Article

Assessing Map-Reading Skills Using Eye Tracking and Bayesian Structural Equation Modelling

Weihua Dong [1],*, Yuhao Jiang [1], Liangyu Zheng [1] , Bing Liu [1] and Liqiu Meng [2]

[1] Beijing Key Laboratory for Remote Sensing of Environment and Digital Cities, Research Center of Geospatial Cognition and Visual Analytics and Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China; mjiang@mail.bnu.edu.cn (Y.J.); zhengly@mail.bnu.edu.cn (L.Z.); liubing_geo@mail.bnu.edu.cn (B.L.)

[2] Chair of Cartography, Technical University of Munich, 80333 Munich, Germany; liqiu.meng@tum.de

* Correspondence: dongweihua@bnu.edu.cn; Tel.: +86-010-5880-9246

check for
updates

Abstract: Map reading is an important skill for acquiring spatial information. Previous studies have mainly used results-based assessments to learn about map-reading skills. However, how to model the relationship between map-reading skills and eye movement metrics is not well documented. In this paper, we propose a novel method to assess map-reading skills using eye movement metrics and Bayesian structural equation modelling. We recruited 258 participants to complete five map-reading tasks, which included map visualization, topology, navigation, and spatial association. The results indicated that map-reading skills could be reflected in three selected eye movement metrics, namely, the measure of first fixation, the measure of processing, and the measure of search. The model fitted well for all five tasks, and the scores generated by the model reflected the accuracy and efficiency of the participants' performance. This study might provide a new approach to facilitate the quantitative assessment of map-reading skills based on eye tracking.

Keywords: map reading; eye tracking; structural equation model; geography education

1. Introduction

Map reading is a fundamental skill required by many daily tasks, such as map-based navigation [1], the acquisition of spatial information from a map [2], and map-based education [3–5]. It is very important to assess map-reading skills, which can contribute to identifying the current status and evaluating the outcomes of geographical education.

There are different components and levels of competence regarding map reading. Board identified navigation, measurement, and visualization as the three main groups of map-reading tasks that should be addressed when assessing map-reading skills [6]. Clarke defined "functional map literacy" as "the ability to understand and use maps in daily life, for work and in the community" and identified levels of competence concerning map literacy [2]. Research has been conducted to measure these skills for purposes such as assisting education and aiding navigation.

Previous studies have developed various standardized tests to assess map-reading skills. Topographic map reading has been a major subject in map-reading tests. Carswell used the devised and validated Test of Topographic Map Skills (TTMS) to investigate the topographic map-reading abilities of children [7]. The TTMS assesses the ability to read symbols, direction, scale, elevation, grid systems, and map-based information interpretation. Gilhooly et al. used a 7-item contour-map-reading test to assess the skill level of test subjects [8]. The questions included spot height, intervisibility, cross-section identification, direction of flow, and distance. Pederson et al. implemented an 8-question test to examine the outcome of topographic map-reading sessions [9]. The test involved knowledge

of topographic map symbols and scale, ability to identify certain landforms from topographic representations, and interpretation of topographic contour lines. Jacovina et al. developed the Topographic Map Assessment (TMA) method to explore an individual's understanding of topographic maps [10]. The TMA consists of 18 questions, including open-ended responses. The assessment was developed to sample the different ways topographic maps are used in field work [11]. Ooms et al. assessed the map-reading skills of secondary school students using 20 questions based on five topological maps [12]. The test included sub-tasks such as visualization, measurement, and navigation. All the tests mentioned above are scored in terms of the number of correctly answered questions.

Apart from topographic maps, navigational map-reading ability is another important issue that has been investigated. Streeter and Vitello conducted a corresponding test in which subjects were asked to choose and trace a driving route from point A to point B on a black-and-white county map [13]. The assessment of performance included the distance of the chosen route, the number of turns, and map-reading errors, such as illegal entries. Kovach et al. asked subjects to follow a street map to a specific destination and used driving time as a measure to assess the navigational map-reading skills of the subjects [14]. Lobben developed the Navigational Map Reading Ability Test (NMRAT) [15], which consists of five sections (map rotation, place recognition, self-location, route memory, and a way-finding exercise), to predict a person's ability to read maps and navigate through an environment. The NMRAT was designed to be a speed test; only the response time was considered in scoring.

Further tests involve several different kinds of map-reading tasks. To investigate gender and cultural differences in map reading, Chang and Antes used a 45-question test related to a reference map, a topographic map, and a street map [16]. Gillmartin and Patton devised a test where they presented students with maps of a country's population distribution, topography, and climate and asked them to answer questions related to the content of the maps [17].

The studies mentioned above identified different components of map-reading skills and measured them as a whole or by focusing on one particular component, such as topology or navigation. In these tests, performance was assessed based on the results, i.e., either the answers or response time. However, map-reading skills have not been assessed from a cognitive perspective.

Eye tracking has provided insight into the human cognitive process [18,19]. An eye tracker samples and records the user's eye movement (gaze) data at a particular rate (e.g., 60 Hz) and enables researchers to examine the user's visual and attention processes [20,21]. Fixations and saccades are the two basic components of eye movement; thus, they are of major concern in eye tracking analysis. With fixations (the position on a screen where the eye pauses for a certain period), visual information is processed, whereas during saccades (rapid movement between fixations), no visual information is processed [21,22]. The perception of the entire scene is completed by alternating between fixations and saccades [23].

Eye tracking has been widely applied to design geo-visualizations suitable for both the user and the context [24]. However, the comprehension of spatial information depends on both the medium and the ability of the individual [25]. Studies of different user groups have reported that people with different skill levels showed different visual attentive behaviours when they were presented with the same map-related task [26–28]. Hermans and Laarni found that when searching for information on screen maps, experienced users had shorter fixation durations and fewer fixations [29]. When studying eye movements of expert and novice map users using a basic-designed map, Ooms et al. reported significant differences between both user groups, and experts were able to interpret the map contents more efficiently than were novice users [26]. Further research conducted with more complex maps [27] confirmed the existence of differences in visual attentive behaviours between people possessing different skill levels. These reported differences were related to fixation duration, dwell time, fixation count, and scan paths. Stofer and Che [28] also reported different approaches between experts and novices who were tasked with viewing scaffolded global ocean data visualizations. In these studies, the expertise of the map user was reflected by eye movement metrics. These studies indicate that eye tracking has potential to assess users' map-reading skills.

In this study, we used eye tracking and structural equation modelling to quantitatively assess map-reading skills. We collected eye movement data from 258 undergraduate students at Beijing Normal University. We used measures related to the first fixation, processing and search to evaluate map-reading tasks that included map visualization, topology, navigation, and spatial association.

2. Methodology

2.1. Participants

We recruited 272 participants (aged 20 ± 2 years) from Beijing Normal University, all of whom were undergraduate students in the Faculty of Geographical Science. Participants with myopia were allowed to participate with glasses. However, 14 individuals were excluded from the experiment due to failures during the calibration phase, while the remaining 258 participants successfully continued with the experiment.

2.2. Apparatus

We used a Tobii 120 eye tracker with a 17-inch monitor and a sample rate of 60 Hz. The eye tracker recorded with an accuracy of $0.5°$ and a spatial resolution of $0.2°$. The monitor had a resolution of 1280×1024 pixels and display colours of 16.7 M (true 8-bit) to display the stimuli. Data were recorded with Tobii Studio. The experiment was conducted in a well-lit room on campus, and no disruptions occurred during the experimental period. Recordings with sample rates lower than 70% were removed before further analysis.

2.3. Materials

Five tasks were presented to the participants in sequence, and all tasks included multiple-choice questions that were based on a map. The descriptions of the tasks and the sample maps can be found in Table 1. All materials were written in Chinese, which was the native language of the participants.

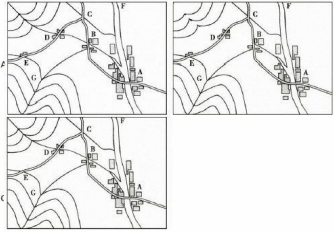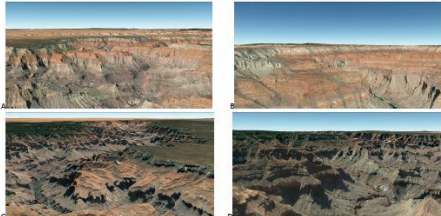**Table 1.** Task description and map sample.
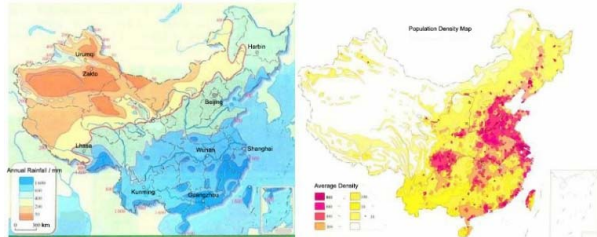
Task	Description	Skills Involved	Map Sample
#1	Choose the map that corresponds to the verbally expressed spatial information	Comprehension of integrated geographic features [30]	
#2	Choose the image that represents the 3D view observed from the given spot and direction on a bird's-eye view image	Topology and transformation from one dimension to another [30]	

Table 1. *Cont.*

Task	Description	Skills Involved	Map Sample
#3	Navigate with a subway map using verbally expressed instructions	Map-based navigation [1]	
#4	Navigate with a campus map using verbally expressed instructions	Map-based navigation [1]	
#5	Identify the spatial correlation between a population map and a precipitation map	Making spatial comparisons and assessing spatial associations [30]	

2.4. Procedure

Participants were guided to complete and sign an entry form with their basic information once they entered the room. Then, they were required to participate in a calibration to guarantee the accuracy of the recordings. They were next instructed to submit their answers and advance to the next task using the computer mouse and keyboard. They were also told that the test had no time limit and that they could not return to the previous page once they had advanced. The participants then proceeded to finish the map-reading tasks.

2.5. Analysis Famework

2.5.1. Eye Movement Metrics

Previous research reported that expert and novice map readers show differences in terms of fixations and scan paths, indicating differences in both the processing and searching of information [26–28]. To assess the performance of the participants, we selected the following metrics (Table 2).

Table 2. Selected eye movement metrics used to assess the performance of participants.

Measure of First Fixation	Measure of Processing
Time to first fixation (AOI [1])	Fixation count percentage
First-fixation duration (AOI)	Fixation duration percentage
Measure of Search	General Performance
Saccade count	Accuracy (whether the answer is correct)
Saccade length	Efficiency (response time)

[1] AOI is used as an abbreviation for "area(s) of interest".

- Measure of first fixation: Metrics related to first fixation are additional important metrics that illustrate visual behaviour. Two metrics were selected regarding first fixation: the time to the first fixation of the AOIs and the first-fixation duration within the AOIs. The time to the first fixation of an area can indicate the saliency of that area [31]. Quickly fixating on a particular area also indicates deliberately directed attention that might be a result of expertise [27]. First-fixation duration indicates the interest of the participant in a particular area, and it also indicates potential difficulties in interpretation.

- Measure of processing: Two fixation-related metrics were selected as the measures of processing: percentage of total fixation duration and percentage of fixation count. The total fixation duration (dwell time) is the sum of the duration of all fixations; the total fixation on an area could indicate interest or difficulty in interpretation [27]. Fixation count can also represent the interest of the participant [22]. Because we selected areas of interest (AOIs) for each map, the percentage of total fixation duration and the fixation count inside the AOIs were calculated as measures of processing.

- Measure of search: The measure of search included two saccade-related metrics, i.e., saccade count and scan-path length (in screen pixels). Saccade count is the number of saccades recorded. More saccades indicate that more effort was spent on searching. Scan-path length is the sum of all scan (saccade) paths, and a longer scan path suggests a less efficient search [31]. (It is noteworthy that the saccade count and scan-path length are not limited to the AOIs; they are calculated from the entire stimuli.)

- General performance: Effectiveness (accuracy) and efficiency are the two major measures used to assess the general performance of a participant [23]. In our case, accuracy was measured based on whether the answer submitted by the participant was correct. Efficiency was measured based on the response time. These metrics were not included in the modelling process, but they were used to evaluate the model (see Section 5).

2.5.2. Bayesian Structural Equation Modelling and Data Imputation

We used structural equation modelling to model the relationship between eye movement metrics and map-reading skills. The structural equation modelling was conducted using the AMOS program [32].

The proposed model to associate eye movement metrics and map-reading skills is presented as "Proposed model" in Figure 1. The ellipses represent the latent variables (e.g., measure of procession, measure of search, measure of first fixation, and map-reading skill), and the rectangles represent the observed variables (e.g., eye movement metrics). We proposed a reflective model where variance in map-reading skills would result in variance in the three measures, and each measure had two eye movement metrics as indicators. Before we fitted the model, all observed variables were standardized. In addition, for ease of interpretation, three variables (i.e., time to first fixation, saccade count, and saccade length) were transformed (by changing the sign of the variables); thus, higher scores indicated better performance [33].

To assess performance, we calculated scores on latent variables for each participant. This was achieved using data imputation. Before the data imputation, we toggled the latent variables into observed variables with missing values [32], as shown in "Model for data imputation" in Figure 1. Bayesian imputation was applied to address the non-numeric variables.
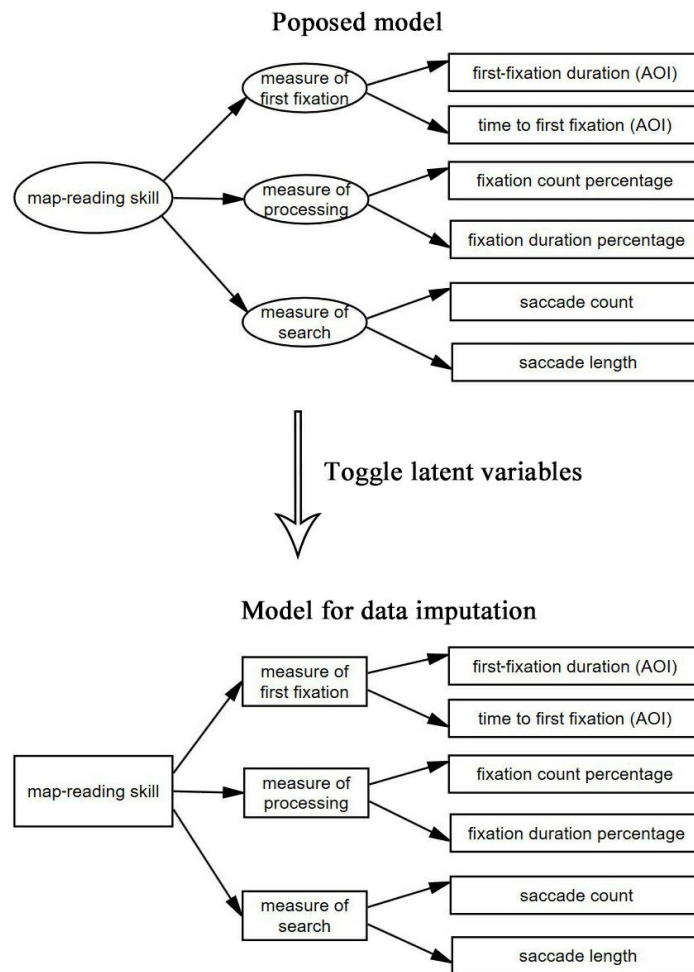
Figure 1. From the proposed model to the model for data imputation using toggling [32].

After imputation, we combined the multiple-imputed results before the regression weights (i.e., path coefficients) were calculated. The multiple-imputed estimation of regression weights was the mean of the estimated regression weights of all completed datasets [32,34].

We used several indices to evaluate the fit of the models. While the most common index is the χ^2 statistic, as a significant χ^2 suggests a poor fit, the χ^2 statistic is correlated with the sample size; thus, this statistic will suggest significance when a large sample size is used, even when the difference between the data and the model is small [35]. Thus, the χ^2/df statistic was adopted (where df stands for "degree of freedom"), and a value less than 2.0 suggested a good fit. Another fit measure is the standardized root mean square error of approximation (RMSEA), where a value less than 0.08 suggests a fair fit, and a value less than 0.05 suggests a good fit [36]. The normal fit index (NFI) compares the proposed model against the null model, and a value above 0.95 is considered to indicate good model fit [37].

3. Results

3.1. Model Fit

The fit measures for each model are shown in Table 3. For all the models, the χ^2/df statistics were less than 2, the NFI was greater than 0.95, and the RMSEA was below 0.05. These results suggest good fits in all five models.

Table 3. Fit measures for each model.

	χ^2	df	χ^2/df	NFI	RMSEA
Task #1	399.091	360	1.109	0.985	0.008
Task #2	690.723	360	1.919	0.975	0.022
Task #3	388.839	360	1.080	0.981	0.007
Task #4	574.751	360	1.597	0.982	0.018
Task #5	534.695	360	1.485	0.984	0.016

3.2. Path Coefficients

According to the proposed model, map-reading skills are reflected by the measure of first fixation, the measure of processing, and the measure of search. The path coefficients for the model after imputation can be found in Table 4. The imputed models with coefficients can be found in the Appendix A (Figures A1–A5).

Table 4. Path coefficients for each model.

Task Path	Task #1	Task #2	Task #3	Task #4	Task #5
Map-reading skill → measure of first fixation	0.341	0.725	0.403	0.548	0.393
Measure of first fixation → time to first fixation (AOI)	0.618	0.708	0.641	0.658	0.665
Measure of first fixation → first-fixation duration (AOI)	−0.249	−0.034	−0.165	0.104	−0.117
Map-reading skill → measure of processing	0.376	0.602	0.482	0.892	0.744
Measure of processing → fixation count percentage	0.978	0.960	0.949	0.973	0.985
Measure of processing → fixation duration percentage	0.981	0.953	0.928	0.971	0.984
Map-reading skill→measure of search	0.097	0.210	0.322	0.301	−0.300
Measure of search → saccade count	0.963	0.979	0.962	0.986	0.970
Measure of search → saccade length	0.980	0.969	0.963	0.990	0.986

The path coefficients from map-reading skill to measure of first fixation varied within the five tasks, ranging from 0.341 to 0.725 (mean = 0.482, SD = 0.139).

The path coefficients from map-reading skill to measure of processing also fluctuated but had greater variance (mean = 0.619, SD = 0.183). These path coefficients had larger absolute values than those from map-reading skill to measure of first fixation and to measure of search, although the differences were not significant.

Regarding the path coefficients from map-reading skill to measure of search, a negative value was reported for task #5 ($\beta 5 = -0.300$), while tasks #1, #2, #3, and #4 all had positive values.

Regarding the measurement model, there was a clear pattern concerning the path coefficients from measure of processing and measure of search to their own indicators. The path coefficients from measure of processing to its indicators (fixation count percentage and fixation duration percentage) were above 0.92 for all five tasks. This was a consistent pattern, as all path coefficients from measure of search to its indicators were greater than 0.96. However, the path coefficients from measure of first fixation to its indicators showed another pattern. The path coefficient from measure of first fixation to time to first fixation was approximately 0.65 ($\beta 1 = 0.618$, $\beta 2 = 0.708$, $\beta 3 = 0.641$, $\beta 4 = 0.658$, $\beta 5 = 0.665$), but the path coefficient from measure of first fixation to first-fixation duration was approximately 0 ($\beta 1 = -0.249$, $\beta 2 = -0.034$, $\beta 3 = -0.165$, $\beta 4 = -0.104$, $\beta 5 = -0.117$).

3.3. Imputed Scores

To evaluate the scores generated by the models, we tested them against the answers and response times of the participants.

3.3.1. Imputed Scores and Answers Given by Participants

For each task, the participants were divided into two groups based on whether they answered the questions correctly or not. After the variance homogeneity test, we performed an independent t-test to examine the differences in the map-reading skill scores between these two groups (Table 5). Participants who answered the questions correctly scored significantly higher in terms of map-reading skill ($mean_1 = 0.002$, $SD_1 = 0.107$; $mean_2 = 0.250$, $SD_2 = 0.229$; $mean_3 = 0.012$, $SD_3 = 0.149$; $mean_4 = 0.176$, $SD_4 = 0.762$; $mean_5 = 0.282$, $SD_5 = 0.0180$) than their counterparts ($mean_1' = -0.065$, $SD_1' = 0.116$; $mean_2' = -0.091$, $SD_2' = 0.333$; $mean_3' = -0.125$, $SD_3' = 0.174$; $mean_4' = *-0.450$, $SD_4' = 0.739$; $mean_5' = -0.171$, $SD_5' = 0.198$). This significant difference was observed for all five tasks ($p_1 = 0.002 < 0.01$, $t_1 = 3.188$; $p_2 = 0.000 < 0.01$, $t_2 = 7.034$; $p_3 = 0.000 < 0.01$, $t_3 = 3.631$; $p_4 = 0.000 < 0.01$, $t_4 = -4.995$; $p_5 = 0.000 < 0.01$, $t_5 = 3.752$; $p3 = 0.000 < 0.01$, $t3 = 3.631$; $p4 = 0.000 < 0.01$, $t4 = -4.995$; $p5 = 0.000 < 0.01$, $t5 = 3.752$).

Table 5. T-test results for map-reading skill scores for the two groups of participants.

		N	Mean	SD	t	p
Task #1	Correct	160	0.002	0.107	3.188	0.002 *
	Incorrect	31	−0.065	0.116		
Task #2	Correct	53	0.250	0.229	7.034	0.000 *
	Incorrect	136	−0.091	0.333		
Task #3	Correct	155	0.012	0.149	3.631	0.000 *
	Incorrect	18	−0.125	0.174		
Task #4	Correct	142	0.176	0.762	4.995	0.000 *
	Incorrect	49	−0.450	0.739		
Task #5	Correct	72	0.282	0.180	3.752	0.000 *
	Incorrect	117	−0.171	0.198		

* $p < 0.05$.

A similar pattern was observed for the measure of processing scores (Table 6). There were significant differences in the measure of processing scores for all five questions. Participants who answered a given question correctly scored significantly higher on the measure of processing ($mean_1 = 0.111$, $SD_1 = 0.909$; $mean_2 = 0.843$, $SD_2 = 0.756$; $mean_3 = 0.041$, $SD_3 = 0.883$; $mean_4 = 0.361$, $SD_4 = 1.503$; $mean_5 = 0.639$, $SD_5 = 1.502$) than their counterparts who answered incorrectly ($mean_1' = -0.633$, $SD_1' = 0.971$; $mean_2' = -0.351$, $SD_2' = 0.792$; $mean_3' = -0675$, $SD_3' = 1.0074$; $mean_4' = -0.994$, $SD_4' = 1.414$; $mean_5' = -0.416$, $SD_5' = 1.650$). While the significance level in task #3 was slightly lower, the differences were generally more significant than those observed for the map-reading skills ($p_1 = 0.000 < 0.01$, $t_1 = 4.130$; $p_2 = 0.000 < 0.01$, $t_2 = -9.425$; $p_3 = 0.001 < 0.01$, $t_3 = 3.210$; $p_4 = 0.000 < 0.01$, $t_4 = 5.520$; $p_5 = 0.000 < 0.01$, $t_5 = 4.359$).

Table 6. T-test results for measure of processing scores for two groups of participants.

		N	Mean	SD	t	p
Task #1	Correct	160	0.111	0.909	4.130	0.000 *
	Incorrect	31	−0.633	0.971		
Task #2	Correct	53	0.843	0.756	9.425	0.000 *
	Incorrect	136	−0.351	0.792		
Task #3	Correct	155	0.041	0.883	3.210	0.001 *
	Incorrect	18	−0.675	1.007		
Task #4	Correct	142	0.361	1.503	5.520	0.000 *
	Incorrect	49	−0.994	1.414		
Task #5	Correct	72	0.639	1.502	4.359	0.000 *
	Incorrect	117	−0.416	1.650		

* $p < 0.05$.

3.3.2. Imputed Scores and Response Times

No strong correlation was observed between map-reading skill scores and response times (Table 7). The absolute values of the correlations between the map-reading skill scores and the response times for all five questions were less than 0.6, and both positive and negative values were reported ($r_1 = -0.073$; $r_2 = -0.358$; $r_3 = -0.531$; $r_4 = -0.431$; $r_5 = 0.273$).

Strong negative correlations were reported between the subjects' scores on the measure of search and response time (Table 7). Negative correlations with absolute values greater than 0.8 were reported for all five questions ($r_1 = -0.917$; $r_2 = -0.927$; $r_3 = -0.926$; $r_4 = -0.928$; $r_5 = -0.845$).

Table 7. Correlations: map-reading skill, measure of search and response time.

	Task #1	Task #2	Task #3	Task #4	Task #5
corr (map-reading skill, response time)	−0.073	−0.358	−0.531	−0.431	0.273
corr (measure of search, response time)	−0.917	−0.927	−0.926	−0.928	−0.845

4. General Discussion

4.1. Eye Movement Metrics and Map-Reading Skills: Path Coefficients

The participants' map-reading skills were primarily reflected by the measure of processing. This result suggests that participants with better map-reading skills focused more on important information, which is consistent with the finding of Ooms et al. that experts tended to fixate more on major structural elements [27].

The path coefficients from map-reading skill to measure of first fixation differed among the tasks, ranging from 0.341 to 0.725. This result suggested that being able to locate important information quickly contributed to better performance. This was supported by the results for task #2 (the topology task), where the path coefficient was 0.725. This result could possibly be explained by the ability of participants with better map-reading skills to quickly identify the wrong answers and select the correct answer. Ooms et al. also suggested that novices were more easily distracted at the beginning of the task and thus were not able to fixate on key information immediately [27]. Additionally, inability to locate key information might suggest that a participant was confused by the map [38].

The path coefficients from map-reading skills to measure of search varied greatly among the five tasks, indicating that participants with better map-reading skills did not necessarily search with higher efficiency. For task #1 (the visualization task), the path coefficient was 0.097, indicating that map-reading skill was not reflected by the measure of search. This result might be explained by the findings of Ooms et al. [27], who concluded that people with different skill levels might still have a similar scan path during certain map tasks. Tasks #2 (the topology task), #3, and #4 (navigation tasks) reported similar path coefficients. This result suggests that for these three tasks, the participants with better skills searched with relatively high efficiency. A negative path coefficient was found for task #5 (the spatial association task). To solve task #5, participants had to determine the spatial association between the two maps. To achieve this, they had to view each map multiple times to match the pattern on one map with that on the other. Frequent switching between maps could result in more saccades and longer saccade paths, which could lead to lower search efficiency. In short, the difference in path coefficients might be explained by differences in the tasks themselves and the skills involved. Stofer and Che suggested that when experts were asked specifically about the visualization, they tended to have more fixations per visualization [28], which indicated greater meaning making [19]. Since more fixations per visualization result in more saccades, this might explain why participants with better map-reading skills did not always search with higher efficiency.

The path coefficients from the measure of processing and measure of search to the eye movement metrics that measured them were all above 0.9. This result is consistent with previous studies showing that large numbers of saccades and lengthy scan paths suggest inefficient searches. However, the two

metrics for the measure of first fixation had very different path coefficients. The path coefficients from the measure of first fixation to time to first fixation were above 0.6 for all five tasks, while the path coefficients from the measure of first fixation to first-fixation duration hovered around approximately 0. This difference in the path coefficients suggests that while these two metrics are both related to first fixation, they may not share a common nature [39].

4.2. Reflections of Accuracy and Efficiency: Imputed Scores

Both the accuracy and efficiency of the participants were reflected in the scores generated by the model. In our model, accuracy weighed more than efficiency in the score, as the path coefficient from the map-reading skill to measure of processing was larger than that from the map-reading skill to measure of search in all tasks. This difference in weight is sensible in practice because selecting the correct answer is typically weighted more than completing the task in a short time.

Instead of the map-reading skill score, the efficiency of the participant was better reflected by the score related to the measure of search. A strong negative correlation was observed between the measure of search score and response time. This result suggested that the efficiency in solving the task was related to the efficiency in searching for information. The results from tasks #3 and #4 (the navigation tasks) are consistent with the finding of Liao et al. that when participants spent less effort searching, they completed the navigation task more quickly [23].

The correlation between the map-reading skill score and response time was weak for task #1, implying that participants who finished the task faster were not necessarily rewarded with higher scores. This result might be explained by the path coefficient from the map-reading skill to measure of search in task #1 (the map visualization task), in which the participants scanned the maps in similar ways.

For task #5 (the spatial association task), the positive correlation between the map-reading skill and response time revealed that participants who completed the task in a shorter time actually received lower scores. A closer examination of their answers revealed that among all 117 participants who submitted an incorrect answer, 68 of them had chosen the incorrect statement (population falls as precipitation decreases), and 44 out of those 68 participants had response times that were shorter than the average. However, while the correct solution (*population first rises but then falls as precipitation continues to increase*) reflected the spatial association for the whole area, this incorrect statement only reflected a part of it. Therefore, we could infer that participants needed to study the details of the map more closely, rather than scanning it briefly, to select the correct answer. Thus, for task #5, many participants answered quickly but incorrectly, ultimately selecting an incorrect answer and achieving a lower map-reading score.

5. Conclusions

This study proposed a model that quantitatively associated map-reading skills with eye movement metrics using structural equation modelling. An eye tracking study was conducted with 258 participants who were responsible for completing five map-reading tasks, including visualization, topology, map-based navigation, and spatial association. Map-reading skills were indicated by the measure of first fixation, measure of processing, and measure of search. The map-reading skill scores for each task were calculated for each participant using Bayesian imputation. The scores were then tested against the answers provided by the participants. The model fitted well for all five tasks. The path coefficients indicated that map-reading skills could be reflected by measures related to first fixation, processing and search. It is noteworthy that eye movement metrics used to assess map-reading skills can be applied to further examples. The scores generated by the model generally reflected the performance of the participants in terms of both accuracy and efficiency.

It should be noted that the eye movement metrics are not limited to the ones we selected. Further analysis with more eye movement metrics (such as mean fixation duration), and perhaps new grouping methods, may contribute to the accuracy of the model. Future work would also include the analysis

of gender effects and the possible effect of the different courses participants have taken. This could be achieved with data from the entry forms of the participants. Additionally, modelling a task with multiple (for example, 3–5) stimuli would increase consistency and confidence. We intend to address this in future experiments, along with the validation of the model with more stimuli.

Conflicts of Interest: The authors declare no conflicts of interest.
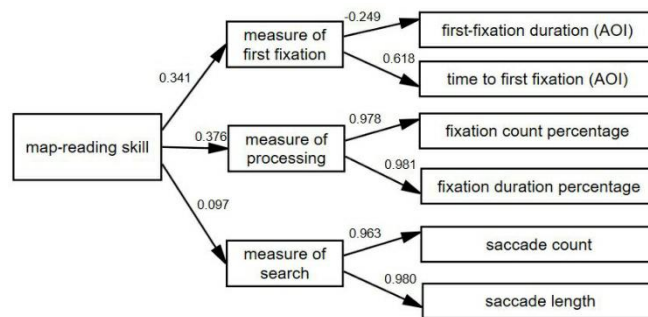
Appendix A

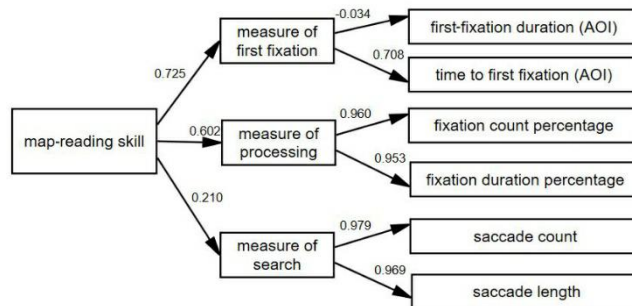**Figure A1.** Path coefficients for task #1.

Figure A2. Path coefficients for task #2.
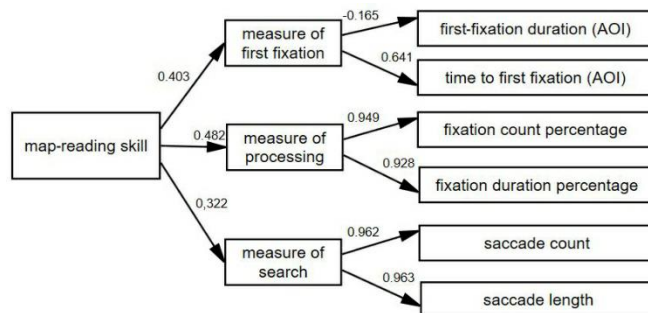

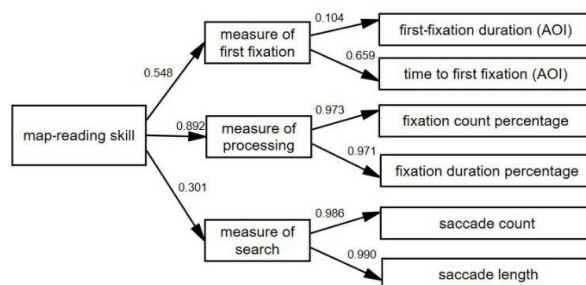**Figure A3.** Path coefficients for task #3.

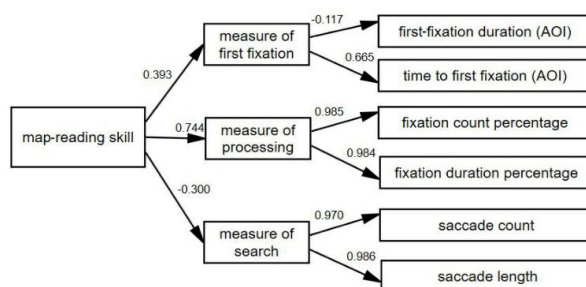**Figure A4.** Path coefficients for task #4.

Figure A5. Path coefficients for task #5.

References

1. Lobben, A.K. Tasks, strategies, and cognitive processes associated with navigational map reading: A review perspective. *Prof. Geogr.* **2004**, *56*, 270–281.
2. Clarke, D. Are You Functionally Map Literate? Available online: lazarus.elte.hu/cet/publications/088.pdf (accessed on 23 August 2018).
3. Ikonovic, V. Importance of Education in Cartography. Available online: https://icaci.org/files/documents/ ICC_proceedings/.../file/f01009.pdf (accessed on 23 August 2018).
4. Phillips, R.J. Are maps different from other kinds of graphic information? *Cartogr. J.* **1989**, *26*, 24–25. [CrossRef]
5. Patterson, T.C. Google Earth as a (not just) geography education tool. *J. Geogr.* **2007**, *106*, 145–152. [CrossRef]
6. Board, C. Map reading tasks appropriate in experimental studies in cartographic communication. *Cartogr. Int. J. Geogr. Inf. Geovis.* **1978**, *15*, 1–12. [CrossRef]
7. Carswell, R.J. The role of the user in the map communication process: Children's abilities in topographic map reading. *Cartogr. Int. J. Geogr. Inf. Geovis.* **1971**, *8*, 40–45. [CrossRef]
8. Gilhooly, K.J.; Wood, M.; Kinnear, P.R.; Green, C. Skill in map reading and memory for maps. *Q. J. Exp. Psychol. A* **1988**, *40*, 87–107. [CrossRef]
9. Pedersen, P.; Farrell, P.; McPhee, E. Paper versus pixel: effectiveness of paper versus electronic maps to teach map reading skills in an introductory physical geography course. *J. Geogr.* **2005**, *104*, 195–202. [CrossRef]
10. Jacovina, M.E.; Ormand, C.; Shipley, T.F.; Weisberg, S.M. Topographic Map Assessment [Measurement Instrument]. 2014. Available online: http://www.silccenter.org/media/silc_pdfs/resources/ testsandinstruments/tandi-new/TMA_Assessment.pdf (accessed on 23 August 2018).
11. Atit, K.; Weisberg, S.M.; Newcombe, N.S.; Shipley, T.F. Learning to interpret topographic maps: Understanding layered spatial information. *Cogn. Res. Princ. Implic.* **2016**, *1*. [CrossRef] [PubMed]
12. Ooms, K.; De Maeyer, P.; Dupont, L.; Van der Veken, N.; Van de Weghe, N.; Verplaetse, S. Education in cartography: What is the status of young people's map-reading skills? *Cartogr. Geogr. Inf. Sci.* **2016**, *43*, 134–153. [CrossRef]
13. Streeter, L.A.; Vitello, D. A profile of drivers' map-reading abilities. *Hum. Factors* **1986**, *28*, 223–239. [CrossRef] [PubMed]
14. Kovach, J.R.C.; Surrette, M.A.; Aamodt, M.G. Following informal street maps: Effects of map design. *Environ. Behav.* **1988**, *20*, 683–699. [CrossRef]

15. Lobben, A.K. Navigational map reading: Predicting performance and identifying relative influence of map-related abilities. *Ann. Assn. Am. Geogr.* **2007**, *97*, 64–85. [CrossRef]

16. Chang, K.; Antes, J.R. Sex and cultural differences in map reading. *Am. Cartogr.* **1987**, *14*, 29–42. [CrossRef]

17. Gilmartin, P.P.; Patton, J.C. Comparing the sexes on spatial abilities: Map-use skills. *Ann. Assn. Am. Geogr.* **1984**, *74*, 605–619. [CrossRef]

18. Rayner, K. Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* **1998**, *124*, 372–422. [CrossRef] [PubMed]

19. Holmqvist, K.; Nyström, M.; Andersson, R.; Dewhurst, R.; Jarodzka, H.; Van de Weijer, J. *Eye Tracking: A Comprehensive Guide to Methods and Measures*; Oxford University Press: Oxford, UK, 2011.

20. Opach, T.; Nossum, A. Evaluating the Usability of Cartographic Animations with Eye-Movement Analysis. Available online: https://scholar.google.com.hk/scholar?q=Evaluating+the+usability+of+cartographic+animations+with+eye-movement+analysis&hl=zh-TW&as_sdt=0&as_vis=1&oi=scholart (accessed on 23 August 2018).

21. Duchowski, A.T. *Eye Tracking Methodology: Theory and Practice*, 3rd ed.; Springer: Berlin, Germany, 2007.

22. Just, M.A.; Carpenter, P.A. Eye fixations and cognitive processes. *Cogn. Psychol.* **1976**, *8*, 441–480. [CrossRef]

23. Liao, H.; Dong, W.; Peng, C.; Liu, H. Exploring differences of visual attention in pedestrian navigation when using 2D maps and 3D geo-browsers. *Cartogr. Geogr. Inf. Sci.* **2017**, *44*, 474–490. [CrossRef]

24. Kiefer, P.; Giannopoulos, I.; Raubal, M.; Duchowski, A. Eye tracking for spatial research: Cognition, computation, challenges. *Spat. Cogn. Comput.* **2017**, *17*, 1–19. [CrossRef]

25. Hegarty, M.; Just, M. Constructing mental models of machines from text and diagrams. *J. Mem. Lang.* **1993**, *32*, 717–742. [CrossRef]

26. Ooms, K.; De Maeyer, P.; Fack, V.; Van Assche, E.; Witlox, F. Interpreting maps through the eyes of expert and novice users. *Int. J. Geogr. Inf. Sci.* **2012**, *26*, 1773–1788. [CrossRef]

27. 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]

28. Stofer, K.; Che, X. Comparing experts and novices on scaffolded data visualizations using eye-tracking. *J. Eye Mov. Res.* **2014**, 7. [CrossRef]

29. Hermans, O.; Laarni, J. Searching Information from Screen Maps. Available online: https://pdfs.semanticscholar.org/32c1/1023fb69eaadea4b3fcdb290120a2d3c9d2b.pdf (accessed on 23 August 2018).

30. Golledge, R.G. The nature of geographic knowledge. *Ann. Assn. Am. Geogr.* **2002**, *92*, 1–14. [CrossRef]

31. Goldberg, J.H.; Kotval, X.P. Computer interface evaluation using eye movements: methods and constructs. *Int. J. Ind. Ergonom.* **1999**, *24*, 631–645. [CrossRef]

32. Arbuckle, J.L. IBM SPSS AMOS 22 Users' Guide: IBM Corp. Available online: https://www.google.nl/url?sa=t&source=web&rct=j&url=http://www.sussex.ac.uk/its/pdfs/SPSS_Amos_User_Guide_22.pdf&ved=2ahUKEwiSgt2k4YXdAhVN46QKHZHTDKQQFjAAegQIAhAB&usg=AOvVaw0ok61Qf--VXoQmnvZc6glH (accessed on 23 August 2018).

33. Hegarty, M.; Montello, D.R.; Richardson, A.E.; Ishikawa, T.; Lovelace, K. Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence* **2006**, *34*, 151–176. [CrossRef]

34. Schafer, J.L. *Analysis of Incomplete Multivariate Data*; Chapman and Hall/CRC: New York, NY, USA, 1997.

35. Kline, R.B. *Principles and Practice of Structural Equation Modelling*, 4th ed.; Guilford publications: New York, NY, USA, 2015.

36. Hu, L.; Bentler, P.M. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychol. Methods* **1998**, *3*, 424–453. [CrossRef]

37. Bentler, P.M.; Bonett, D.G. Significance tests and goodness of fit in the analysis of covariance structures. *Psychol. Bull.* **1980**, *88*, 588–606. [CrossRef]

38. 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]

39. Coltman, T.; Devinney, T.M.; Midgley, D.F.; Venaik, S. Formative versus reflective measurement models: Two applications of formative measurement. *J. Bus. Res.* **2008**, *61*, 1250–1262. [CrossRef]