and end point), such as Dalian and Qiqihar in Figure 6a,b, which means that users focused more on the terminals of straight flows. However, in maps using curves, the opacities at target points were not as low as those in maps using straight lines. While in the flow body area, maps using curves have a much lower opacity than maps using straight lines, especially for flows that do not come from or to the target points, such as Dalian and Hulunbuir in Figure 6c,d. This finding means that users were more concerned about flows in maps using curves than in maps using straight lines. We inferred that for straight-line features, the participants knew that the coordinates of the target points should be along the same straight line, so they focused more on the target points. In this case, the participants were able to locate the target flows (correct answers) by quickly scanning the straight lines. On the other hand, for curved features, the participants needed to pay more attention to scanning along the curves to judge the directions of flows. This result is consistent with previous findings that straight lines do not detour and have a narrower visual search range [22]. Thus, maps using straight lines lead to a higher efficiency than maps using curves.

Although the participants performed less efficiently while reading the curve maps, curve maps contribute to a more effective performance than straight-line maps, which is similar to the findings of previous studies [1]. This finding is mainly reflected in the results of accuracy, as the curve group exhibited a greater number of correct answers than the straight-line group (Table 3 and Figure 3).

According to the questionnaire results (Figure 5a), users think maps using curves are clearer than maps using straight lines. For maps using straight lines, there were more overlapping or intersecting line features, which made it more difficult to effectively retrieve the required information

and even caused participants to misunderstand the maps. For maps using curves, these overlapping or intersecting line features were smoothed with different curvatures, which helped reduce the number of overlapping and intersecting features and prevented excessive line curvature, making the maps distinguishable. However, the results of finish time and fixation count show that participants spent less time and effort on straight-line maps. Participants tended to read maps using straight lines less carefully and were not aware that there would be overlapping or intersecting line features in maps that might lead them to be misunderstood. They failed to obtain correct answers in maps using straight lines. This result is consistent with previous findings that overlapping or intersecting features may cause a decrease in task correctness [1,21]. Thus, maps using straight lines lead to a lower effectiveness than maps using curves.



**Figure 6.** Samples of gaze opacity maps (generated by the fixation count of all participants with a radius of 50 pixels in Tobii Studio). SL = straight line; CV = curve; LT = line thickness; and CG = color gradient. (a) SL + LT; (b) CV + LT; (c) SL + CG; and (d) CV + CG.

There was no significant difference in the time to first fixation between maps using curves and maps using straight lines. This means that two types of line features have the same contributions to visual guidance in representing flows.

## 5.2. Comparison between the Line Thickness and the Color Gradient

Table 4 shows that the color gradient yielded a better effectiveness than different line thicknesses based on the greater number of correct answers, which is contrary to the finding of a previous study by Garlandini and Fabrikant [15]. They conducted an eye-tracking experiment using maps that contained points with different sizes and areas with different color gradients and found that size was more effective and efficient than color gradients for change detection. In our study, according to the questionnaire results (Figure 5b), users think maps using color gradients are clearer than maps using line thicknesses, and they feel that it is more difficult to identify volume correctly on maps using line thicknesses. For the line-thickness group, the differences of line features were not salient. For the color gradient group, the changes in the volumes were clearer, and the participants were able to retrieve the correct answers more easily. In addition, the gaze opacity maps show that participants paid attention to legends in maps using color gradients, but very few fixations covered legends in maps using line thicknesses. Participants naturally believed that thick line features represent large volumes, and they may not spend extra efforts on interpreting legends, while participants need to refer to legends for the color-gradient maps because they are less sensitive to color gradients than line thicknesses. We infer that enlarging the difference between line thicknesses may increase participants' sensitiveness to them. Moreover, for maps using line thickness, some flows with very small volumes become overly thin, and their arrowheads may be covered by nearby flows or be too small to be observed. This problem will lead participants to misunderstand maps. This result is consistent with previous findings that the sizes of arrowheads being proportional to line thicknesses is improper for application with thin lines because of a decrease in readability [1].

However, there was no significant difference in the finish time, fixation count, percentage of fixations in the AOIs, or time to first fixation between using variable line thicknesses and a red color gradient to represent the flow volume (Table 4, Figure 4). Although participants could obtain more correct answers when reading maps using color gradients, the participants of the two groups spent the same amount of work and time obtaining the information, which can be seen from the results of the fixation count and finish time (Table 4 and Figure 4). In addition, line thicknesses and color gradients contributed neither to increasing the searching efficiency for flows corresponding to correct answers, nor to shortening the time required to guide attention to them due to the similar result for fixation in the AOIs and time to first fixation. We infer that there were no significant differences between the difficulties of decoding information from maps using the two different methods to represent flow volume; thus, users had the same burden to finish the tasks. Therefore, although the participants of the line thickness group exerted the same effort as the participants in the color gradient group, their endeavors did not produce a significant improvement in the availability of effective information and occasionally caused them to misread the map.

## 6. Conclusions and Future Work

This study was intended to explore the effects of visual variables on the usability of flow maps. The results of indoor eye tracking experiments show that the use of curves instead of straight lines to visualize flows benefits the effectiveness of flow maps because using curves reduces the number of overlapping and intersecting features, thus improving the clarity of maps. For maps using curves, participants paid more attention to effective areas while performing tasks that involved reading the maps and searching for information, and they were more likely to collect accurate information. However, straight-line maps tend to exhibit more overlapping and intersecting features, which could hinder the perception of accurate information and even cause maps to be misread. In addition, the use of color gradients could significantly improve the effectiveness of maps compared to the use of different line thicknesses.

Notably, our study only provides a comparison between a single color gradient and different line thicknesses. Therefore, the influence of other color gradients and relevant variables on the usability of

flow maps should be tested in further investigations with larger datasets that contain more nodes and flows and more participants with different backgrounds.

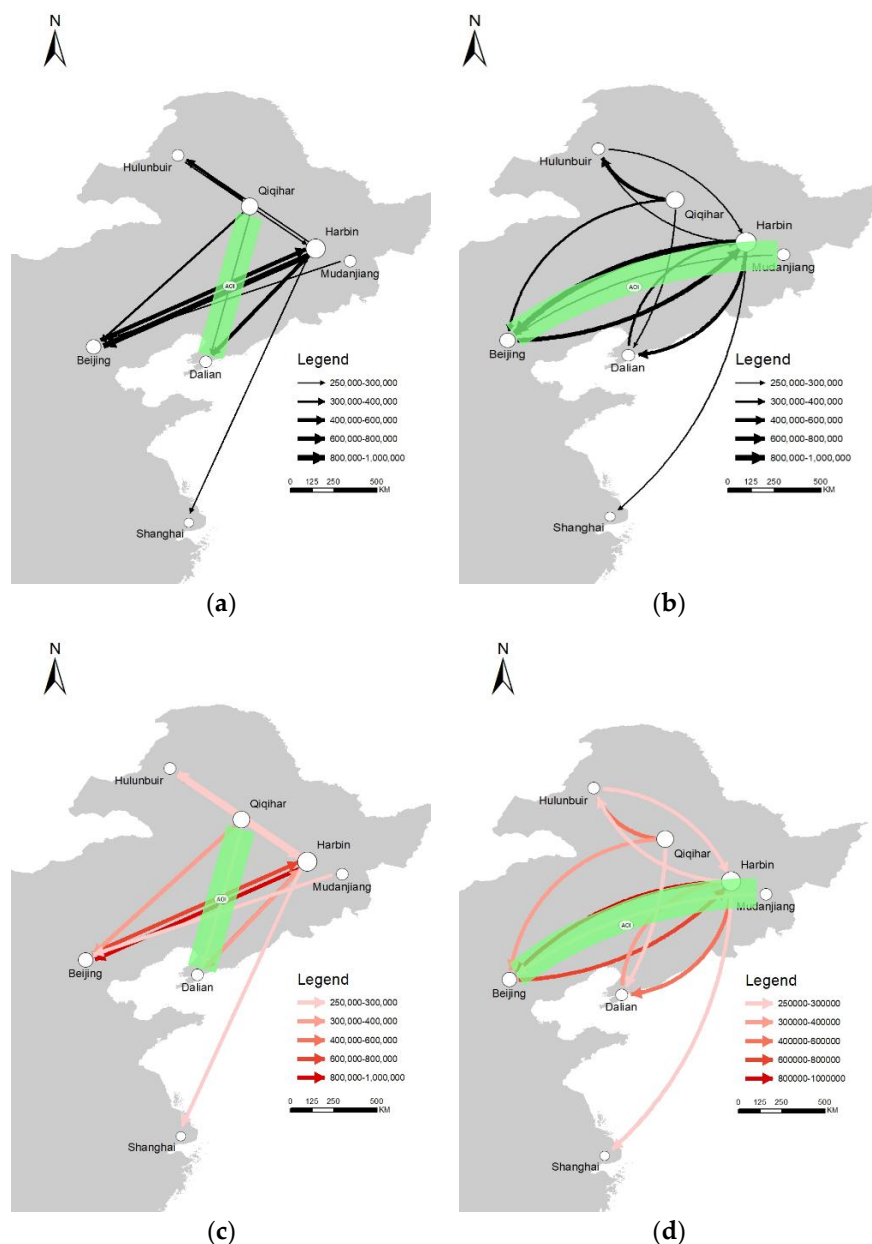
**Author Contributions:** Conceptualization, Weihua Dong; Methodology, Shengkai Wang; Formal Analysis, Yizhou Chen; Writing—Original Draft Preparation, Shengkai Wang; Writing—Review & Editing, Weihua Dong and Liqiu Meng; Visualization, Yizhou Chen; Supervision, Weihua Dong.

**Funding:** This research is supported by the National Key Research and Development Program of China (Grant No. 2017YFB0503602), National Natural Science Foundation of China (NSFC, Grant No. 41471382).

**Acknowledgments:** The authors would like to thank all the reviewers for their helpful comments and suggestions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A



**Figure A1.** Samples of maps with AOIs. SL = straight line; CV = curve; LT = line thickness; and CG = color gradient. (a) SL + LT; (b) CV + LT; (c) SL + CG; and (d) CV + CG.



## References

- Jenny, B.; Stephen, D.M.; Muehlenhaus, I.; Marston, B.E.; Sharma, R.; Zhang, E.; Jenny, H. Design principles for origin-destination flow maps. *Cartogr. Geogr. Inf. Sci.* **2018**, *45*, 62–75. [CrossRef]
- Boyandin, I.; Bertini, E.; Lalanne, D. Using flow maps to explore migrations over time. In Proceedings of the Geospatial Visual Analytics Workshop in Conjunction with the 13th AGILE International Conference on Geographic Information Science, Guimaraes, Portugal, 11–14 May 2010.
- Grandjean, M. Connected World: Untangling the Air Traffic Network. Available online: <http://www.martingrandjean.ch/connected-world-air-traffic-network/> (accessed on 20 July 2018).
- Balcan, D.; Colizza, V.; Gonçalves, B.; Hu, H.; Ramasco, J.J.; Vespignani, A. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc. Nat. Acad. Sci. USA* **2009**, *106*, 21484–21489. [CrossRef] [PubMed]
- Guo, D. Flow mapping and multivariate visualization of large spatial interaction data. *IEEE Trans. Vis. Comput. Gr.* **2009**, *15*. [CrossRef]
- Padmanabhan, A.; Wang, S.; Cao, G.; Hwang, M.; Zhao, Y.; Zhang, Z.; Gao, Y. Flumapper: An interactive cybergis environment for massive location-based social media data analysis. In Proceedings of the Conference on Extreme Science and Engineering Discovery Environment: Gateway to Discovery, San Diego, CA, USA, 22–25 July 2013; p. 33.
- Jenny, B.; Stephen, D.M.; Muehlenhaus, I.; Marston, B.E.; Sharma, R.; Zhang, E.; Jenny, H. Force-directed layout of origin-destination flow maps. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 1521–1540. [CrossRef]
- Zhou, H.; Xu, P.; Yuan, X.; Qu, H. Edge bundling in information visualization. *Tsinghua Sci. Technol.* **2013**, *18*, 145–156. [CrossRef]
- Vrotsou, K.; Fuchs, G.; Andrienko, N.; Andrienko, G. An interactive approach for exploration of flows through direction-based filtering. *J. Geovisual. Spat. Anal.* **2017**, *1*. [CrossRef]
- Rae, A. Flow-data analysis with geographical information systems: A visual approach. *Environ. Plan. B Plan. Des.* **2011**, *38*, 776–794. [CrossRef]
- Koylu, C.; Guo, D. Design and evaluation of line symbolizations for origin—Destination flow maps. *Inf. Vis.* **2017**, *16*, 309–331. [CrossRef]
- Bertin, J. *Graphics and Graphic Information Processing*; Walter de Gruyter: Berlin, Germany, 1981.
- Wolfe, J.M.; Horowitz, T.S. What attributes guide the deployment of visual attention and how do they do it? *Nat. Rev. Neuros.* **2004**, *5*, 495–501. [CrossRef] [PubMed]
- Dong, W.; Ran, J.; Wang, J. Effectiveness and efficiency of map symbols for dynamic geographic information visualization. *Cartogr. Geogr. Inf. Sci.* **2012**, *39*, 98–106. [CrossRef]
- Garlandini, S.; Fabrikant, S.I. Evaluating the effectiveness and efficiency of visual variables for geographic information visualization. In Proceedings of the International Conference on Spatial Information Theory, Aber Wrach, France, 21–25 September 2009; Springer: Berlin, Germany, 2009; pp. 195–211.
- Kubíček, P.; Šašinka, Č.; Stachoň, Z.; Štěrbá, Z.; Apeltauer, J.; Urbánek, T. Cartographic design and usability of visual variables for linear features. *Cartogr. J.* **2017**, *54*, 91–102. [CrossRef]
- Holten, D.; Isenberg, P.; Van Wijk, J.J.; Fekete, J.-D. An extended evaluation of the readability of tapered, animated, and textured directed-edge representations in node-link graphs. In Proceedings of the 2011 IEEE Pacific Visualization Symposium (PacificVis), Hong Kong, China, 1–4 March 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 195–202.
- Holten, D.; van Wijk, J.J. A user study on visualizing directed edges in graphs. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, 4–9 April 2009; pp. 2299–2308.
- Imhof, E. *Thematische Kartographie*; Walter de Gruyter: Berlin, Germany, 1972; Volume 10.
- Dent, B.D.; Torguson, J.S.; Hodler, T.W. *Cartography: Thematic Map Design*; WCB/McGraw-Hill Boston: Boston, MA, USA, 1999; Volume 5.
- Xu, K.; Rooney, C.; Passmore, P.; Ham, D.-H.; Nguyen, P.H. A user study on curved edges in graph visualization. *IEEE Trans. Vis. Comput. Gr.* **2012**, *18*, 2449–2456. [CrossRef] [PubMed]
- Purchase, H.C.; Hamer, J.; Nöllenburg, M.; Kobourov, S.G. On the usability of lombardi graph drawings. In Proceedings of the International Symposium on Graph Drawing, Redmond, WA, USA, 19–21 September 2012; Springer: Berlin, Germany, 2012; pp. 451–462.

23. Steinke, T.R. Eye movement studies in cartography and related fields. *Cartogr. Int. J. Geogr. Inf. Geovisual.* **1987**, *24*, 40–73. [[CrossRef](#)]
24. Reichenbacher, T.; Swienty, O. Attention-guiding geovisualization. In Proceedings of the 10th AGILE International Conference on Geographic Information Science, Aalborg, Denmark, 8–11 May 2007; Aalborg University: Aalborg, Denmark, 2007.
25. Webb, N.; Renshaw, T. Eyetracking in hci. *Res. Methods Hum. Comput. Interact.* **2008**, 35–69.
26. Çöltekin, A.; Fabrikant, S.I.; Lacayo, M. Exploring the efficiency of users' visual analytics strategies based on sequence analysis of eye movement recordings. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 1559–1575. [[CrossRef](#)]
27. Çöltekin, A.; Heil, B.; Garlandini, S.; Fabrikant, S.I. Evaluating the effectiveness of interactive map interface designs: A case study integrating usability metrics with eye-movement analysis. *Cartogr. Geogr. Inf. Sci.* **2009**, *36*, 5–17. [[CrossRef](#)]
28. Ooms, K.; De Maeyer, P.; Fack, V.; Van Assche, E.; Witlox, F. Interpreting maps through the eyes of expert and novice users. *Int. J. Geogr. Inf. Sci.* **2012**, *26*, 1773–1788. [[CrossRef](#)]
29. Ooms, K.; Maeyer, P.D.; Fack, V.; Assche, E.V.; Witlox, F. Investigating the effectiveness of an efficient label placement method using eye movement data. *Cartogr. J.* **2012**, *49*, 234–246. [[CrossRef](#)]
30. Dong, W.; Liao, H.; Roth, R.E.; Wang, S. Eye tracking to explore the potential of enhanced imagery basemaps in web mapping. *Cartogr. J.* **2014**, *51*, 313–329. [[CrossRef](#)]
31. Popelka, S.; Brychtova, A. Eye-tracking study on different perception of 2 d and 3 d terrain visualisation. *Cartogr. J.* **2013**, *50*, 240–246. [[CrossRef](#)]
32. Liao, H.; Wang, X.; Dong, W.; Meng, L. Measuring the influence of map label density on perceived complexity: A user study using eye tracking. *Cartogr. Geogr. Inf. Sci.* **2018**. Available online: <https://www.tandfonline.com/doi/full/10.1080/15230406.2018.1434016> (accessed on 20 July 2018). [[CrossRef](#)]
33. Alaçam, Ö.; Dalci, M. A usability study of webmaps with eye tracking tool: The effects of iconic representation of information. In Proceedings of the International Conference on Human-Computer Interaction, San Diego, CA, USA, 19–24 July 2009; Springer: Berlin, Germany; pp. 12–21.
34. Dong, W.; Liao, H.; Xu, F.; Liu, Z.; Zhang, S. Using eye tracking to evaluate the usability of animated maps. *Sci. China Earth Sci.* **2014**, *57*, 512–522. [[CrossRef](#)]
35. Ho, H.Y.; Yeh, I.C.; Lai, Y.C.; Lin, W.C.; Cherng, F.Y. Evaluating 2d flow visualization using eye tracking. *Comput. Gr. Forum* **2015**, *34*, 501–510. [[CrossRef](#)]
36. Blascheck, T.; Burch, M.; Meisel, T.; Schneider, T.; Mumin, S. Exploring eye movements with node-link graph layouts. In *Eye Tracking for Spatial Research, Proceedings of the 3rd International Workshop, Zurich, Switzerland, 14 January 2018*; ETH Zurich: Zurich, Switzerland, 2018.
37. Netzel, R.; Burch, M.; Weiskopf, D. Comparative eye tracking study on node-link visualizations of trajectories. *IEEE Trans. Vis. Comput. Gr.* **2014**, *20*, 2221–2230. [[CrossRef](#)] [[PubMed](#)]
38. Goldberg, J.H.; Kotval, X.P. Computer interface evaluation using eye movements: Methods and constructs. *Int. J. Ind. Ergon.* **1999**, *24*, 631–645. [[CrossRef](#)]
39. Byrne, M.D.; Anderson, J.R.; Douglass, S.; Matessa, M. Eye tracking the visual search of click-down menus. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Pittsburgh, PA, USA, 15–20 May 1999; pp. 402–409.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).