International Journal of
Geo-Information

MDPI

Article

Evaluating the Usability of a Gaze-Adaptive Approach for Identifying and Comparing Raster Values between Multilayers

Changbo Zhang [1,2], Hua Liao [1,2,*], Yongbo Huang [3] and Weihua Dong [4]

1 School of Geographic Sciences, Hunan Normal University, Changsha 410081, China; cat_zcb@hunnu.edu.cn
2 Hunan Key Laboratory of Geospatial Big Data Mining and Application, Hunan Normal University, Changsha 410081, China
3 School of Geography, South China Normal University, Guangzhou 510631, China; 2021022876@m.scnu.edu.cn
4 Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China; dongweihua@bnu.edu.cn
* Correspondence: liaohua@hunnu.edu.cn

Abstract: Raster maps provide intuitive visualizations of remote sensing data representing various phenomena on the Earth's surface. Reading raster maps with intricate information requires a high cognitive workload, especially when it is necessary to identify and compare values between multiple layers. In traditional methods, users need to repeatedly move their mouse and switch their visual focus between the map content and legend to interpret various grid value meanings. Such methods are ineffective and may lead to the loss of visual context for users. In this research, we aim to explore the potential benefits and drawbacks of gaze-adaptive interactions when interpreting raster maps. We focus on the usability of the use of low-cost eye trackers on gaze-based interactions. We designed two gaze-adaptive methods, gaze fixed and gaze dynamic adaptations, for identifying and comparing raster values between multilayers. In both methods, the grid content of different layers is adaptively adjusted depending on the user's visual focus. We then conducted a user experiment by comparing such adaptation methods with a mouse dynamic adaptation method and a traditional method. Thirty-one participants (n = 31) were asked to complete a series of single-layer identification and multilayer comparison tasks. The results indicated that although gaze interaction with adaptive legends confused participants in single-layer identification, it improved multilayer comparison efficiency and effectiveness. The gaze-adaptive approach was well received by the participants overall, but was also perceived to be distracting and insensitive. By analyzing the participants' eye movement data, we found that different methods exhibited significant differences in visual behaviors. The results are helpful for gaze-driven adaptation research in (geo)visualization in the future.

Keywords: gaze adaptation; gaze-based HCIs; eye tracking; visual behavior; low-cost eye tracker

Citation: Zhang, C.; Liao, H.; Huang, Y.; Dong, W. Evaluating the Usability of a Gaze-Adaptive Approach for Identifying and Comparing Raster Values between Multilayers. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 412. https://doi.org/10.3390/ijgi12100412

Academic Editors: Wolfgang Kainz and Florian Hruby

Received: 6 July 2023
Revised: 25 September 2023
Accepted: 6 October 2023
Published: 8 October 2023

1. Introduction

Raster maps have high information density. Interpreting raster maps requires a high cognitive workload because a raster consists of many regular grid cells representing a specific theme of objects (e.g., population, land cover/use, precipitation and PM2.5 concentration) at a specific time in space. A raster can be visualized using continuous, stratified or discrete colormaps (e.g., Figure 1) [1]. Moreover, multiple raster layers (usually with the same spatial extent and grid cell size) can be overlaid to represent time-series thematic information.

Identifying and comparing raster cell values between layers are basic interactive operations for raster map interpretation. For instance, as shown in Figure 1, a user may want to know the population densities of Point A in 2000, 2010 and 2020 and then compare the values to find the increasing/decreasing population trend. In many GIS software applications (e.g., ArcGIS, QGIS and ENVI), geo-applications and web maps, users can use the *identify* and *swipe* tools to perform identification and comparison operations (e.g.,

Figure 1b,c). The *identify* tool can show the raster value(s) of the clicked cell from one or all layers, but cannot indicate its stratified class (e.g., Point A in 2020 belongs to the 0–10,000 class). The *swipe* tool can easily reveal the differences between the visible layer and the underlying layer. For both tools, users need to click or drag their mouse while repeatedly switching their visual attention between the target grid cell (e.g., Point A) and the legend/identification panel. This process is more cognitively demanding when the number of layers or the number of stratified/discrete classes increases. During such repeated attention switches, users may lose their visual context [2]. Moreover, moving their eyes back and forth may lower the raster map comprehension efficiency.

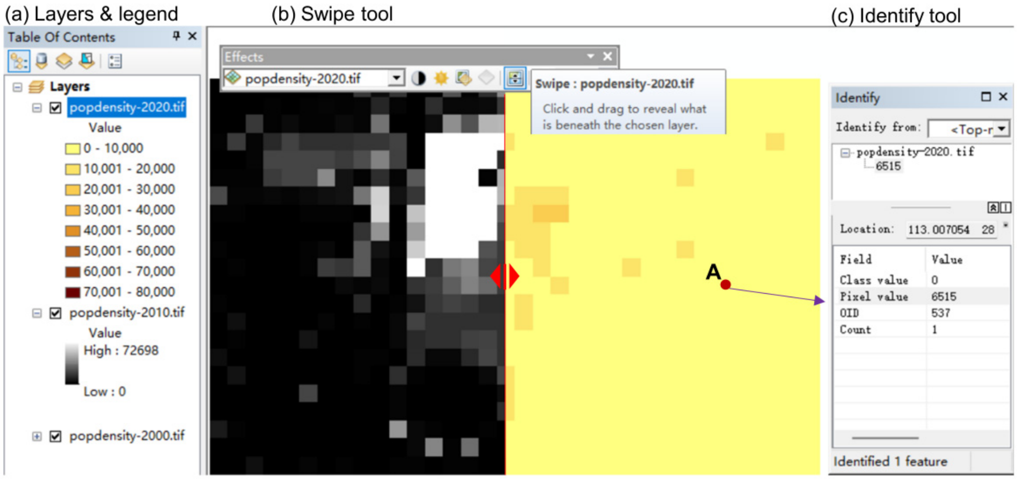

Figure 1. Identifying and comparing population densities between 2000, 2010 and 2020 using the swipe tool and identify tool. (**a**) Layers and legend, (**b**) swipe tool and (**c**) identify tool. Although this example was captured from ESRI ArcMap 10.2, similar tools can be found in many other GIS software applications (e.g., QGIS and ENVI), geo-applications and web maps.

Gaze behavior has exhibited the potential to provide adaptive human–computer interactions (HCIs) because it has been shown to be an effective indicator of user intention [3–11]. By tracking a user's eye movements, a computer system can respond in real time to the user's visual attention [12]. Eye tracking has become a promising new human–computer interaction modality [13]. With improvements in eye tracker portability, affordability and tracking accuracy (e.g., Tobii Eye Tracker 5, ~EUR 259), eye trackers can be easily mounted on or embedded in personal computers (e.g., Lenovo Legion 9000 K, China, https://shop.lenovo.com.cn/, accessed on 8 October 2023). Eye trackers may become ubiquitous sensors, leading to pervasive, gaze-based HCIs in the future. However, gaze interaction remains an ongoing research area [14,15]. Evidence of gaze adaptation application to raster map interpretation is rare.

We contribute to this body of research by suggesting a gaze-adaptive approach for identifying and comparing raster values between multiple layers. We focus on using a low-cost eye tracker for gaze interaction and its usability (i.e., effectiveness, efficiency, visual behavior and user experience) in real-world map-reading scenarios [16]. Only identification and comparison were considered to control the usability experiment, leaving more complex gaze interactions for future investigations.

After reviewing related work in the next section (Section 2), we detail the design and implementation of our gaze-adaptive approach in Section 3. An evaluation experiment is presented in Section 4, and the results are analyzed in Section 5. We discuss the benefits and drawbacks of the presented approach in Section 6 and draw conclusions and present future work in Section 7.

2. Background and Related Work

2.1. Eye tracking for Human–Computer Interaction

Eye tracking has been used as an input modality for decades [17,18]. There is increasing interest in gaze interaction, especially in virtual/augmented reality (VR/AR) [19–23], partly due to improvements in tracking accuracy, portability and affordability. One critical issue in gaze interaction is avoiding unintended selection (i.e., a user looking at an object does not mean he or she wants to interact with it), which is referred to as the *Midas Touch* problem [17]. A second issue is that gaze tracking is spatially inaccurate because high visual acuity is limited to the central 2° of the human foveal region [12]. Gaze pointing thus has less spatial accuracy than traditional mouse pointing. The typical tracking accuracy of current eye trackers ranges from 0.5°~1°.

Various solutions have been proposed to mitigate the Midas Touch problem and limited spatial accuracy. Using a longer gaze time on the target (dwell) and a larger button size are considered effective approaches [24]. Dwell time ranges from 100 ms to 3000 ms depending on specific applications [25]. For example, Hansen et al. [26] utilized a 500 ms dwell time for typing Japanese sentences. Dunphy et al. [27] also used 500 ms for their graphical passwords to unlock a simulated ATM. Feit et al. [28] found that tracking accuracy decreases when eyes approach the edge of a screen. Gaze gestures are another robust approach that has gained research attention in recent years [29,30]. Other solutions include combining eye movements and head movements [31,32].

In cartography and GIScience, Göbel et al. [33] developed a gaze interactive campus map that enables users to select buildings by fixating on rectangular menus rather than directly selecting area features on a map. More recently, Zhu et al. [34] selected an area by determining the area that contains a user's current gaze. To select points, lines and areas for gaze interaction, Liao et al. [35] examined various combinations of dwell time and buffer size. They discovered that buffer sizes of 1.5° and 0.7° are suitable for selecting points and lines, respectively, and that a dwell time of 600 ms provides a better balance between accuracy and efficiency than 200 ms and 1000 ms.

Despite the diversity of approaches for gaze-based target selection, dwell selection is still a mainstream approach in gaze interaction because it is easy and straightforward [21,28]. Therefore, in the present study, we follow this research line by applying dwell time to gaze-based raster map reading.

2.2. Gaze-Driven Adaptive (Geo)visualization

Using eye tracking as a modality to provide adaptations for (geo)visualizations is natural. To date, however, gaze-driven adaptations have been investigated in only a few studies. Bektaş and Çöltekin [36] were among the first to explore gaze-driven adaptations in GIScience. They proposed foveation algorithms to build a multiresolution geospatial data structure (representing different levels of details) and visualize them on gaze-contingent displays (GCDs). The GCD system offers high-resolution data visualization around a user's central field of view, as indicated by the user's eye movements, while providing a low resolution in peripheral areas [36].

Giannopoulos et al. [37] introduced a 'GeoGazemarks' system that records and clusters the users' eye movements on smartphone maps and visualizes their fixation clusters as a visual clue to help users find their previously visited places. The authors found that this system can lower the response time and improve accuracy in searching map targets. In a 'GazeGIS' application developed by Tateosian et al. [38], when users read narrative passages and gaze at a place name, the application geolocates the place, shows its position on a map, and provides related photos.

Göbel et al. [39] suggested a gaze-adaptive legend approach that can display the meaning of a symbol that a user is looking at (content adaptation). Therefore, users do not need to spend time searching and matching symbols from a full list of symbols on a legend panel. Moreover, they compared legends with a fixed location and an adjustable location (i.e., position adaptation: showing the legend near the user's gaze position) and found that users preferred a fixed location legend.

More recently, Lallé et al. [40] and Barral et al. [41] extended gaze-driven adaptations to 'magazine-style narrative visualizations' (MSNVs). They suggested adaptations that emphasize related data points in bar charts when users read associated narratives in MSNVs. They found that the adaptations can guide users' visual attention to salient visualization elements and, therefore, are helpful for low-literacy users in interpreting data visualizations. Keskin and Kettunen [42] suggested combining eye tracking and machine learning methods to build gaze-aware interactive map systems. They emphasize simplifying, detecting and visualizing vector data based on what users focus on. For example, the system detects the map features that the user is attending to, selects other map features that have the same style and highlights these features on the map.

While previous studies have focused on developing gaze-driven adaptations for vector data maps, the present study concentrates on exploring raster data interpretation aided by eye tracking. We emphasize its use with user-affordable eye trackers and its usability in raster map reading. The details are presented in Sections 3 and 4.

3. Gaze-Adaptive Approach: Design

3.1. Design Considerations

Following Göbel et al. [39]'s idea of content adaptation and position adaptation, we designed two gaze-adaptive methods for identifying and comparing raster values between multiple layers.

- **Gaze dynamic adaptation (GD)**. In this method, the grid information viewed by the user is displayed in a dynamic window that is always near the gaze focus (Figure 2a). Based on the results of [39], we placed the window at the right-bottom corner of the user's gaze point at approximately 2.7° (≈2.36 cm, 180 px). Displaying the information window besides the user's gaze point is considered intuitive because it can decrease the visual search distance between the current gaze and the legend. Note that the window is always visible to ensure that users can obtain the grid information as quickly as possible. In the dynamic information window, the items, including the year, color block and label of the gaze position, were automatically extracted from the layers and their symbology, meaning that the information in the window is consistent with the layer panel (shown on the left). The gaze dynamic adaptation is able to deal with discrete, stratified and continuous raster maps (see Figure 3).
- **Gaze fixed adaptation (GF)**. Different from gaze dynamic adaptation, in this method, the position of the information window of the grid is fixed at the top-left corner of the screen, but its content is adapted to gaze (Figure 2b). The other settings are the same as the GD. Göbel et al. [39] found that participants preferred fixed adaptation rather than dynamic adaptation. Therefore, in this study, we tested whether a fixed information window is preferred in gaze-based raster map reading.

In addition to the two gaze-adaptive methods, the experiment also includes two other methods: traditional identification and mouse dynamic adaptation.

- **Traditional identification (TR)**. This method was used as a baseline for the comparison experiment. No adaptation was provided in this method and the participants needed to use the *identify* tool to obtain the raster values. By clicking the visible layer, the information (including the layer names, classes and colormaps) of the clicking grid is displayed in the form of a pop-up window (Figure 2c). Then, users can interpret the raster maps by combining the information in the layer control and the pop-up window. When users want to view the information of other layers, however, they have to switch the visible layer and repeat the previous operation.
- **Mouse dynamic adaptation (MD)**. In mouse dynamic adaptation, we used a mouse pointer to replace gaze, but kept other settings unchanged, as in GD. The grid information that was pointed to by the mouse was displayed in the dynamic window

(Figure 2d). Since a mouse is more accurate than gaze in pinpointing and without the consideration of the Midas contact problem, this method seemed to perform the best. We designed this method to include a high-precision input as a benchmark to better understand the limitations of eye-tracking technology in practical scenarios. Additionally, the comparison with MD may help identify the unique advantages of a gaze-based interaction, such as its potential for hands-free and more intuitive interactions in certain contexts.

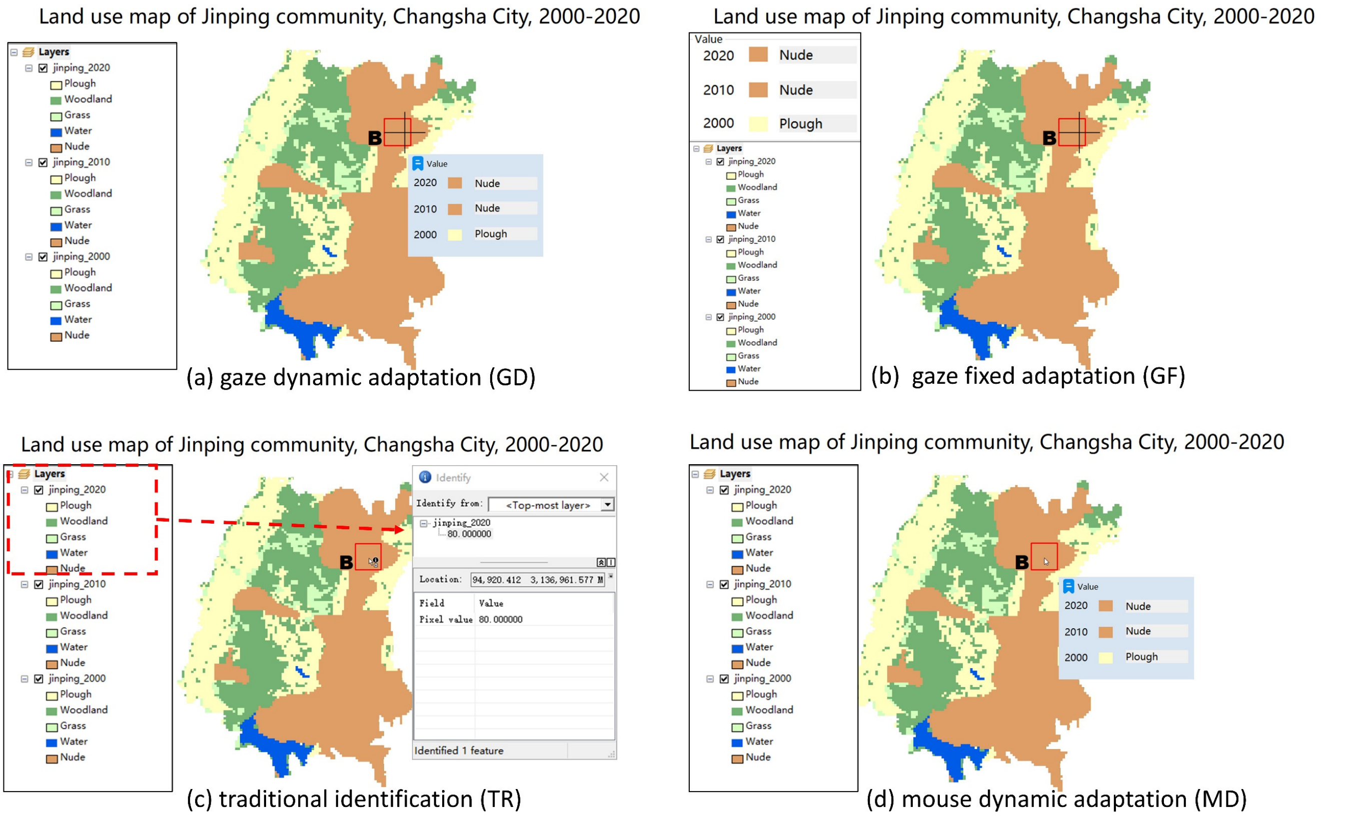

Figure 2. Illustration of (**a**) gaze dynamic adaptation (GD), (**b**) gaze fixed adaptation (GF), (**c**) traditional identification (TR) and (**d**) mouse dynamic adaptation (MD).

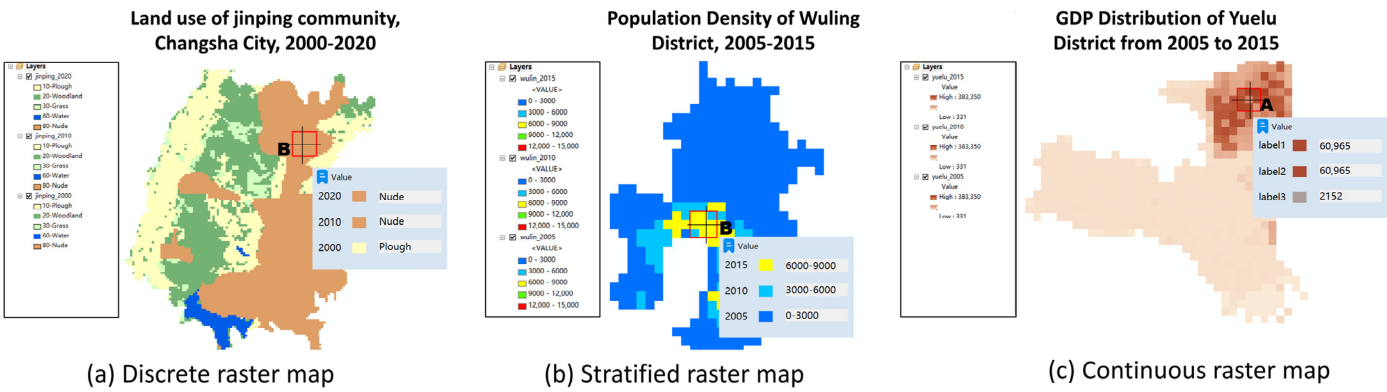

Figure 3. Example of a (**a**) discrete, (**b**) stratified and (**c**) continuous map in gaze dynamic adaptation.

3.2. Technical Framework

Figure 4 shows a prototype of the adaptive system and Figure 5 shows the technical framework with three components.

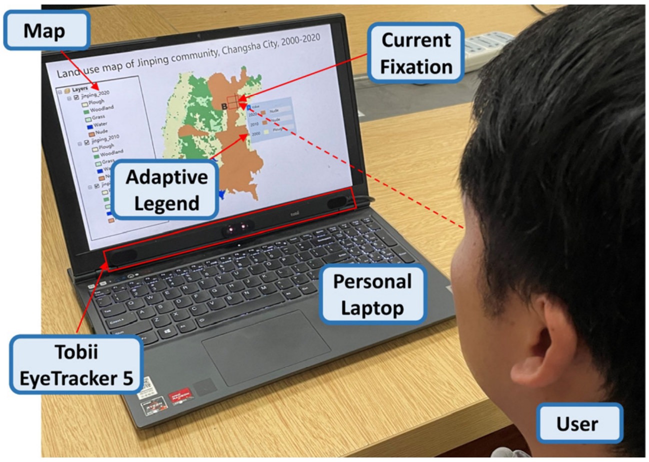

Figure 4. A prototype of the system.

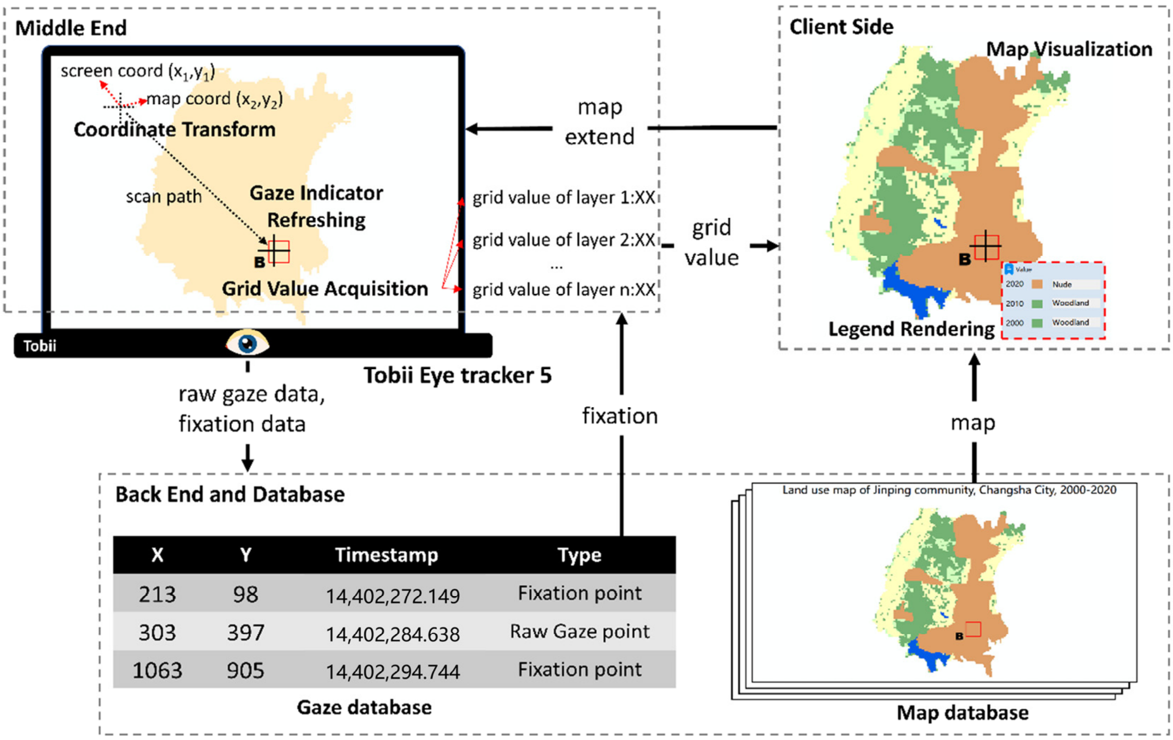

Figure 5. The technical framework of the system.

- The **backend and database** are connected with the Tobii Eye Tracker 5 and are responsible for real-time gaze data collection and map storage. Two kinds of gaze data were provided by the Tobii API: the raw gaze data stream and the fixation data stream. We adopted the fixation data stream that was calculated by the built-in Tobii I-VT algorithm from the raw gaze data in real time [43].
- The **middle end** manages the delivery of the adaptation based on the fixations generated by the backend. When the system is working, the user's fixation is indicated as a black crosshair on the screen as feedback of the user's fixation position. Fixations from screen coordinates are first converted to map (e.g., georeferenced) coordinates. Then, the system obtains the raster grid values of different layers according to the current fixation position and sends them to the client side.
- The **client side** presents the data and adaptation. It first displays maps from the database. After receiving the raster grid values from the middle end, the client side

then renders the legend to present the layer information to users. A "+" marker is displayed on the screen to show the user's current gaze position.

4. Evaluation

A user study based on a comparison experiment was conducted to gather insights into the performance of the proposed gaze-adaptive methods and to gather user feedback.

4.1. Experiment

4.1.1. Participants

Thirty-one undergraduate and postgraduate student participants (14 males and 17 females) aged between 18 and 25 years (M = 22.26 years, SD = 1.84 years) were recruited for the experiment from the School of Geographical Sciences, Hunan Normal University. All participants had normal or corrected-to-normal vision. The study was conducted in accordance with the Declaration of Helsinki and approved by the local institutional review board (IRB). They all signed informed consent forms and were compensated for their participation.

4.1.2. Apparatus and Software

A Tobii Eye Tracker 5 (Tobii, Sweden, www.tobii.com, accessed on 14 June 2022) and a HUAWEI MateBook 13 laptop (HUAWEI, China, consumer.huawei.com, accessed on 14 June 2022; Intel i7 8565U CPU, 1.8 GHz and 16 GB RAM) were used. The eye tracker ran at 90 Hz and was connected to the laptop with a USB 2.0 interface. The laptop had a 13-inch LED screen with a full HD resolution of 2160 × 1440 px (28.6 cm × 21.1 cm). The experimental platform was developed using ESRI ArcObjects 10.2 and Tobii Interactor APIs 0.7.3 with the C# programming language. All equipment was placed in the laboratory to ensure constant light conditions and a noise-free environment.

4.1.3. Materials and Tasks

The participant's task was to answer a four-choice question by reading a thematic map that was associated with the task. All the thematic maps used in this experiment consisted of three raster layers representing the mapping area in different years. The thematic contents of these maps included land use type, population density and GDP distributions, which were represented using discrete, stratified and continuous symbology, respectively. These three types of symbology constitute most of the raster representations. *Discrete* usually shows categorial variables (here, land use type) with each color representing a category; *stratified* divides the values into several (usually graduated) groups (classes) (here, the population density levels), with each color representing a group; and *continuous* displays the data using stretched color scheme (here, GDP). The dynamic information window shows the categorial names (discrete maps), stratified class names (stratified maps) or the grid values (continuous maps) depending on the type of symbology. Examples are shown in Figure 3.

The mapping areas were from different community/district administrative regions in China. A questionnaire showed that the participants had no prior knowledge of the mapping area or the thematic data in these regions. The grid size of the land use, population density and GDP maps is 30, 1000 and 1000 m, respectively. Since the mapping areas are from various regions, their spatial distributions are different from each other.

For each task, three phases were included: a question-reading phase, a map-reading phase and a question-answer phase (Figure 6).

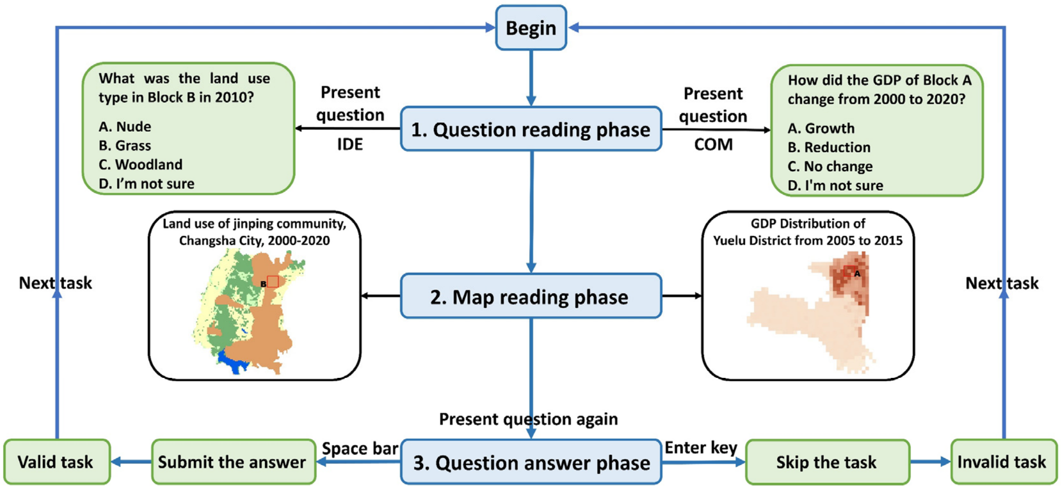

Figure 6. Procedure of the three task phases.

- **Question-reading phase.** For each task, the question and its four possible answers were first displayed in the center of the screen. According to the description of the questions, these tasks were divided into two types: single-layer identification tasks (IDE) and multilayer comparison tasks (COM). For example, "How was the land use type of block B in 2010? (Identification task)" and "How did the GDP of Block A change from 2000 to 2020? (Comparison task)". In this phase, participants had enough time to read the question and the choices, and then they could press the space bar to switch to the map-reading phase.
- **Map-reading phase.** In this phase, participants had time to read the map associated with the task. Meanwhile, they also needed to collect the grid information using different identification methods to complete the tasks. Once participants felt they found the answers, they were required to press the space bar to switch to the question-answer phase as soon as possible.
- **Question-answer phase**. In the question-answer phase, the task question and the four possible answers were displayed on the screen again. Participants had enough time to consider and make their choice and then submit it by pressing the space bar. In addition, participants were asked to speak their answers aloud before submitting them. This was to ensure that the choice they submitted was what they were thinking to avoid misoperations during this phase. Note that participants could also press the enter key to skip the task if they forgot the answer or for any other reason. Whether participants pressed the space bar or the enter key, the next task was presented.

4.1.4. Procedure

First, we gave participants a brief overview of the experiment and the eye tracker technology when they arrived at the lab. Then, participants were asked to sit in front of a laptop at a distance of approximately 50 cm. Next, participants could adjust their sitting posture to a comfortable position to calibrate the eye tracker. The calibration was conducted with a 6-point calibration method supported by Tobii Experience software 1.69.32600. In addition, a recalibration could be conducted during the experiment when necessary. After calibration, there were some training tasks for participants to become familiar with the experiment and the identification operation for each method (TR, GF, GD and MD). When the participants were prepared, they could proceed to the formal experiment.

In the formal experiment, participants were asked to complete a series of identification and comparison tasks. For each adaptation method, we designed 18 tasks, including 12 comparison tasks and 6 identification tasks, resulting in a total of 72 tasks (4 adaptation methods $\times$ 18 tasks). We used a within-subject design, meaning that each participant needed to perform all 72 tasks. The order of the 4 adaptation methods was counterbalanced based on a Latin square and in each method, the order of the 18 tasks was randomized. The participants' gaze data were recorded in all tasks. During the experiment, the experimenter monitored the experimental process without disturbing the participants.

At the end of the experiment, participants were required to complete two sets of questionnaires: the NASA Task Load Index (NASA-TLX) [44] and a User Experience Questionnaire (UEQ) [45]. The scores for the NASA-TLX and UEQ both ranged from 1 (low) to 7 (high) to measure the task load and user experience of participants, respectively. The NASA-TLX consists of six indicators: mental demand, physical demand, temporal demand, operational performance, effort and frustration. The UEQ consists of five indicators: attractiveness, perspicuity (i.e., how easy it is to learn to use the system), efficiency, stimulation (i.e., whether there is a motivation to use a system) and novelty. Finally, participants were encouraged to express their feelings and provide advice on the gaze adaptations through an open question.

4.2. Data Quality Check

A total of 2232 (31 participants × 72 tasks) task trials (recordings) were collected from our experiment. After checking the data quality, we abandoned 29 invalid trials that were skipped by participants, resulting in a total of 2203 valid trials. The results of the data quality check are shown in Table 1.

Table 1. The number of valid and invalid trials in different methods.

	Task Trials	
Method	**Valid Trials**	**Invalid (Skipped) Trials**
Traditional identification (TR)	555	3
Gaze fixed adaptation (GF)	542	16
Gaze dynamic adaptation (GD)	548	10
Mouse dynamic adaptation (MD)	558	0

4.3. Metrics

4.3.1. Efficiency

We calculated the average task time (s) that the participants spent in the map-reading phase for each task to measure efficiency.

4.3.2. Effectiveness

During the experiment, we also recorded whether the participant submitted the correct choice in the task. For each participant, we calculated the correct rate (%) under different identification methods by dividing the number of correct tasks by the number of his or her valid tasks.

4.3.3. Visual Behavior

To assess the impact of different methods on visual attention, we used the four eye tracking parameters listed below. All the metrics were processed using Python.

- **Mean fixation duration**. Fixation occurs when the gaze focuses on a target and remains relatively still for a period. The fixation duration (milliseconds, ms, for single fixations) indicated how long a fixation lasted. According to Goldberg and Kotval [46], the fixation duration was closely associated with one's interpretation process of visual information. In this study, a longer fixation duration is considered as greater difficulty comprehending visual information. Fixations were obtained using Tobii Interactor APIs (see Section 3.2 for more details).
- **Saccade amplitude**. A saccade is a rapid eye movement from one fixation to another. The saccade amplitude (screen pixels, px) describes the distance of a saccade (between two adjacent fixations) and is related to the efficiency of a visual search [47,48].

- **Proportion of fixation duration on the layer panel**. The layer panel shows the most basic information (e.g., colormap and value labels) of different layers. Since gaze adaptations were utilized in the GF, GD and MD, participants could obtain the grid information without paying attention to the layer panel. Therefore, we first created an area of interest (AOI) on the layer panel and found the fixations that fell in the AOI. We then calculated the proportion of fixation duration on the layer panel to investigate how participants' attention to the layer panel changes in different methods [49].
- **Minimum gaze bounding area**. The minimum gaze bounding area is the area of the smallest convex polygon enclosing all the gaze points of a participant in a task. It is an extensiveness measure that denotes the on-screen search breadth. Combined with the saccade amplitude metric, we can determine whether a visual search covers a broader area or is limited to a smaller region [46]. We first found the convex hull of the gaze points and then calculated the area of the convex hull. This was realized using the Python Scipy ConvexHull function (https://docs.scipy.org/doc/scipy/reference/generated/scipy.spatial.ConvexHull.html, accessed on 5 September 2023) and Shapely (https://shapely.readthedocs.io/en/stable/reference/shapely.area.html, accessed on 5 September 2023).

4.3.4. Questionnaire

We used the NASA-TLX to measure the task load and UEQ to measure the user experience of different methods (described in the Procedure section). The answers to the open questions were also analyzed.

5. Results

For the metrics described in Section 4.3, we applied the following statistical methods to analyze the results. First, we used nonparametric statistical tests since these data do not conform to the Gaussian normal distribution. For multiple comparisons, the Kruskal–Wallis test was employed. The Mann–Whitney U test was used to determine the significance of the differences between two groups and the effect size was reported as *Cohen's d* value (d) [50]. The statistical tests were performed using IBM SPSS Statistics v26 (IBM, USA, https://www.ibm.com/cn-zh/spss, accessed on 5 September 2023).

5.1. Efficiency and Effectiveness

As seen in Figure 7a, there was a significant main effect of the adaptation method on efficiency ($p = 0.000 < 0.001$). Participants using mouse dynamic adaptation had the shortest task time for both the identification and comparison tasks (IDE: M = 3.13 s, SD = 1.56 s; COM: M = 3.35 s, SD = 2.11 s). For both types of tasks, participants using the gaze fixed (IDE: M = 7.48 s, SD = 7.91 s; COM: M = 6.46 s, SD = 5.16 s) and gaze dynamic adaptation (IDE: M = 5.91 s, SD = 5.63 s; COM: M = 5.89 s, SD = 4.36 s) were significantly faster than those using the traditional method (all $ps = 0.000 < 0.001$, GF and TR: IDE: $d = 0.20$, COM: $d = 1.06$; GD and TR: IDE: $d = 0.52$, COM: $d = 1.18$), but slower than those using mouse dynamic adaptation (all $ps = 0.000 < 0.001$, GF and MD: IDE: $d = 0.77$, COM: $d = 0.79$; GD and MD: IDE: $d = 0.70$, COM: $d = 0.73$). However, there was no significant difference between the gaze fixed and gaze dynamic methods for either the identification or comparison tasks. As expected, the traditional method had the lowest efficiency (IDE: M = 8.83 s, SD = 5.59 s; COM: M = 14.87 s, SD = 9.78 s).

As shown in Figure 7b, for all methods, the mean correct rate was above 85%. There was a significant main effect of the adaptation method on effectiveness ($p = 0.001 < 0.01$). Specifically, mouse dynamic adaptation reached the highest correct rate compared to the other methods for both types of tasks (IDE: M = 95.39%, SD = 10.03%; COM: M = 95.89%, SD = 5.67%). Interestingly, the performance of the other three methods exhibited inconsistency between the IDE and COM tasks. For the identification tasks, the mean correct rate for gaze dynamic adaptation was 85.41% (SD = 12.46%), which was significantly lower than that of the traditional method ($p = 0.006 < 0.01$, $d = 0.77$) and lower than that of mouse dynamics ($p = 0.001 < 0.01$, $d = 0.88$). There is no significant difference between the gaze fixed and gaze dynamic methods. For the comparison tasks, the three adaptive methods (GF, GD and MD) all had higher correct rates than TR. The correct rate of mouse dynamics was significantly higher than that of the traditional method ($p = 0.002 < 0.01$, $d = 0.90$).

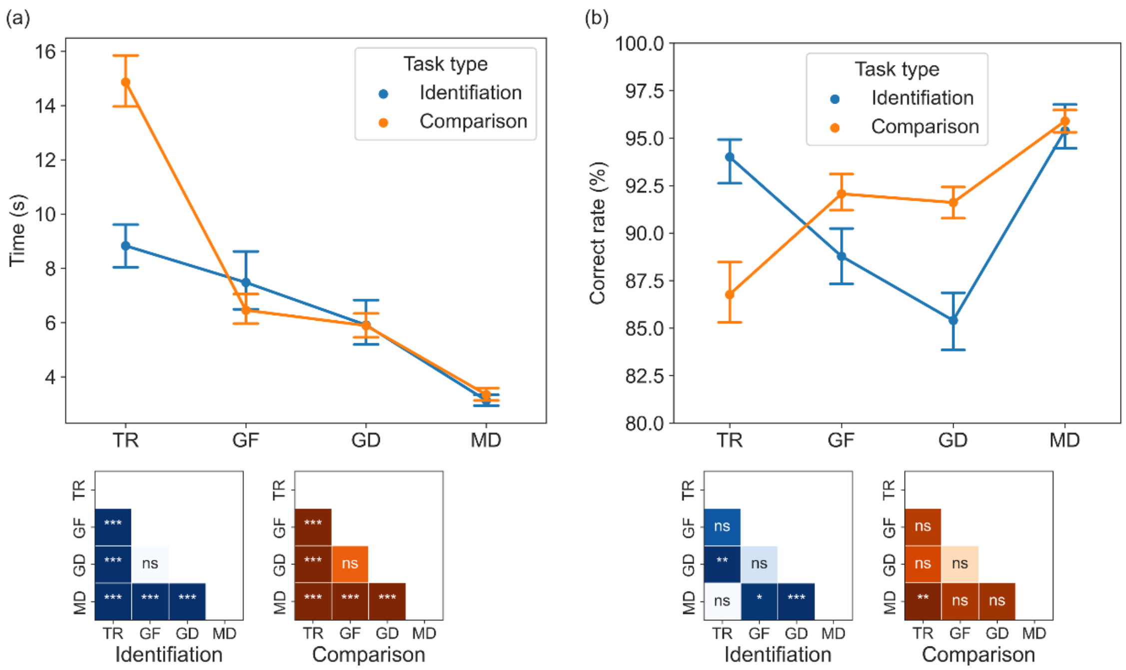

Figure 7. Mean task time (**a**) and correct rate (**b**) of different methods. Note: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

5.2. Visual Behavior

As shown in Figure 8, there is a significant main effect of the adaptation method on both the fixation duration and saccade amplitude (all $ps = 0.000 < 0.001$). The fixation duration using the gazed fixed method was the longest in both the IDE (M = 1404.70 ms, SD = 1451.64 ms) and COM tasks (M = 1386.17 ms, SD = 1573.58 ms), which was significantly longer than that using the traditional method (IDE: M = 946.42 ms, SD = 955.11 ms, $p = 0.000 < 0.001$, $d = 0.39$; COM: M = 902.52 ms, SD = 881.70 ms, $p = 0.000 < 0.001$, $d = 0.44$) and the mouse dynamic method (IDE: M = 1180.76 ms, SD = 1098.69 ms, $p = 0.033 < 0.05$, $d = 0.17$; COM: M = 1112.37 ms, SD = 1077.95 ms, $p = 0.01 < 0.05$, $d = 0.20$). Furthermore, the fixation duration when using the gaze dynamic method (IDE: M = 1269.11 ms, SD = 1285.34 ms; COM: M = 1357.72 ms, SD = 1345.57 ms) was also significantly longer than that when using the traditional method in both tasks (IDE: $p = 0.000 < 0.01$, $d = 0.30$; COM: $p = 0.000 < 0.001$, $d = 0.45$) and was significantly longer than that when using the mouse dynamic method in the COM tasks ($p = 0.000 < 0.001$, $d = 0.20$). In addition, no significant difference was found between the gaze fixed and gaze dynamic methods for both task types in the fixation duration.

As seen in Figure 8b, the traditional method (IDE: M = 784.63 px, SD = 457.81 px; COM: M = 818.60 px, SD = 478.89 px) and gaze fixed methods (IDE: M = 855.98 px, SD = 510.58 px; COM: M = 820.35 px, SD = 502.11 px) all had a significantly greater saccade amplitude than the gaze dynamic (IDE: M = 433.78 px, SD = 306.56 px; COM: M = 441.19 px, SD = 326.25 px) and mouse dynamic methods (IDE: M = 408.89 px, SD = 367.39 px; COM: M = 389.43 px, SD = 301.74 px) in both identification (all $ps = 0.000 < 0.001$, TR and GD: $d = 0.85$; TR and MD: $d = 0.85$; GF and GD: $d = 0.99$; GF and MD: $d = 0.94$) and comparison tasks (all $ps = 0.000 < 0.001$, TR and GD: $d = 0.84$; TR and MD: $d = 0.93$; GF and GD: $d = 0.89$; GF and MD: $d = 0.97$). In summary, the position adaptation in the gaze and mouse dynamic methods reduced the saccade amplitude from approximately 800 px to below 500 px.

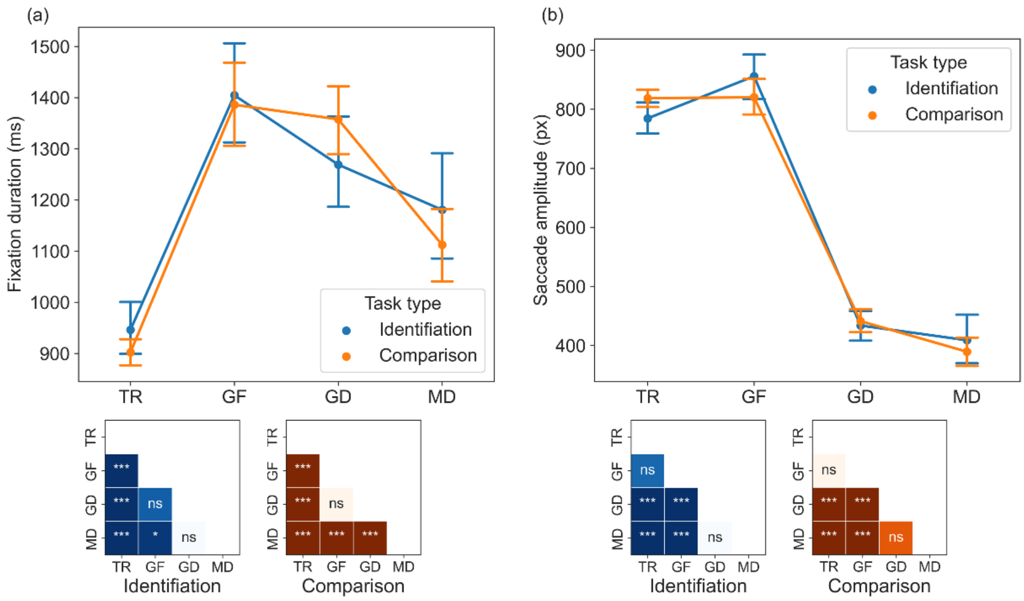

Figure 8. Mean fixation duration (**a**) and saccade amplitude (**b**) of different methods. Note: *: $p < 0.05$, ***: $p < 0.001$.

We also calculated the mean proportion of the fixation duration on the layer panel (Figure 9a). Different adaptation methods had a significant main effect on the proportion of the fixation duration ($p = 0.000 < 0.001$). When using the traditional method, participants needed to switch the visible layer in the layer panel and then use the identify tool to interpret the grid information. Therefore, the proportion of the fixation duration on the layer control of the traditional method is significantly higher (IDE: M = 0.41, SD = 0.22; COM: M = 0.43, SD = 0.21) than that of the other three methods in both identification (GF: M = 0.25, SD = 0.18, $p = 0.000 < 0.001$, $d = 0.76$; GD: M = 0.11, SD = 0.18, $p = 0.000 < 0.001$, $d = 1.35$; MD: M = 0.06, SD = 0.05, $p = 0.001 < 0.01$, $d = 1.56$) and comparison tasks (GF: M = 0.26, SD = 0.20, $p = 0.000 < 0.001$, $d = 0.78$; GD: M = 0.06, SD = 0.04, $p = 0.000 < 0.001$, $d = 1.79$; MD: M = 0.05, SD = 0.04, $p = 0.009 < 0.01$, $d = 1.77$).

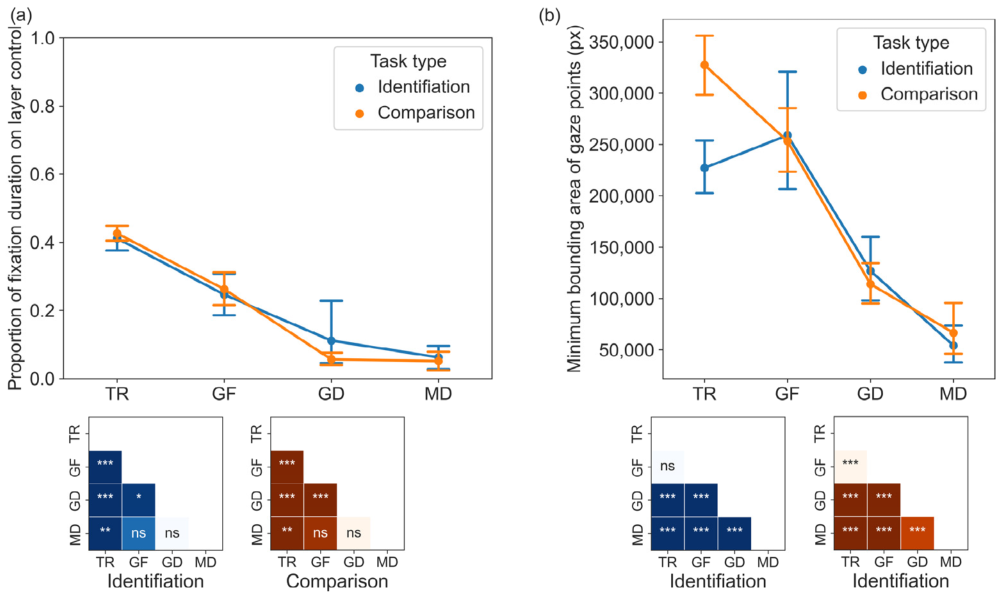

Figure 9. Mean proportion of fixation duration on the layer panel (**a**) and minimum gaze point bounding area (**b**) of different methods. Note: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

There was also a significant main effect of the adaptation method ($p = 0.000 < 0.001$) on the minimum gaze point bounding area. Since the grid information was presented in the top-left corner of the screen in the gaze fixed method, the minimum bounding area in the IDE tasks was the largest (M = 2.6×10^4 px, SD = 3.5×10^4 px). Similarly, the traditional method had the largest minimum gaze point bounding area in the COM tasks (M = 3.3×10^4 px, SD = 2.8×10^4 px) because participants needed to switch visible layers repeatedly. Since the adaptive method was utilized, the minimum gaze point bounding area in the gaze dynamic (IDE: M = 1.3×10^4 px, SD = 1.6×10^4 px; COM: M = 1.1×10^4 px, SD = 1.6×10^4 px) and mouse dynamic methods (IDE: M = 5×10^3 px, SD = 9×10^3 px; COM: M = 7×10^3 px, SD = 1.5×10^4 px) was significantly smaller than the other two methods (all $ps = 0.000 < 0.001$) in both identification (GD and TR: $d = 0.59$; GD and GF: $d = 0.46$; MD and TR: $d = 1.13$; MD and GF: $d = 0.56$) and comparison tasks (GD and TR: $d = 0.89$; GD and GF: $d = 0.68$; MD and TR: $d = 1.03$; MD and GF: $d = 0.87$).

5.3. NASA-TLX and UEQ

As shown in Figure 10, different methods had consistent score distributions (the scores of TR > GF > GD > MD) for all indexes in NASA-TLX (the lower the better). Mouse and gaze dynamic adaptations were both rated significantly lower than the traditional method in mental demand, physical demand, temporal demand and effort. Except for frustration (MD: M = 2.00, SD = 1.15; GD: M = 3.10, SD = 1.33, $p = 0.023 < 0.05$, $d = 0.23$), no significant difference was found in the mouse and gaze dynamic methods in other indexes. This illustrated the similarities between the gaze and mouse dynamic methods in many aspects. Among all six indexes in NASA-TLX, participants rated the four methods most differentially in their temporal demand. The traditional method (M = 5.52, SD = 1.73) was rated significantly higher than the other methods (GF: M = 3.65, SD = 1.31, $p = 0.003 < 0.01$, $d = 1.22$; GD: M = 3.26, SD = 1.18, $p = 0.000 < 0.001$, $d = 1.52$; MD: M = 2.48, SD = 1.41, $p = 0.000 < 0.001$, $d = 1.92$) and the gaze fixed method was rated significantly higher than the mouse dynamic method ($p = 0.028 < 0.05$, $d = 0.85$). For the two gaze interaction methods, no significant difference was found in NASA-TLX.

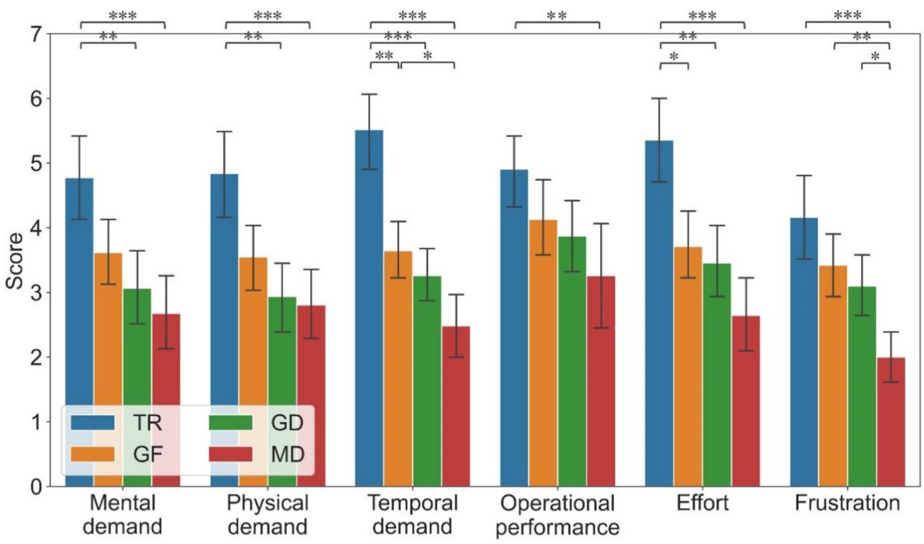

Figure 10. The results of NASA-TLX. Note: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

The results of the UEQ are shown in Figure 11. Since TR received the lowest scores in all five indexes, it performed the worst in both the UEQ and NASA-TLX. This was not surprising because it is the most cumbersome compared with the other three improvement methods. Different from NASA-TLX, participants gave the highest scores for the gaze dynamic method in all indexes except efficiency in the UEQ. For the two gaze interaction methods, there was a significant difference between the gaze dynamic and gaze fixed methods only in attractiveness (GF: M = 4.94, SD = 1.36; GD: M = 6.06, SD = 0.81, $p = 0.014 < 0.05$, $d = 1.00$). Furthermore, GD (M = 6.68, SD = 0.60) and GF (M = 6.00, SD = 0.77) were both rated significantly higher than MD (M = 6.68, SD = 0.60; MD and GD: $p = 0.000 < 0.001$, $d = 2.71$; MD and GF: $p = 0.000 < 0.001$, $d = 1.91$) in novelty.

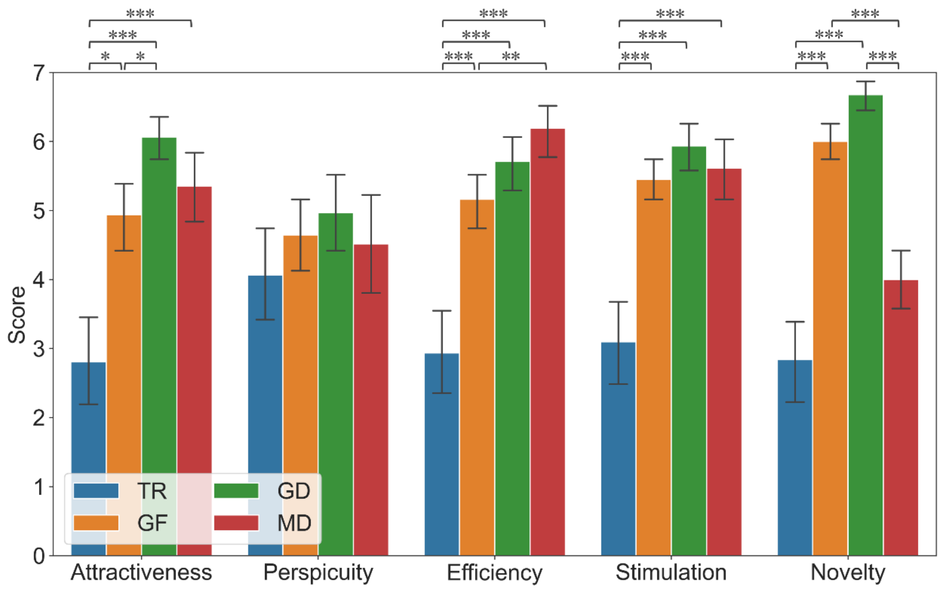

Figure 11. The results of UEQ. Note: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

5.4. User Feedback

After the experiment, 29 participants expressed their opinions and comments. To better understand participants' perceptions, we classified their answers according to their preferences for the two gaze methods. Of all 29 answers, 10 participants showed their preference for the gaze dynamic method, 3 participants preferred the gaze fixed method, and the other 16 participants stated no preference for the two gaze methods. Participants most preferred the gaze dynamic method for "*convenience*" (P4, P8, P14, P18, P25, P31) because they did not need to switch visual attention dramatically. They also mentioned that the gaze dynamic method was "*more efficient*" (P7, P14, P20, P21) and required "*less work*" (P15, P21). However, some participants also struggled with the dynamic information window due to its "*less intelligent flexibility*" (P1). For instance, P3 stated that "*it was difficult to focus because the window may move with my gaze simultaneously sometimes*". In contrast, the gaze fixed method was preferred due to its "*simple logic*" (P3) and "*would not block the map extension compared with the gaze dynamic method*" (P27).

For those participants who showed no preference between the two gaze methods, P17 stated that "*since the target area was small, adjusting the shaky gaze indicator to point at the target took me more time*". Meanwhile, they mostly argued that "*the accuracy of gaze methods was not sufficient*" (P2, P9, P10, P11, P19, P23, P30) to apply to precise interactions in such experiments. Participants also recommended "*increasing the sensitivity*" of the gaze methods (P5, P6, P7, P12, P29). Other comments mainly focused on "*visual fatigue*" (P13), "*light conditions*" (P22), "*need more concentration*" (P26) and "*more unskilled than using a mouse*" (P28).

6. Discussion

6.1. Performance and Visual Behavior

Although the number of participants ($n = 31$) generally aligns with many other eye tracking experiments in cartography and GIScience (e.g., [51–54]), researchers should be careful when interpreting the findings due to the possibility of statistical fluctuations caused by the small sample size. In addition, the limited sample size could decrease the robustness of the results. Therefore, this study is considered a qualitative research study rather than a quantitative one.

The results revealed that the gaze and mouse adaptation methods (GF, GD and MD) were all faster than the traditional identification method in both the single-layer identification (IDE) and multilayer comparison (COM) tasks. In particular, the adaptation methods showed almost the same efficiency in IDE and COM tasks.

However, the gaze-based adaptation methods were still 2 to 3 s slower than the mouse dynamic method in both IDE and COM tasks. This is probably because participants needed extra time to stabilize the 'shaky' gaze indicator (i.e., the '+' indicator showing current gaze position) after they located the target when using gaze interaction. Fundamentally, such shakiness was also a manifestation of the low spatial tracking precision of gaze interaction, which was the most commonly mentioned issue in user feedback. The low spatial tracking precision, further, also leads to longer fixation durations on average for the two gaze adaptation methods. Participants must be cautious in avoiding any shakiness during each fixation, which can prolong the time needed for the fixation and cause visual fatigue according to the feedback. Therefore, some supplementary methods, such as the target acquisition technique [55], seem to be necessary for target selection or pinpointing gaze interaction.

For the correct rate in both the IDE and COM tasks, the mouse dynamic method performed the best, with a stable mean correct rate of more than 95%. The gaze adaptation methods exhibited differences between the IDE and COM tasks. In the COM tasks, there was a slight improvement of the two gaze adaptation methods compared to the traditional method, but such an improvement was not significant. Compared to the traditional method in the IDE tasks, the correct rates of both the gaze fixed and dynamic methods significantly decreased and were even lower than their correct rates in the COM tasks. This was surprising since IDE tasks were theoretically easier than COM tasks because participants only needed to focus on the information of a particular layer. We speculated that when the information of all layers was presented, participants were easily disturbed by misinformation from nontarget layers. Since participants needed to read the map and filter out useless information while interacting using their eyes, the additional cognitive workload made gaze methods less effective. In future research, therefore, minimizing the impact of redundant information on users' cognition will also be a key issue.

6.2. Comparison between Identification and Comparison Tasks

It is noted that both the adaptation approach (i.e., GD, GF, MD and TR) and the task type (i.e., IDE and COM) are two independent variables in our experiment. However, the main focus of this study was to compare the performance of different adaptation approaches and to test whether their performance was consistent across the two task types. We summarize two observations, regarding the task type, that merit further discussion as follows.

First, we observed that the two task types exhibited consistent influences on the participants' visual behavior (see Figures 8 and 9). One exception is that when using the traditional method, comparison tasks resulted in a significantly larger minimum bounding area than the identification tasks (Figure 9b). This was because comparison tasks required more visual switches between the layer control panel and the information window than identification tasks.

Second, participants using the traditional method spent a significantly longer time and achieved a lower correct rate in the comparison tasks than in the identification tasks (Figure 7). This is consistent with the above observation that more visual switches were required for the comparison tasks than for the identification tasks. In addition, using the traditional *identify* tool for the comparison tasks meant needing to turn on/off the layers frequently and memorize the raster values, which could lead to a higher cognitive load (Figure 10) and lower correct rate.

The above two observations imply that for both the identification and comparison tasks, the placement adaptation of the dynamic information window could reduce the saccade amplitude and visual search area, leading to a shorter task time. However, the legend content adaptation did not improve the correct rate of the gaze-based adaptations methods in the identification tasks. As discussed in Section 6.1, this was probably due to the redundant information provided by the dynamic information window.

6.3. Design Issues

As mentioned in Section 3.1, in the gaze dynamic adaptation method, the dynamic information window is always visible. We consider that displaying the window directly is intuitive and can enable users to see the information immediately. It seems that the participants could get used to this method easily during the training session. No usability comment on this design was reported in the questionnaire. Our method differs from Göbel et al. 's [39] implementation, which requires users to first fixate on a circle legend proxy to unfold the legend. When the user's gaze leaves the legend, it changes to the proxy again. In our method, when the dynamic window is fixed, its content and position will remain static. A disadvantage of our method is that it may block the region where the user wants to see. To see the grid below the current information window, the user needs to first leave the window and then return to the region. Therefore, both our method and Göbel et al.'s approach have pros and cons. Further experiments are required to evaluate their performance and user preferences.

6.4. Limitation

In this experiment, we devised tasks to simulate the demand of exploring the local features of the raster map. Based on the tracking accuracy of current devices and the recommended control size from prior research on gaze interaction [35,56], we established a size of 1° for each block that participants were required to observe in this experiment. Meanwhile, we set the spatial resolution of the map to a lower level to ensure that the grid values were consistent within each block, thereby reducing potential ambiguity for participants during the task.

Nevertheless, in real-world grid-reading scenarios, the grid resolution may vary depending on changes in the data sources and map themes. Thus, it is rare for all grid values to be the same even within the same block. Moreover, with the change in the usage scenarios of raster maps, the objects of interest in our research will also be different. For instance, we may need to explore not only the local features (i.e., blocks in our experiment), but also the specific features of individual grid cells, such as their respective values, in certain situations. However, due to the inherent low spatial accuracy of gaze interaction, the methods proposed in this study may not satisfy such a demand. A possible solution is to use a two-step interaction method. For example, the methods proposed in this study can be used first to observe local features and then utilize a zoom-in window with a pinpointing method to explore more detailed features.

As mentioned in Section 6.1, another limitation of this study is the small sample size ($n = 31$), which could restrict the applicability of the findings to wider populations. If more participants are included, the outcomes might change. In addition, different characteristics (e.g., age, gender, map reading ability and normal/corrected-to-normal visual and visual impairments) may affect individuals' performance and visual behavior. The sample size of the study limits the generalizability of the results to other individual groups. In the future, it would be interesting to expand the sample size to explore the influence of user factors on gaze interaction.

Familiarity with gaze interaction might also affect the participants' performance and visual behavior. Although the participants were trained before the formal experiment, it cannot be guaranteed that the participants achieved the same level of familiarity or proficiency of using gaze as of using a mouse. The current experiment could be further improved by adding an evaluation of the participants' familiarity of using gaze/mouse interactions to control the familiarity- or skill-level variance.

7. Conclusions and Future Work

This study explored the potential benefits and drawbacks of gaze-adaptive interactions when reading raster maps. We focused on the usability of the use of low-cost eye trackers on gaze-based interactions. We designed two gaze adaptive methods, gaze fixed (GF) and gaze dynamic adaptations (GD) for identifying and comparing raster values between multilayers. In both methods, the grid legend content of different layers is adaptively adjusted depending on the user's visual focus. To include a high-precision method as a benchmark, we used a mouse to replace the gaze but kept other settings unchanged in the mouse dynamic adaptation (MD). We further conducted a user study by comparing such adaptation methods with the traditional method (TR) through a series of single-layer identification tasks (IDE) and multilayer comparison tasks (COM). We summarize the contributions of this study as follows:

1. Compared to the traditional method, both gaze- and mouse-based adaptations can significantly enhance user efficiency in both the identification and comparison tasks. However, the gaze-based adaptations (GF and GD) had lower efficiency and effectiveness than the mouse dynamic adaptation in both tasks. In the identification tasks, the gaze-based methods even exhibited lower effectiveness than the traditional method. This is probably because the gaze-adaptive legends that contained three layers (i.e., redundant information exists) may confuse the participants when the participants intended to focus on only one certain layer.
2. Despite incorporating both content and placement adaptations, the gaze dynamic method exhibited inferior efficiency compared to the mouse dynamic method. This is primarily due to the lower spatial tracking precision of the low-cost eye tracker which led to longer average fixation durations and visual fatigue. This is the most commonly mentioned issue in the user feedback.
3. Different adaptation methods resulted in different visual behavior characteristics. First, participants switched their visual focus to the layer content panel considerably less under the adaptive methods (GF, GD and MD) than under the traditional method, as we predicted. Second, when using methods with placement adaptation (GD and MD), participants' visual searches covered smaller regions than those without placement adaptation (TR and GF). Third, when using methods based on gaze interaction (GF and GD), participants had longer fixation durations than those using a mouse (TR and MD).
4. The gaze-adaptive methods (GF and GD) were generally well received by the participants, but they were also perceived to be somewhat distracting and insensitive. However, it did not seem to hinder performance or the user experience in this study, but left further improvement to reduce the negative perceptions.

There are also several other steps of our future research for further investigation. First, the content contained in the adaptive legend should be designed more intuitively. This can help users reduce the cognitive workload of gaze interaction without becoming confused. This is especially true in situations where users need to focus on only the grid information of a certain layer. Second, all the maps used in this study consisted of three raster layers with low spatial resolutions. Thus, we limited the difficulty and scenarios of raster map reading in our user study. Due to the potential variability in the performance of different methods under different experimental settings, we aim to extend our research to more cases of map-reading scenarios with varying difficulty levels in the future. Finally, mouse movement data were recorded to conduct a more quantitative analysis of the gaze dynamic and mouse dynamic methods. Therefore, mouse tracking and quantitative dynamic interactions [57] between the tracing of the gaze and the mouse are required in the mouse dynamic method to provide more specific improvement suggestions for gaze dynamic adaptations and for gaze-based map interaction.

Author Contributions: Conceptualization, Hua Liao, Changbo Zhang and Weihua Dong; Funding acquisition, Hua Liao; Project administration, Hua Liao; Software, Yongbo Huang; Supervision, Hua Liao and Weihua Dong; Validation, Hua Liao; Writing—review and editing, Changbo Zhang and Hua Liao. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the National Natural Science Foundation of China [NSFC, Grant Nos. 42371455, 42001410 and 42230103] and the Natural Science Foundation of the Hunan Province [Grant No. 2021JJ40350].

Data Availability Statement: The data supporting the findings of this investigation are accessible from the author upon reasonable request. The raw data are not available to the public because they contain information that could compromise the participants' privacy.

Acknowledgments: The authors are grateful for the comments from the reviewers, which helped improve the article's quality. The authors thank all the participants in the experiment.

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

References

1. Robinson, A.H.; Morrison, J.L.; Muehrcke, P.C.; Kimmerling, A.J.; Guptil, S.C. Elements of cartography. *Geod. List* **1995**, *50*, 408–409.
2. Kubicek, P.; Sasinka, C.; Stachon, Z.; Sterba, Z.; Apeltauer, J.; Urbanek, T. Cartographic Design and Usability of Visual Variables for Linear Features. *Cartogr. J.* **2017**, *54*, 91–102. [CrossRef]
3. Bednarik, R.; Vrzakova, H.; Hradis, M. What do you want to do next: A novel approach for intent prediction in gaze-based interaction. In Proceedings of the Symposium on Eye Tracking Research and Applications, Santa Barbara, CA, USA, 28–30 March 2012; pp. 83–90.
4. Liao, H.; Dong, W.H.; Huang, H.S.; Gartner, G.; Liu, H.P. Inferring user tasks in pedestrian navigation from eye movement data in real-world environments. *Int. J. Geogr. Inf. Sci.* **2019**, *33*, 739–763. [CrossRef]
5. David-John, B.; Peacock, C.; Zhang, T.; Murdison, T.S.; Benko, H.; Jonker, T.R. Towards gaze-based prediction of the intent to interact in virtual reality. In Proceedings of the ACM Symposium on Eye Tracking Research and Applications, Virtual Event, 24–27 May 2021; pp. 1–7, Article 2.
6. Chen, X.; Hou, W. Gaze-Based Interaction Intention Recognition in Virtual Reality. *Electronics* **2022**, *11*, 1647. [CrossRef]
7. Stachoň, Z.; Šašinka, Č.; Čeněk, J.; Angsüsser, S.; Kubíček, P.; Štěrba, Z.; Bilíková, M. Effect of Size, Shape and Map Background in Cartographic Visualization: Experimental Study on Czech and Chinese Populations. *Isprs Int. J. Geo Inf.* **2018**, *7*, 427. [CrossRef]
8. Cybulski, P.; Krassanakis, V. The effect of map label language on the visual search of cartographic point symbols. *Cartogr. Geogr. Inf. Sci.* **2022**, *49*, 189–204. [CrossRef]
9. Keskin, M.; Ooms, K.; Dogru, A.O.; De Maeyer, P. Exploring the Cognitive Load of Expert and Novice Map Users Using EEG and Eye Tracking. *Isprs Int. J. Geo Inf.* **2020**, *9*, 429. [CrossRef]
10. Popelka, S.; Herman, L.; Reznik, T.; Parilova, M.; Jedlicka, K.; Bouchal, J.; Kepka, M.; Charvat, K. User Evaluation of Map-Based Visual Analytic Tools. *Isprs Int. J. Geo Inf.* **2019**, *8*, 363. [CrossRef]
11. Edler, D.; Keil, J.; Tuller, M.C.; Bestgen, A.K.; Dickmann, F. Searching for the 'Right' Legend: The Impact of Legend Position on Legend Decoding in a Cartographic Memory Task. *Cartogr. J.* **2020**, *57*, 6–17. [CrossRef]
12. Duchowski, A.T. Gaze-based interaction: A 30 year retrospective. *Comput. Graph.* **2018**, *73*, 59–69. [CrossRef]
13. Kasprowski, P.; Harezlak, K.; Niezabitowski, M. Eye movement tracking as a new promising modality for human computer interaction. In Proceedings of the 17th International Carpathian Control Conference (ICCC), High Tatras, Slovakia, 29 May–1 June 2016; pp. 314–318.
14. Singh, R.; Miller, T.; Newn, J.; Velloso, E.; Vetere, F.; Sonenberg, L. Combining gaze and AI planning for online human intention recognition. *Artif. Intell.* **2020**, *284*, 103275. [CrossRef]
15. Fairbairn, D.; Hepburn, J. Eye-tracking in map use, map user and map usability research: What are we looking for? *Int. J. Cartogr.* **2023**, *9*, 1–24. [CrossRef]
16. Ooms, K.; Krassanakis, V. Measuring the Spatial Noise of a Low-Cost Eye Tracker to Enhance Fixation Detection. *J. Imaging* **2018**, *4*, 96. [CrossRef]
17. Jacob, R.J.K. The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Trans. Inf. Syst.* **1991**, *9*, 152–169. [CrossRef]
18. Ware, C.; Mikaelian, H.H. An evaluation of an eye tracker as a device for computer input2. In Proceedings of the SIGCHI/GI Conference on Human Factors in Computing Systems and Graphics Interface, Toronto, ON, Canada, 1 May 1986; pp. 183–188.
19. Zhang, H.; Hu, Y.; Zhu, J.; Fu, L.; Xu, B.; Li, W. A gaze-based interaction method for large-scale and large-space disaster scenes within mobile virtual reality. *Trans. GIS* **2022**, *26*, 1280–1298. [CrossRef]
20. Piumsomboon, T.; Lee, G.; Lindeman, R.W.; Billinghurst, M. Exploring natural eye-gaze-based interaction for immersive virtual reality. In Proceedings of the 2017 IEEE Symposium on 3D User Interfaces (3DUI), Los Angeles, CA, USA, 18–19 March 2017; pp. 36–39.

21. Isomoto, T.; Yamanaka, S.; Shizuki, B. Interaction Design of Dwell Selection Toward Gaze-based AR/VR Interaction. In Proceedings of the 2022 Symposium on Eye Tracking Research and Applications, Seattle, WA, USA, 8 June 2022; pp. 31–32, Article 39.
22. Deng, C.; Tian, C.; Kuai, S. A combination of eye-gaze and head-gaze interactions improves efficiency and user experience in an object positioning task in virtual environments. *Appl. Ergon.* **2022**, *103*, 103785. [CrossRef] [PubMed]
23. Hirzle, T.; Gugenheimer, J.; Geiselhart, F.; Bulling, A.; Rukzio, E. A Design Space for Gaze Interaction on Head-mounted Displays. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 2 May 2019; p. 625.
24. Majaranta, P.; Ahola, U.-K.; Špakov, O. Fast gaze typing with an adjustable dwell time. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, 4 April 2009; pp. 357–360.
25. Paulus, Y.T.; Remijn, G.B. Usability of various dwell times for eye-gaze-based object selection with eye tracking. *Displays* **2021**, *67*, 101997. [CrossRef]
26. Hansen, J.P.; Johansen, A.S.; Hansen, D.W.; Itoh, K.; Mashino, S. Command without a Click: Dwell Time Typing by Mouse and Gaze Selections. In *Human-Computer Interaction INTERACT'03*; Rauterberg, M., Ed.; IOS Press: Amsterdam, The Netherlands, 2003; pp. 121–128.
27. Dunphy, P.; Fitch, A.; Olivier, P. Gaze-contingent passwords at the ATM. In Proceedings of the 4th Conference on Communication by Gaze Interaction—Communication, Environment and Mobility Control by Gaze COGAIN 2008, Prague, Czech Republic, 2–3 September 2008.
28. Feit, A.M.; Williams, S.; Toledo, A.; Paradiso, A.; Kulkarni, H.; Kane, S.; Morris, M.R. Toward Everyday Gaze Input: Accuracy and Precision of Eye Tracking and Implications for Design. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, New York, NY, USA, 2 May 2017; pp. 1118–1130.
29. Drewes, H.; Schmidt, A. Interacting with the computer using gaze gestures. In Proceedings of the IFIP Conference on Human-Computer Interaction, Rio de Janeiro, Brazil, 10 September 2007; pp. 475–488.
30. Hyrskykari, A.; Istance, H.; Vickers, S. Gaze gestures or dwell-based interaction? In Proceedings of the Symposium on Eye Tracking Research and Applications, Santa Barbara, CA, USA, 28–30 March 2012.
31. Kytö, M.; Ens, B.; Piumsomboon, T.; Lee, G.; Billinghurst, M. Pinpointing: Precise Head- and Eye-Based Target Selection for Augmented Reality. In Proceedings of the Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018.
32. Barz, M.; Kapp, S.; Kuhn, J.; Sonntag, D. Automatic Recognition and Augmentation of Attended Objects in Real-time using Eye Tracking and a Head-mounted Display. In Proceedings of the ACM Symposium on Eye Tracking Research and Applications, Virtual Event, 24–27 May 2021; pp. 1–4, Article 3.
33. Göbel, F.; Bakogioannis, N.; Henggeler, K.; Tschümperlin, R.; Xu, Y.; Kiefer, P.; Raubal, M. A Public Gaze-Controlled Campus Map. In Proceedings of the Eye Tracking for Spatial Research, Proceedings of the 3rd International Workshop, ETH, Zurich, Switzerland, 14 January 2018.
34. Zhu, L.; Wang, S.; Yuan, W.; Dong, W.; Liu, J. An Interactive Map Based on Gaze Control. *Geomat. Inf. Sci. Wuhan Univ.* **2020**, *45*, 736–743. (In Chinese) [CrossRef]
35. Liao, H.; Zhang, C.B.; Zhao, W.D.; Dong, W.H. Toward Gaze-Based Map Interactions: Determining the Dwell Time and Buffer Size for the Gaze-Based Selection of Map Features. *Isprs Int. J. Geo-Inf.* **2022**, *11*, 127. [CrossRef]
36. Bektaş, K.; Çöltekin, A.; Krüger, J.; Duchowski, A.T.; Fabrikant, S.I. GeoGCD: Improved visual search via gaze-contingent display. In Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications, Denver, CO, USA, 25–28 June 2019; pp. 1–10, Article 84.
37. Giannopoulos, I.; Kiefer, P.; Raubal, M. GeoGazemarks: Providing gaze history for the orientation on small display maps. In Proceedings of the 14th ACM international conference on Multimodal interaction, Santa Monica, CA, USA, 22 October 2012; pp. 165–172.
38. Tateosian, L.G.; Glatz, M.; Shukunobe, M.; Chopra, P. GazeGIS: A Gaze-Based Reading and Dynamic Geographic Information System. In Proceedings of the ETVIS 2015: Eye Tracking and Visualization, Chicago, IL, USA, 25 October 2015; pp. 129–147.
39. Göbel, F.; Kiefer, P.; Giannopoulos, I.; Duchowski, A.T.; Raubal, M. Improving map reading with gaze-adaptive legends. In Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications, Warsaw, Poland, 14 June 2018; pp. 1–9.
40. Lallé, S.; Toker, D.; Conati, C. Gaze-Driven Adaptive Interventions for Magazine-Style Narrative Visualizations. *IEEE Trans. Vis. Comput. Graph.* **2021**, *27*, 2941–2952. [CrossRef]
41. Barral, O.; Lallé, S.; Iranpour, A.; Conati, C. Effect of Adaptive Guidance and Visualization Literacy on Gaze Attentive Behaviors and Sequential Patterns on Magazine-Style Narrative Visualizations. *ACM Trans. Interact. Intell. Syst.* **2021**, *11*, 1–46. [CrossRef]
42. Keskin, M.; Kettunen, P. Potential of eye-tracking for interactive geovisual exploration aided by machine learning. *Int. J. Cartogr.* **2023**, *9*, 1–23. [CrossRef]
43. Olsen, A. The Tobii IVT Fixation Filter Algorithm Description. Available online: http://www.vinis.co.kr/ivt_filter.pdf (accessed on 18 April 2023).
44. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*; Hancock, P.A., Meshkati, N., Eds.; North Holland: Amsterdam, The Netherlands, 1988; Volume 52, pp. 139–183.

45. Laugwitz, B.; Held, T.; Schrepp, M. Construction and Evaluation of a User Experience Questionnaire. In *HCI and Usability for Education and Work*; Lecture Notes in Computer Science; Holzinger, A., Ed.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 63–76.
46. Goldberg, J.H.; Kotval, X.P. Computer interface evaluation using eye movements: Methods and constructs. *Int. J. Ind. Ergon.* **1999**, *24*, 631–645. [CrossRef]
47. Just, M.A.; Carpenter, P.A. Eye fixations and cognitive processes. *Cogn. Psychol.* **1976**, *8*, 441–480. [CrossRef]
48. Yang, N.; Wu, G.; MacEachren, A.M.; Pang, X.; Fang, H. Comparison of font size and background color strategies for tag weights on tag maps. *Cartogr. Geogr. Inf. Sci.* **2023**, *50*, 162–177. [CrossRef]
49. Jia, F.; Wang, W.; Yang, J.; Li, T.; Song, G.; Xu, Y. Effectiveness of Rectangular Cartogram for Conveying Quantitative Information: An Eye Tracking-Based Evaluation. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 39. [CrossRef]
50. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*; Academic Press: New York, NY, USA, 2013.
51. 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]
52. 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]
53. Ooms, K.; De Maeyer, P.; Fack, V. Study of the attentive behavior of novice and expert map users using eye tracking. *Cartogr. Geogr. Inf. Sci.* **2014**, *41*, 37–54. [CrossRef]
54. Kiefer, P.; Giannopoulos, I.; Raubal, M. Where Am I? Investigating Map Matching During Self-Localization with Mobile Eye Tracking in an Urban Environment. *Trans. GIS* **2014**, *18*, 660–686. [CrossRef]
55. Grossman, T.; Balakrishnan, R. The bubble cursor: Enhancing target acquisition by dynamic resizing of the cursor's activation area. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Portland, OR, USA, 2–7 April 2005; pp. 281–290.
56. Niu, Y.F.; Gao, Y.; Zhang, Y.T.; Xue, C.Q.; Yang, L.X. Improving Eye-Computer Interaction Interface Design: Ergonomic Investigations of the Optimum Target Size and Gaze-triggering Dwell Time. *J. Eye Mov. Res.* **2019**, *12*, 8. [CrossRef] [PubMed]
57. Demsar, U.; Coltekin, A. Quantifying gaze and mouse interactions on spatial visual interfaces with a new movement analytics methodology. *PLoS ONE* **2017**, *12*, e0181818. [CrossRef]