An eye-tracking study examining information search in transit maps. Using China's high-speed railway map as a case study

Wang, Z, Lonsdale MDS and Cheung, V

Zheng Wang, is a PhD student of Information and Communication Design, University of Leeds, UK. He received his Master Degree in 2017 from the School of Design, University of Leeds, UK. His research focuses on infographic design, cartographic design and data visualization.

Professor Maria dos Santos Lonsdale, PhD, is a Professor of Information and Communication Design, University of Leeds, UK. She received her PhD in 2006 from the Department of Typography and Graphic Communication, University of Reading, UK. She has experience of teaching design theory, design practice and research methods for Design in higher education. Her main areas of research are are: [more specific] Information Design and Typographic Design; [wider] Communication Design and Graphic Design. Her research is notable in the field of Design for its interdisciplinary nature, with collaboration across sectors such as Healthcare, National and Global Security, Education, etc. Her research is also notable in the field of Design for its research based user-centred approach using a wide range of research methods to develop, iterate, test and validate design solutions to particular problems encountered in real-life contexts. At its core, her research focuses on the effective use of design to improve user performance and wellbeing.

Dr Vien Cheung, PhD, is an Associate Professor in Color and Imaging Science in the School of Design, University of Leeds. She is currently the President of the International Colour Association (AIC) which comprises of 28 national colour association members representing 5 continents. Her research interests are color vision, color imaging, cross-media color reproduction and color design.

Version: Accepted Version: 25th August 2020.

Abstract

This study investigates the legibility of China's high-speed railway map through eye-tracking measurement. The information searching process was identified by conducting: 1) Scoping stage – a user performance test and interview to inform the design of the eye-tracking study; 2) In-depth stage – an eye-tracking study. Many visual design problems with the map were subsequently identified. This research helps to explore user-centered design map solutions and build up detailed design guidance for transit maps in the future. It also demonstrates that eye-tracking is an effective method to show design limitations and user needs when evaluating the design quality of a transit map.

Keywords: Information design, visualization, legibility, eye tracking, map design

1. Introduction

According to information from the Ministry of Transport of the People's Republic of China (2020), in 2019 around 60 million international tourists traveled in China by railway (95% of 63.1 million international tourists who visited China in 2019). This indicates that the railway is the most popular transportation choice for international travelers and, consequently, provides evidence of the need for a well-designed transit map of China's high-speed railway system. Due to such high demand and use, the Ministry of Transport (PRC) acknowledged the need to evaluate the design quality of China's current high-speed railway map and funded this research. The aim is to tackle any problems with the map in order to improve its effectiveness.

China's current high-speed railway map (English version) was updated in 2017 (Figure 1) and it has been widely promoted online and displayed in railway stations since then. However, questions remain regarding the legibility of the map: a) whether it is a good transit map design in terms of information structure and visual design; b) whether there are any design limitations that may affect user reading efficiency in terms of speed and accuracy; etc. As shown by a pilot test conducted for this study, legibility is compromised in this version of the transit map.

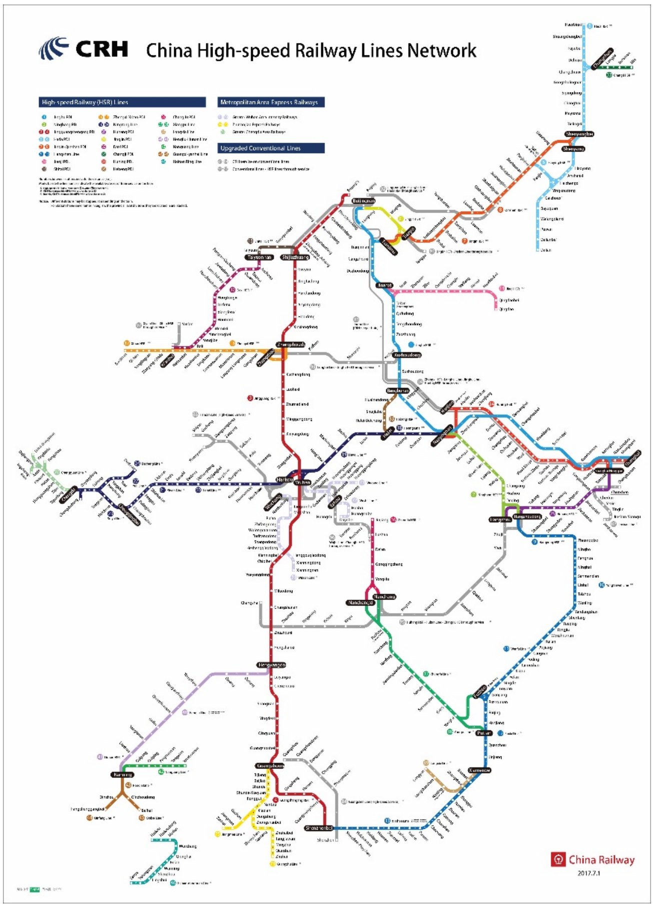


Figure 1. China's high-speed railway map

The aim of this study is to evaluate the information design structure and visual design quality of China's high-speed railway map. To this end, the following research questions have been explored:

1. How do participants find out the information they need from China's high-speed railway map according to different information needs?
2. Which design limitations can be identified in the transit map by analyzing user information searching behavior?
3. What design solutions could be implemented in the future to improve the current map design?

In order to address these questions, a user performance test was conducted using an eye tracker to investigate the potential design problems of the map and explore how the visual design of a transit map affects the information searching behavior of users. The eye tracker can record the eye-movement path of a user, its fixation points and fixation duration, and demonstrate the results with gaze plots and a heatmap. These advantages make the eye-tracking test a reliable physiological measuring method to record the reading and searching behavior of a user (Holmqvist et al., 2011). This will allow for a better understanding of human performance and perception and the underlying cognitive process. It can also be used to evaluate the legibility and usability of visual design materials (Kang et al., 2016).

There are examples of studies using eye-tracking tests to evaluate the design quality of maps, with most studies focusing on people's attention, reading speed, areas of interest, and/or people's reading strategies. For example, Brychtova and Coltekin (2016) investigated the influence of color distance and font size on map readability based on user reading speed and fixation duration. Netzel et al. (2016) compared the user's reading strategy of a map when a color system or task difficulty changed. Both studies showed that eye-tracking data can help evaluate the map's visual design quality, which provides valuable insight to inform our research methodology. However, there are very few articles focusing on how the design limitations of a map might affect user reading speed and accuracy based on user information searching behavior and how to improve these problems following a user-centered approach.

This research took China's high-speed railway map as a case study, mainly focusing on the evaluation of its design quality based on transit map design principles and information design theories. It then explored ways to improve the legibility of the map to inform the optimum design of transit maps

Modern interactive tools make a series of wayfinding decisions for users, which is both a significant aid, but also a potential for impairing the understanding of the overall network. In this case, a well-designed transit map can cater for this need (Bain, 2010). Some universal design principles can be followed when designing transit maps, but research to date mainly focuses on its cartographic function. According to Mollerup (2015), a well-designed transit map needs to display the rail lines, stations, stops, and other relevant travel information clearly and do so systematically. A transit map focuses on the relationship between stations and the connections between them, which allows the users to have a better understanding of the transportation system rather than the geographic information (Mollerup, 2015). Several design factors may affect the legibility of a transit map, such as color coding, the visual layout of the lines and stations, as well as interchange signs. Additional information includes important zones, attractions, and other transportation choices that could affect the user's decisions (Wagner, Wolff, Kapoor, & Strijk, 2001). Table 1 summarizes the general guidance for transit map design.

Information Structure	Establishing clear direction instructions that can help users locate where they are and how to find a transit stop to begin their journey, decide how to travel on the network, and then navigate from the transit system stop or station to their final destination (Bain, 2010). Making sure that all information displayed is related to traveling because irrelevant information may cause information overload (Chang et al., 2018). Other relevant information such as airports and landmarks can be added, which is helpful for tourists to have an overall understanding of the area and other main transportation choices (Bain, 2010).
Map Legend	Information in a map legend should be grouped into themes, and relevant information should be displayed close to each other. Key information needs to be more legible for users to find it easily (Buard and Ruas, 2009).
Transit Lines	Color difference between transit lines should be as clear as possible to enhance the visual conspicuity of the target information in a route searching (Ou, 2019). A reasonable geographic distortion to adjust positions of stations is acceptable to help to simplify the routes and span the breadth of each network (Bertin and Berg, 2011). Transit lines can be simplified with diagonals constrained at 45° to the verticals and horizontals Holscher et al. (2011).
Stations	Stations need to be represented with simple dots or evenly spaced tick marks, and station names should be marked uncrowned with legible typefaces (Bain, 2010). Instead of names, some stations may be represented with numbers, or stations can be paired with letters for short lines. Station stops and interchanges can be denoted by a range of graphic indicators (Sort, 2006).
Interchanges	Transfers or interchanges between lines need to be demonstrated with visual signs, such as a "white-line connector" designed by Beck in 1933. This helps users to understand the complicated transit system and avoid possible mistakes (Garland, 2008).

Table 1. Principles of transit map design

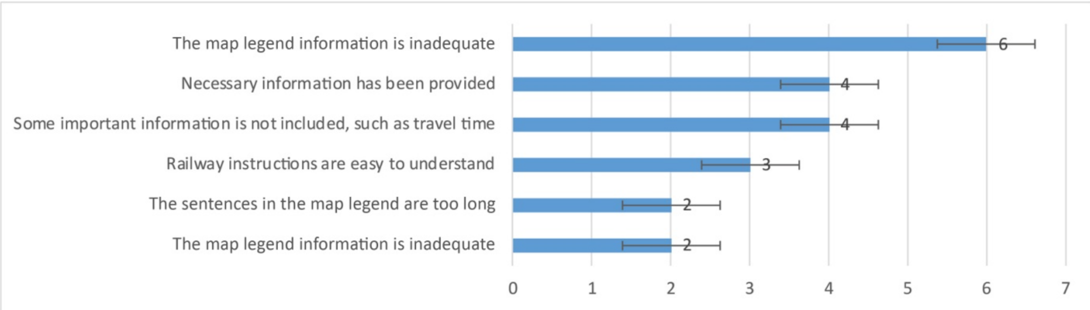
2. Eye-tracking user performance test

2.1. Scoping

In order to inform the design of the eye-tracking user performance test, an interview was conducted to identify potential design problems with the map. It involved ten international tourists who were identified by the Chinese Transportation Department as frequent users of the high-speed railway, i.e. who had experience using the railway map (average age of 32.2). The questions and results of the interview are shown in Figure 2.

	N	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Q1: Is the map easy to use?	10	20%	30%	20%	20%	10%
		50%			30%	
Q2: Is the information structure clear?	10	0%	20%	20%	30%	30%
		20%			60%	
Q3: Is the visual design good?	10	10%	30%	40%	20%	0%
		40%			20%	

Q4: Why do you think the information structure is clear/unclear?



Q5: Why do you think the visual design is good/not good?

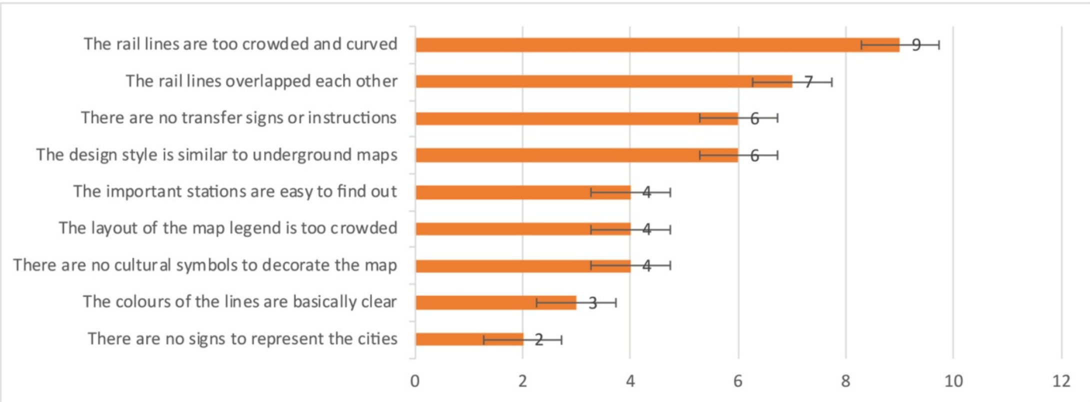


Figure 2. Results of the interview – scoping stage

A scoping user performance test was also conducted to identify further potential design problems. All 20 participants were university students with an average age of 20.6 years old. None of them had experience using China's high-speed railway map. Participants had to complete four information searching tasks with different levels of difficulty. The tasks had no time limit and focused on participants' user experience. The results of the user test are summarized in Table 2, which shows whether each design element present in the current transit map meets the design principles of map design.

Design principles		Potential limitations raised by users
Clear and understandable instructions	√	-
Relevant information are linked to each other	√	-
Sufficient information displayed on the map	×	Additional information, such as travel time may be needed
Other necessary travel information also included	×	No other important transportation information is included, such as airports
A functional and understandable map legend	×	Long sentences used, which are difficult to read
Simplified shapes of transit lines	×	Too many unnecessary curves
Uncrowded layout of transit lines	×	Transit lines are crowded and even overlapped
Distinct colors between lines	?	The design needs further research to test user performance
Transit lines grouped with numbers	√	-
Stations should be represented with visual marks	√	-
Stations evenly spaced	√	-
Uncrowded layout of station names presented in legible typefaces	×	Station names are crowded and difficult to read because of the messy layout
Interchange points should be displayed with visual signs	×	No clear interchange signs displayed

Table 2. Potential design problems of the map identified in the user test–scoping stage

The results of the interview and scoping user test show clearly that the map has legibility problems and that many design elements do not follow transit map design principles. Most design problems mentioned by the participants related to the layout of transit lines and train stations, as well as the transfer system. These findings helped to focus on specific design problems that were then tested in more depth through an eye-tracking user performance test investigating the legibility of the current high-speed railway map (Figure 1).

2.2. Participants

In order to test the user behavior of international travelers objectively, all participants chosen were non-Chinese citizens. Participants had to answer three screening questions before the test in order to confirm that no one had knowledge of Chinese geography. A total of 30 participants were tested, all of whom were from different countries and had no cultural relationship with China. Although the gender difference factor was not considered in this study, the numbers of male and female participants were kept the same to allow further analysis of findings (15 male participants and 15 female participants).

The average age of the participants was 26.1 years old. This is was to replicate the fact that young people between 20–30 years old make up the largest group (31.4%) visiting China in 2019.

All participants took the Ishihara test for color blindness to make sure that they each had normal color vision.

2.3. Test environment, procedure, and materials

The testing was conducted in a lighting laboratory, using an D65 (average daylight) setting, which is similar to the lightness level applied in a railway station. Participants were asked to complete several tasks on a computer monitor with a screen resolution of 1920 x 1280 pixels. This setting was used to simulate the conditions in the China's railway stations where transit maps are displayed on computer monitors to provide travel information. All participants sat in front of the screen at around 80 cm distance, and their eye-movements were recorded using a Tobii Pro 2C eye tracker. The eye-tracking software Tobii Studio 3.01 was used, which enabled the reading process of each participant to be analyzed via gaze plots and heatmaps. Additional data included the time participants took to complete each task.

To increase the validity and reliability of the findings, other analysis methods such as SPSS calculation were used, as described later in the article. In order to make sure that different groups of data were normally distributed, detailed data distribution analysis, including histograms, was conducted. In order to make clear comparisons between variables a paired sample *t*-test was used.

2.4. Task settings

To evaluate the design quality of China’s high-speed railway map thoroughly, participants had to complete four tasks, pointing out their path to finding the information by using a mouse. Each task corresponds to a specific research target as shown in Table 3.

Tasks in eye-tracking test

· Task 1
Find the train line: Huning PDL.
To test the efficiency of the color coding used for the transit lines, and observe how people locate a specific transit line on the map.

· Task 2
Find out the distance between Jinanxi and Zhangqiu after reading the map legend.
To test the legibility of the map legend, especially the layout structure and visual design of the key information.

· Task 3
Plan a route between two stations: Beijing and Shijiazhuang.
To assess participant reading strategy when planning a simple route, and to investigate the potential layout problems of the cities and/or transit lines.

· Task 4
Plan a route between two stations: Nanjing and Ningbo.
To assess participant reading strategy when planning a complicated route that contained more than three interchange points. This task focused on the transfer system and transfer signs of the map.

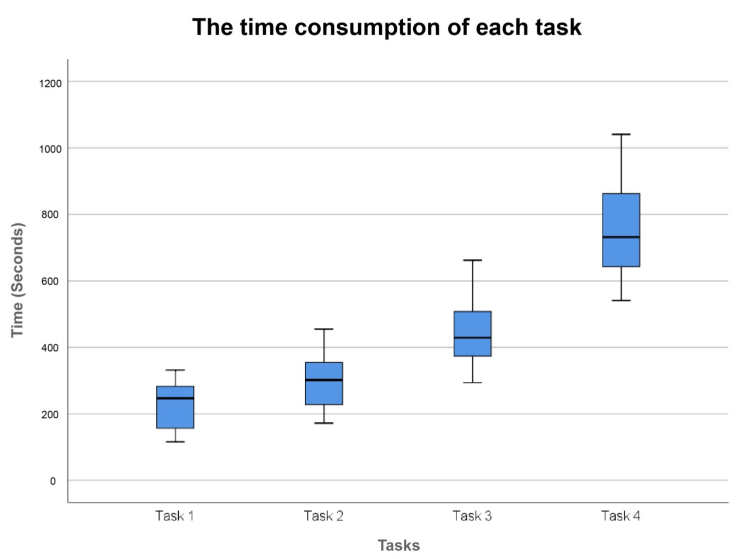
Table 3. Research target of each task

In addition, after each task was completed, participants were asked to answer some questions based on their experience. Questions were relevant to each individual's information searching behavior. For example, when participants spent a long time reading the information in a certain area, they were asked why they had focused on this area for such a long time. This was important in order to increase objectivity when analyzing and understanding the findings.

In an eye-tracking test, the longer the fixation on a particular area of information, the higher the interest in that area (Dwyer, 2018). On the other hand, the longer the gaze duration is the more difficulties the user is having in finding/reading/understanding the information (William, 2018). Therefore, the gaze duration is used in several studies as an indicator and to measure the difficult parts in reading experiments. Additional data regarding participant/user experience was collected through interviews after each task.

3. Results and analysis

The average time to complete the test was 1715s with the fastest participant taking 1367s and the slowest taking 2126s. According to Figure 3, as the complexity of the tasks increased, the time participants spent on the task increased as well.



Tasks	N	Mean	Std. Deviation	Std. Error
Task 1	30	228.47	70.434	12.859
Task 2	30	293.50	78.917	14.408
Task 3	30	437.43	95.384	17.415
Task 4	30	755.20	141.565	25.846

Figure 3. Total time of each task

Most participants finished Task 1 (*M*=228s) and Task 2 (*M*=294s) quickly and spent most of their time on Task 3 (*M*= 437s) and Task 4 (*M*=755s).

3.1. Result of Task 1 (Find the train line: Huning PDL)

The accuracy of this task was 57% (approx.). A total of 17 participants found the target railway line correctly (approx. 57%), while ten participants (approx. 33%) made the same mistake, i.e., they assumed another line (Jingguang HSR) was the target line. According to the interviews following this task, more than half of the participants (16, approx. 53%) mentioned that the colorof the main misleading line (Jingguang HSR) looked similar to the target line (Huning PDL). This suggests that the color-coding of the railway lines caused confusion and leads to reading mistakes.

Figure 4 shows the most recurrent types of information searching process used by 14 (approx. 47%) participants through gaze plots. The gaze plot of a representative participant (i.e., representing those 14 participants who used a similar information searching process) is used as an example

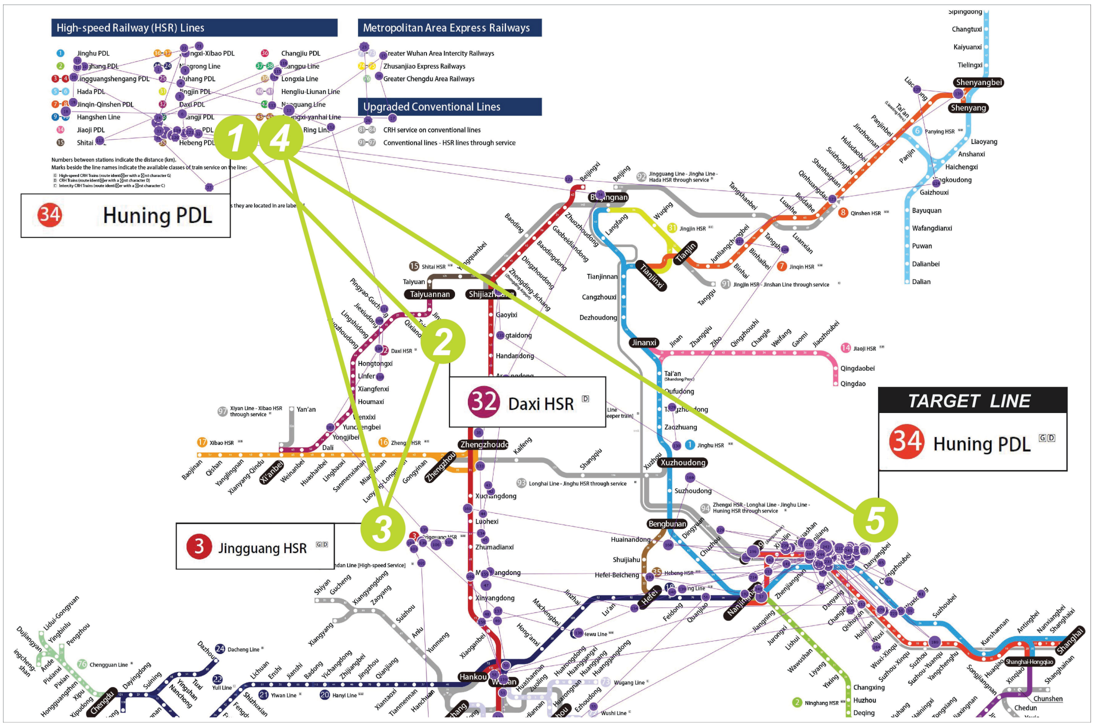


Figure 4. The gaze plot of representative participants in Task 1

As seen in Figure 3, participants started from the map legend (Step 1). The gaze points show that they read all the railway line instructions and found the target line instruction. Then they looked into the railway system (Steps 2 and 3) to check the lines with similar colors (Daxi HSR and Jingguang HSR). The gaze points show that they navigated along the misleading line (Jingguang HSR) and tried to find useful information but failed. Participants looked back to the map legend (Step 4) and double-checked the number (No. 34) of the line, which then helped them find the target line (Huning PDL) on the transit map (Step 5). This reading process shows that the design inefficiency of the color system causes frequently re-reading, wastes user time, and allows users to make errors. The frequent re-reading behavior in the eye-tracking test indicates that participants were not confident they were looking at the right information and could also indicate that the information was wrong or difficult to understand (De Smet, Leijten, & Van Waes, 2018).

Another interesting finding in this task is the relationship between colors and sentences. In order to show what elements participants reviewed the most, a sample comparison was made between the color of the target line and the name of the target line. Eight variables were compared based on fixation counts and fixation duration (see Tables 4 and 5):

Fixation Count Paired Samples Test

		Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pair 1	V1 - V3	-2.167	2.890	.528	-4.107	29	.000
Pair 2	V2 - V4	-1.933	4.675	.854	-2.265	29	.031

Fixation counts (n):

V1 - Fixation counts of the target line's ***name*** on map legend.

V2 - Fixation counts of the target line's ***name*** on map.

V3 - Fixation counts of the target line's ***color*** on map legend.

V4 - Fixation counts of the target line's ***color*** on map.

Table 4. Fixation count paired samples of a target line's color and its name

According to Table 4, there was a significant difference between V1 and V3 ($p < .01$), which shows that fixations on color (*M*=3.27, *SD*=2.53) of the target line in the map legend were higher than fixations on its name (*M*=1.1, *SD*=1.24). Similarly, when comparing V2 and V4, the fixations on color (*M*=3.4, *SD*=4.18) of the target line in the transit system were significantly higher ($p < .05$) than fixations on its name (*M*=3.4, *SD*=4.18). This indicates that participants looked at the color of the target line more often than the name in both the map legend and the transit system.

Fixation Duration Paired Samples Test

		Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pair 1	V5 - V7	-.77345	1.10397	.20500	-3.773	28	.001
Pair 2	V6 - V8	-.59133	1.53587	.28041	-2.109	29	.044

Fixation duration (Sec.):

V5 - Fixation duration of the target line's ***name*** on map legend.

V6 - Fixation duration of the target line's ***name*** on map.

V7 - Fixation duration of the target line's ***color*** on map legend.

V8 - Fixation duration of the target line's ***color*** on map.

Table 5. Fixation duration paired samples of a target line's color and its name

The fixation duration calculated through the paired samples *t*-test (Table 5) shows that the fixation duration on color (V7, M=1.15, SD=1.02) of the target line in the map legend was significantly longer ($p < .01$) than on the name (V5, M=0.38, SD=0.56). In addition, there was a significant difference between V6 and V8 ($p < .05$), which shows that fixation duration on color (V8, M=1.15, SD=1.51) of the target line in the map legend was significantly longer ($p < .05$) than on the name (V6, M=0.42, SD=0.27). This indicates that participants spent more time looking at the color of the target line than at the name, not only in the map legend, but also in the whole transit system.

3.2. Result of Task 2 (What is the distance between Jinanxi and Zhangqiu?)

The accuracy of this task is approximately 37%, i.e., only 11 of the 30 participants answered the question correctly. There are two stops between the stations, and therefore the correct answer should be the sum of the two distances: 20 km + 49 km = 69 km. The most common incorrect answer was 49 km, with 16 (approx. 53%) participants making the same mistake. In addition, three participants failed to find this information.

The reasons for this result are various. A representative participant gaze plot (Figure 5) is used again as an example, which is representative of 20 participants. In the map legend, participants had to find key information (Figure 5, Part A). First participants had to find numbers between stations that indicate the distance (km). Then they investigated the target stations: Jinanxi and Zhangqiu on the Jiaoji HSR line (Figure 4, Part B). At this point, participants felt uncertain about the answer and went back to read the sentences. A large number of refixation paths can be seen between Part A and Part B in Figure 5, showing that participants double-checked the key information several times and still gave the incorrect answer (49 km). Participants mentioned that they had not realized that the distances between stations needed to be added together because no example is given in the map legend to show this.

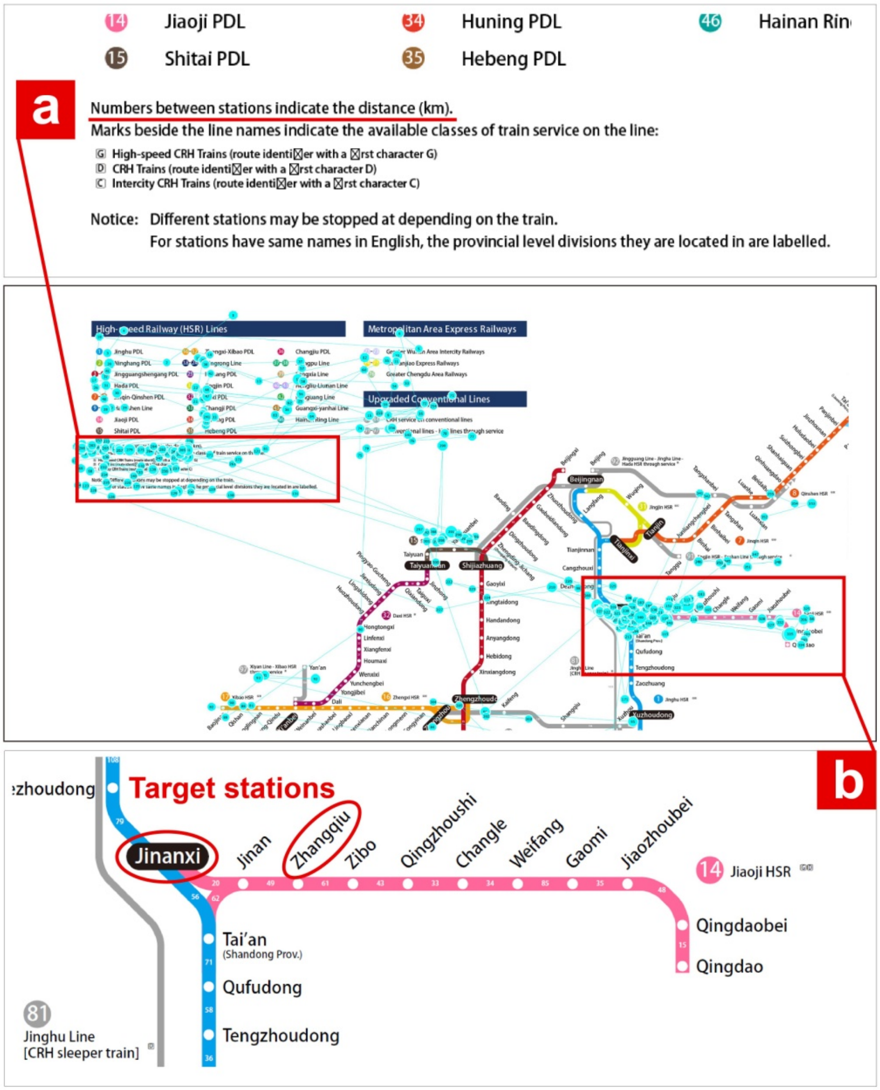


Figure 5. The gaze plot of a representative participant in Task 2

From the searching behavior of the 20 participants, design limitations of the map legend can be identified as follows:

1. The information structure of the map legend is not clear enough to display the different categories of information together. This leads to most participants finding the key information after searching the whole map legend, which wastes travelers' time in a time-critical situation of catching a train.
2. The key information for this task is not easy to understand. More than half of the participants answered wrongly due to misunderstanding the information. Compared with the current text-only sentences, other design solutions such as adding iconic examples to display the key information could be considered.

3.3. Result of Task 3 (Plan a route between Beijing & Shijiazhuang)

At the beginning of this task, an interesting instance of reading behavior occurred whereby half of the participants spent a long time reading the map legend trying to find relevant useful information. However, detailed information on specific stations is not included in the map legend. Participants mentioned in the interview that they assumed the stations and lines could be found directly in the map legend. This indicates that after two tasks, some participants were still unfamiliar with the map legend. This supports the previous finding in Task 2 that the information structure of the map legend is not clear enough.

By analyzing the overall gaze plots of this task and the interview records, most participants started searching for Beijing Station at the beginning as they assumed that they could find Beijing easily, thinking that the station of the capital city would stand out. However, Beijing Station is not represented in a different way to normal stations on the current map. According to Harris (2007), if a design is not as users assume or expect it to be, it will cause confusion. This task, therefore, supports Harris's opinion that participants wasted time in this stage.

Figure 6 provides details of where Beijing Station is located. Beijing Station is one of the three stations in Beijing City, but another station—Beijingnan—became the most misleading station in this task. Participants noticed Beijingnan Station quickly because this important station is emphasized with a black block with white text. Participants mentioned in the interview that Beijingnan Station looked as if it could have relevance to Beijing Station, but it was not the target station they were looking for. They then wondered whether looking at Beijingnan station was misleading and some participants left the area and started searching for Beijing in other parts of the map.

Figure 6. Stations in Beijing city: Beijing, Beijingnan, and Beijingxi

Table 6 shows more details of the fixation numbers and fixation duration on the two target stations (Shijiazhuang, Beijing) and the misleading station (Beijingnan).

The fixation number on Beijingnan (M=35, SD=9.04) is significantly higher than ($p < .01$) the target stations Beijing (M=8, SD=3.7) and Shijiazhuang (M=8.37, SD=3.42). The fixation duration on Beijingnan (*M*=14.37, *SD*=4.4) is also significantly higher than ($p < .01$) Beijing (*M*=2.8, *SD*=1.27) and Shijiazhuang (*M*=2.95, *SD*=1.29). This shows that participants spent most of their time in this task hesitating about an unwanted target.

Paired Samples Test (Task 3)

		Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Pair 1	V1 - V3	-27.433	9.339	1.705	-16.089	29	.000
Pair 2	V2 - V3	-27.067	10.329	1.886	-14.353	29	.000
Pair 3	V4 - V6	-11.57333	4.25593	.77702	-14.894	29	.000
Pair 4	V5 - V6	-11.42133	4.56201	.83291	-13.713	29	.000

Fixation counts (n):
V1 - Fixation Count (Beijing).
V2 - Fixation Count (Shijiazhuang).
V3 - Fixation Count (Beijingnan).

Fixation duration (Sec.):
V4 - Fixation Duration (Beijing).
V5 - Fixation Duration (Shijiazhuang).
V6 - Fixation Duration (Beijingnan).

Table 6. Fixation count and duration between Beijing, Shijiazhuang, and Beijingnan

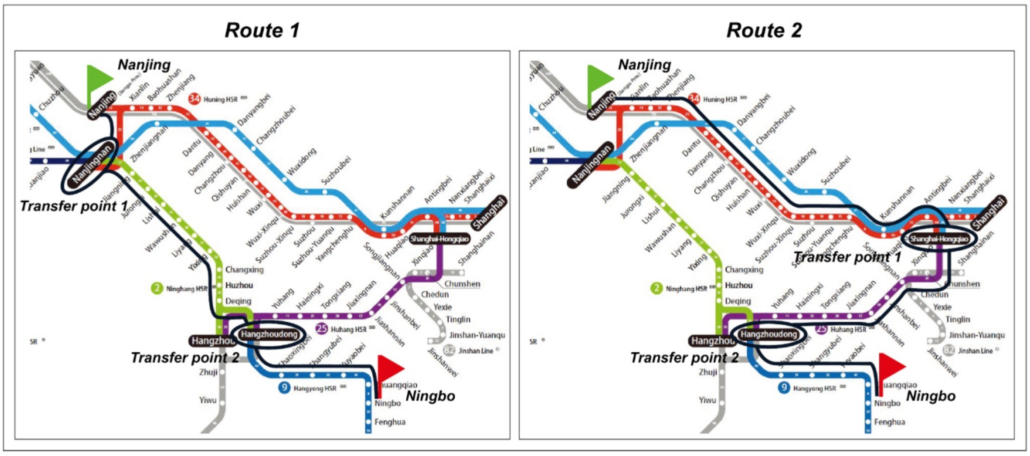
A question that arose in this task is the relationship between cities and stations, i.e., is it necessary to display the spatial relationship between cities and stations? In other words, will linking all three stations in Beijing help users to distinguish the differences more easily? This question is further discussed in the next section.

3.4. Result of Task 4 (Plan a route between Nanjing & Ningbo)

This task was the most complicated of all the four tasks, although the mission of the task is almost the same as Task 3: plan a route between two stations, but with different complexities. In Task 4, there were at least two interchange points between the target stations, and the main research target focused on the design quality of the transfer system.

There were 14 participants (approx. 47%) who made mistakes when planning the routes, and two participants (approx. 7%) who failed to find the cities at all. This shows low information searching accuracy in this task. Only 14 participants (approx. 46.67%) finished the task with no mistakes. The main factor that caused the information searching errors in this task may have been the poor design of the transfer system. Many participants transferred to other lines at points where a transfer cannot happen. The following instruction shows more details of the routes in this task (Figure 7):

Two main routes between the target stations: Nanjing & Ningbo



Route 1 is to take Huning HSR from Nanjing (green flag) to Nanjingnan, **transfer** Ninghang HSR to Hangzhoudong, and then **transfer** Hangyong HSR to the final destination – Ningbo (red flag).

Route 2 is to take Huning HSR from Nanjing (green flag) to Shanghai-Hongqiao, **transfer** Huhang HSR to Hangzhoudong, and then **transfer** Hangyong HSR to the final destination – Ningbo (red flag).

Both main routes contained 2 transfer points.

Figure 7. Two main routes between Nanjing and Ningbo

Figure 8 (Part A) shows the most confusing area in this task that caused ten participants (approx. 33%) to make the same mistake. The Ninghang HSR (green line) and the Huhang HSR (purple line) overlapped and their intersected point led users to assume that it was a transfer point (red circle), which lead them to choose to transfer at that point. However, in reality, passengers cannot transfer between the two lines at the overlapped area. They can only transfer at Hangzhoudong Train Station further down this area.

The misleading position in Task 4 where transit lines are overlapped

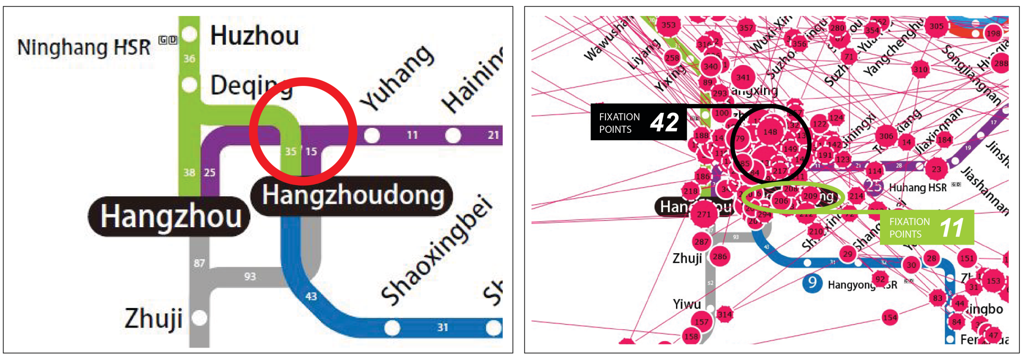


Figure 8. Interchange points where transit lines are overlapped

The gaze plot of a participant's eye-movement in this task (Figure 8) provides more details of user reading behavior. There are 42 fixation points on the overlapped position, much more than on the actual transfer point at Hangzhoudong Railway Station (11 fixation points). A large number of re-fixation paths show frequent re-reading behavior by participants, which means that participants encountered problems in this area and spent a longer time checking the information (Mannion & Leader, 2014). Thus, the lack of transfer signs and the overlapped lines wasted user time and caused various mistakes in this task.

Another interesting example of user behavior found in this task was that participants paid more attention to the destination (Ningbo), rather than the origin station (Nanjing), which can be observed from the heat map in Figure 9. Participants preferred to plan a route that looked closer to the destination. This user behavior can be explained by Holscher et al. (2011) who claimed that people tend to choose the direction closest to the destination as far as possible, even if the route is longer or more time-consuming. This information searching behavior is known as a direction-based strategy. This can also explain why most participants (21, 70%) chose the route downwards from Nanjing (origin station) rather than another route as this route looks closer. Although this does not provide evidence that such a habit may lead to mistakes in route planning, it may limit the user's route choices.

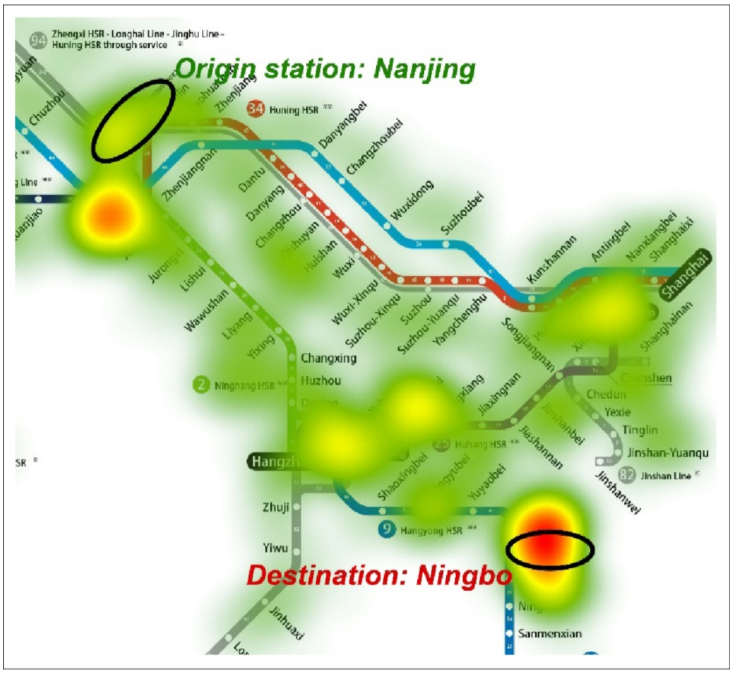


Figure 9. Heat map of Task 4

4. Discussion

Based on the results of the tasks and the information searching behavior of participants, many useful findings about the legibility of China's high-speed railway map need to be further explored. Some limitations in terms of the information design and visual elements have been found in this study. Participants also provided some suggestions based on their user experiences, which may help to establish an initial plan for future improvement of the map design.

4.1. Information design

Dziekan (2008) mentioned that the railway transit line system is the primary source of information for passengers traveling to unfamiliar places. Clear and systematic design of the lines, stops, transfer points, and relevant instructions are the key factors in a transportation map. In Task 3, a large proportion of participants had difficulties in locating Beijing, a station which they assumed would stand out in the map (stations nearby, such as Beijingnan, caused confusion and participants were not sure of the spatial relationship between them). The lack of connections between stations is also a common problem in network transportation map design, such as two stations that are very close to each other in a zone but are displayed separately and far away on a map, which misleads and limits user route choices (Grison et al., 2016). This was the case of the three stations in Beijing City, which are displayed separately with no connections. Participants suggested linking the stations in the same city by using symbols or establishing a zone for the stations in the same city. This would allow grouping the stations in certain areas and would reduce reading and searching confusion. The British national railway map provides a good design solution showing the spatial relationship between cities and stations, all 12 stations in London have been grouped in a gray zone, which makes information clearer.

On the transit lines, numbers are used to represent the travel distance between stations, but because distances between stations are different, the numbers are disordered. In Task 2, participants did not understand these numbers due to poor explanation of the travel distance information. Many participants also indicated that they were more interested in traveling time than traveling distance, which is more helpful when planning their traveling journey. According to Guo (2011), a transit map is a schematic diagram that focuses on locations, directions, and connections of stations and lines in a public transit system. It normally does not include services, such as traveling time or crowding. But, Avineri and Prashker (2006) argue that traveling time information can help travelers to make better route choice decisions. Hadlaw (2003) further explains that a map allows travelers to imagine the route geographically, but more importantly, users wish to know the traveling time to each place on the map as this provides them with more impetus to explore the area. In our study, a few participants suggested adjusting the length between stations to represent travel time instead of distance. According to Harvey (1998), the distance between stations is normally arranged as a certain standard distance, and the author agrees that this action typically represents traveling time rather than traveling distance.

The map legend is essential to the effectiveness of a map, as it contains most of the necessary explanations and examples to show the functions of a map (Meirelles, 2013). Some design problems regarding the design of the information can be identified from the results. In Task 2, the explanation of the travel distance was not accurate. Only 37% of participants understood the meaning correctly, i.e., that the travel distance of two lines should be the summation of the two travel distance numbers in between. However, this complicated explanation could be demonstrated in another way, such as a visual example, as most participants suggested. Visual encoding is a process to provide a suitable type of representation by using graphical elements. For example, MacEachren (2008) supports the idea that good design optimizes the thinking process. Therefore, visualization can help process the map information more efficiently, in contrast to a long descriptive explanation. In Task 3 and Task 4, a lack of necessary instructions caused difficulties where there was no relevant information in the map legend for both tasks, but most participants still spent time looking at the legend as they expected they could find useful information there. This can be called, as defined by Nielsen and Pernice (2010), a desired exploration in an eye-tracking test. Participants' reading process showed that the information structure of the map legend is not clear enough. Participants did not have a clear understanding of what kind of information was included in the map legend, which led to frequent re-reading resulting in low reading efficiency and time wasting. Due to the unsatisfactory reading experience, participants suggested regrouping the map legend information. This suggestion is in line with Meirelles' (2013) claim that the map legend should follow the grouping principle to avoid forcing users to search for information through exhaustive re-reading. Visualized instructions or symbols were recommended by participants to be included in the map legend as they have become increasingly prominent tools for making information "meaningful" and for discovering important relationships within data sets (Salvo, 2012).

4.2. Visual design

According to Garland et al. (1979), the more detailed the map is, the more important the color-coding of transit lines will become for route planning accuracy. A well-designed color system on a map can provide users with more confidence, less confusion, and frustration. Netzel et al.'s (2016) research showed that the difference in response time and correctness between colored and gray-scale maps drops as complexity increases. That is, the use of color could, at a certain complexity level, no longer be useful. In Task 1, the use of a similar color of a misleading railway line led many participants to make mistakes. This problem may be because the contrast between colors is not strong enough in hues and values, which is a very common problem in map design (Buard & Ruas, 2012). A detailed evaluation of the color system needs to be further explored. In addition, the reading behavior of participants showed that they preferred to locate the target line according to its color in the first place, then the number, and then the name. This trend can be supported by the results of Task 1 where the color of the target line got more fixation points and a longer fixation duration than words. However, although color is important in the first instance, i.e. to grab attention and direct the user, when analyzing the reading process it can be seen that most participants still double-checked the target line by reading its number and name. Participants explained that reading certain numbers and sentences made them feel more confident that they had found the correct information. This indicates that the color system is not enough to represent the lines. Key information on a map should always be supported by numbers or sentences.

In Task 4, the layout of transit lines caused some reading difficulties in the route plan. The main problem is that a large number of curved lines overlapped each other. Overlapping is a normal problem in transit system design and refers to the crossing of routes running on different places that do not actually overlap (Sadahiro et al., 2015). Another common problem is acute bend, i.e. the bend of routes at an acute angle. This made participants assume that there was a stop or a transfer station, which happened many times in Task 4. Acute bends clutter route maps, especially when many routes are running in parallel (Agrawala & Stolte, 2001). Another problem is that some railway lines are irregularly curved and look disordered. In order to find out the reason for the use of irregularly curved lines, an interview was conducted with an officer who is currently working in the Chinese Railway Department. According to the officer, the main reason for the disordered layout is that the current design is largely based on real geographic information. That is, all the cities, stations, and lines on the map are as geographically accurate as possible. Is it necessary to consider geographic accuracy in transit map design? According to Marcus (2016), transit lines can be seen as an independent system apart from the physical geographic environment, and geographic accuracy is normally not a necessary requirement in a transit map. Therefore, some distortion and simplification of the direction and distance can be used to improve the information. It can also help to simplify the user's information analysis process and enhance route decision-making efficiency (Hochmair, 2009). Mollerup (2015) suggests only using horizontal and vertical lines as well as diagonal lines of 45 degrees. This indicates that it is not necessary to make the railway lines geographically correct on a transit map. The need to follow Mollerup's suggestion can be supported by the gaze duration in the test, where there are more fixation points and longer fixation duration on the over-curved lines than the straight lines. This means that these areas wasted user time when route planning.

In both Task 3 and Task 4, participants felt confused when route planning because the transfer points were not shown clearly. This was the case in Task 4 where participants needed to plan routes that included at least two transfer points, and many participants transferred in the wrong place. A well-organized transfer system can display the connections between lines or transport modes systematically. However, this is often poorly represented and can be interpreted incorrectly in transit map design (Grison et al., 2016). In the current map, there are no transfer signs nor any symbols with which to visualize the transfer points. This led to participants feeling unconfident about the routes they planned because they were not sure that they could transfer to the correct train at confusing transfer points. Based on this problem, participants suggested establishing a visual system for transfer points and demonstrating it clearly on the map legend. According to Guo (2011), designers need to understand different ways to represent transfers and visualize them with different symbolized elements. Most of these cases are overlapped stations, although there are also semi-overlapped stations and separate stations connected by a link. It is necessary to clearly separate lines or stations with no connections because a confusing layout can easily lead users into going the wrong way.

Moreover, around half the participants suggested adding cultural symbols for popular tourist cities. They believed that these improvements could make it easier to find important destinations, as well as more travel information compared with the current design. However, Avineri and Prashker (2006) argue that spatial distinctions should be ignored on a transit map. For example, in the London Underground map designed by Beck, localities were essentially "standardized" for reasons of visual clarity and balance. There is no other unnecessary information or decoration, because these may be redundant when reading.

However, because a railway map contains more information, the spatial relationship between cities and stations is more complicated than an underground system. According to Chang et al. (2018), travelers need to know some basic spatial and travel service information. As participants suggested previously, adding other transportation information, such as airports and undergrounds, can provide travelers with more transportation choices. This should, therefore, be considered in future research.

5. Conclusion

This study evaluated the design quality of China's high-speed railway map based on user behavior by collecting quantitative data using eye tracking followed by interviews to collect qualitative data. The different categories of information and visual elements of the map were tested in a total of four tasks that were set to answer the research questions defined for this study. The results of the test showed some design limitations in terms of information structure and visual elements, such as confusion with colors, the overlap of railway lines, the lack of transfer signs, etc. These design limitations affected the legibility of the map and caused information searching errors. With the help of the eye tracker, it was possible to observe the information searching process of the users, to see which areas of the map or what visual elements made users feel interested or confused, and how design limitations affected user experience. The results of the tests also helped to understand user habits and potential needs.

This study shows that the legibility of China's high-speed railway map is limited and that many improvements are needed. The main problem of the map is that the poor information structure weakened the relevance between the visual elements, such as railway lines, train stations, and number systems, etc. Some other visual design problems like the color system and transfer signs also need to be further explored and improved. Furthermore, the eye-movement data and design suggestions from participants gave rich details about the user reading experience. Improving the map design by taking into account user needs will reduce error and improve user acceptance and satisfaction with the design solution (Endsley & Jones, 2012).

In eye-movement data analysis, the heat maps and gaze plots visually explain how design limitations affect user behavior. The present study shows that eye-tracking is an effective design evaluation method of transit map design. Our study also shows that, at least for design-related research, eye-tracking should be complemented by qualitative research methods to help explore user behavior further, which cannot be observed from eye-tracking data directly. This is helpful to reduce ambiguity in the conclusions made from the data analysis. For example, the long gaze duration could reflect users' interest areas, but it could also indicate that users feel confused.

A potential limitation with this research is that, although participants had no user experience of China's railway map or in-depth knowledge of Chinese geography, different experiences of using transit maps could affect the accuracy and objectivity of the results. For example, a participant who frequently takes public transportation has more experience in reading transit maps. According to Beck et al. (2012), a user's information searching efficiency using a transit map can be affected by their experience, knowledge of the targets, or familiarity of the map system. In addition, similar background (such as educational level and familiarity with digital technology) among participants may limit the generalization of the test results. Therefore, different user groups of a more varied background should be included in future research. Similarly, this study only used one type (and size) of device in the tests. But, there are opportunities in the future to investigate impact when using different type or size of devices, and different media such as printed copy.

Overall, this study is an important contribution to knowledge in the field of information design, more specifically in the evaluation of transit maps from different angles, i.e. not only in terms of information structure and legibility but also in terms of information visualization and the application of visual elements.

References

Agrawala, M. & Stolte, C. (2001). Rendering effective route maps: Improving usability through generalization. *Computer Graphics and Interactive Techniques*, *21*(3), 12–17.

Avineri, E. & Prashker, J. (2006). The impact of travel time information on travelers' learning under uncertainty. *Transportation*, *33*(4), 393–408.

Bain, P. (2010). Aspects of transit map design. *Parsons Journal for Information Mapping*, *2*(3), 6–11.

Beck, M. R., Trenchard, M., van Lamsweerde, A., Goldstein, R. R., & Lohrenz, M. (2012). Searching in clutter: Visual attention strategies of expert pilots. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 56, 1411–1415.

Bertin, J. & Berg, W. (2011). *Semiology of graphics*. Redlands, ESRI Press.

Brychtova, A., & Coltekin, A. (2016). An empirical user study for measuring the influence of colour distance and dont size in map reading using eye tracking. *The Cartographic Journal*, *53*(3), 202-212.

Buard, E. & Ruas, A. (2009). *Processes for improving the colours of topographic maps in the context of Map-on-Demand*. https://www.researchgate.net/publication/228488125_Processes_for_improving_the_colours_of_topographic_maps_in_the_context_of_Map-on-Demand

Chang, T., Yang, D., Yang, Y., Huo, J., Wang, G. & Xiong, C. (2018). Impact of urban rail transit on business districts based on time distance: Urumqi Light Rail. *Journal of Urban Planning and Development*, *144*(3), 18–24.

De Smet, M., Leijten, M. & Van Waes, L. (2018). Exploring the process of reading during writing using eye tracking and keystroke logging. *Written Communication*, *35*(4), 411–447.

De Smet, B., Lempereur, L., Sharafi, Z., Guéhéneuc, Y., Antoniol, G. & Habra, N. (2014). Taupe: Visualizing and analyzing eye-tracking data. *Science of Computer Programming*, *1*(79), 260–278.

Dwyer, T. (2018). *Seeing into screens: eye tracking and the moving image*. Bloomsbury Academic.

Dziekan, L. (2008). The transit experience of newcomers to a city – learning phases, system difficulties, and information search strategies. *The 87th Meeting of the Transportation Research Board*, Washington DC.

Endsley, M. & Jones, D. (2012). *Designing for situation awareness*. CRC Press.

Garland, K. (2008). *Mr Beck's underground map*. Capital Transport.

Grison, E., Gyselinck, V., Burkhardt, J., & Wiener, J. (2016). Route planning with transportation network maps: an eye-tracking study. *Psychological Research*, *81*(5), 1020-1034.

Guo, Z. (2011). Mind the map! The impact of transit maps on path choice in public transit. *Transportation Research Part A: Policy and Practice*, *45*(7), 625–639.

Hadlaw, J. (2003). The London underground map: Imagining modern time and space. *Design Issues*, *19*(1), 25–35.

Harris R. (2007). *The elements of visual style*. Houghton Mifflin.

Harvey, D. (1996). *Justice, nature and the geography of difference*. Blackwell.

Hochmair, H. (2009). The Influence of Map Design on Route Choice from Public Transportation Maps in Urban Areas. *The Cartographic Journal*, *46*(3), 242-256.

Holmqvist, K., Nyström, N., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6, 65–70.

Holscher, C., Tenbrink, T., & Wiener, J. M. (2011). Would you follow your own route description? Cognitive strategies in urban route planning. *Cognition*, *121*(2), 228–247.

Kang, Z., Mandal, S., Crutchfield, J., Millan, A. & McClung, S. (2016). Designs and algorithms to map eye tracking data with dynamic multielement moving objects. *Computational Intelligence and Neuroscience*, *16*(1), 1–18.

MacEachren, A. (2008). *How maps work: Representation, visualization & design*. Guilford Press.

Mannion, A. & Leader, G. (2014). Attention-deficit/hyperactivity disorder (AD/HD) in autism spectrum disorder. *Research in Autism Spectrum Disorders*, *8*(4), 432–439.

Marcus, A. (2016). Underground maps unravelled: Explorations in information design. *Information Design Journal*, *21*(1), 64–66.

Meirelles, I. (2013). *Design for information*. Rockport Publishers.

Mollerup, P. (2015). *Data design*. Bloomsbury Academic.

Mot.gov.cn. 2020. *Ministry of Transport of the People's Republic of China*. http://www.mot.gov.cn/ [Accessed 19 May 2020]

Netzel, R., Ohlhausen, B., Kurzhals, K., Woods, R., Burch, M. & Weiskopf, D. (2016). User performance and reading strategies for metro maps: An eye tracking study. *Spatial Cognition & Computation*, *17*(1–2), 39–64.

Nielsen, J. & Pernice, K. (2010). *Eyetracking web usability*. New Riders.

Ou, L. (2019). Influence *of colour on visual conspicuity*: Taking *subway route map as an example*. International Commission on Illumination. https://www.researchgate.net/profile/Li_Chen_Ou [Accessed 13 May 2020]

Sadahiro, Y., Tanabe, T., Pierre, M. & Fujii, K. (2015). Computer-aided design of bus route maps. *Cartography and Geographic Information Science*, *43*(4), 361–376.

Salvo, M. J. (2012). Visual rhetoric and big data: Design of future communication. *Communication Design Quarterly*, *1*(1), 37–40.

Sort, J. (2006). *Metropolitan Networks*. NAI Boekverkopers.

Wagner, F., Wolff, A., Kapoor, V. & Strijk, T. (2001). Three rules suffice for good label placement. *Algorithmica*, *30*(2), 334–349.

William, C. (2018). Typographic emphasis and contrastive focus: an eye tracking study. [PhD thesis] University of Leeds.