A comparison of the performance on extrinsic and intrinsic cartographic visualizations through correctness, response time and cognitive processing

Čeněk Šašinka[1], Zdeněk Stachoň[1,2]*, Jiří Čeněk[1], Alžběta Šašinková[1], Stanislav Popelka[3], Pavel Ugwitz[1], David Lacko[1]

1 Department of Information and Library Studies, Faculty of Arts, Masaryk University, Brno, Czech Republic, 2 Department of Geography, Faculty of Science, Masaryk University, Brno, Czech Republic, 3 Department of Geoinformatics, Faculty of Science, Palacký University Olomouc, Olomouc, Czech Republic

* zstachon@geogr.muni.cz

Abstract

The aim of this study was to compare the performance of two bivariate visualizations by measuring response correctness (error rate) and response time, and to identify the differences in cognitive processes involved in map-reading tasks by using eye-tracking methods. The present study is based on our previous research and the hypothesis that the use of different visualization methods may lead to significant cognitive-processing differences. We applied extrinsic and intrinsic visualizations in the study. Participants in the experiment were presented maps which depicted two variables (soil moisture and soil depth) and asked to identify the areas which displayed either a single condition (e.g., "find an area with low soil depth") or both conditions (e.g., "find an area with high soil moisture *and* low soil depth"). The research sample was composed of 31 social sciences and humanities university students. The experiment was performed under laboratory conditions, and Hypothesis software was used for data collection. Eye-tracking data were collected for 23 of the participants. An SMI RED-m eye-tracker was used to determine whether either of the two visualization methods was more efficient for solving the given map-reading tasks. Our results showed that with the intrinsic visualization method, the participants spent significantly more time with the map legend. This result suggests that extrinsic and intrinsic visualizations induce different cognitive processes. The intrinsic method was observed to generally require more time and led to higher error rates. In summary, the extrinsic method was found to be more efficient than the intrinsic method, although the difference was less pronounced in the tasks which contained two variables, which proved to be better suited to intrinsic visualization.

Humanities Laboratory, Faculty of Arts, Masaryk University.

Introduction

The awareness that maps serve as tools for the creation of mental representations of the world and cannot therefore be considered transparent or direct depictions has long been discussed in cartography [1]. As a research topic, the cognition of maps is rooted in the early twentieth century [2]. A key question is how a particular form of cartographic visualization affects the effectiveness of cartographic communication [3–5]. The same data can be represented by different cartographic visualization methods. An unsuitable method not only reduces performance but also places various requirements which correspond to the type of cartographic visualization on different types of users and tasks [6, 7]. User characteristics (such as cartographic skills, [8, 9]) and the type of task [10, 11] must therefore be considered when we conduct empirical studies on the performance of alternative visualizations.

The primary aim of the present study was an empirical and objective comparison of two alternative bivariate visualizations (Fig 1) to assess the performance of a selected population which possessed a basic level of cartographic skill [12–16] through two different types of task. Another aim was to understand the cognitive processes which underlie the potential differences in objective performance [17–20].

Olson [21] stressed that maps are considered highly valuable visual stimuli in experimental psychology since the variables they represent can be accurately controlled. The manner of presenting geographic information can have a significant effect on user cognitive processing (internal mental processes) during map-related tasks. Larkin and Simon [22] presented the concept of informational and computational equivalence and argued that different visualizations can be informationally equivalent if all the information available in one of them is available in the other, and vice versa. The establishment of informational equivalence between bivariate cartographic visualizations permits us to investigate the extent of computational equivalence between the two.

Cartographic visualization offers numerous methods of presenting geographical data. These methods differ in their ability to visualize certain data types, the level of detail they provide, and the number of variables they simultaneously portray [23]. The graphic display of multiple geographic phenomena is known as multivariate mapping [24], and its purpose is to investigate the relationships between the given phenomena. Bivariate maps encode two separate variables simultaneously [25]. Bivariate mapping can be further divided into extrinsic (the variables carrying the information are visually separable) and intrinsic (the variables are visually inseparable [26]).

The present study applies both extrinsic and intrinsic bivariate encoding of geographic variables (Fig 1) to investigate the cognitive processes of map users.

The extrinsic bivariate method employs two visually distinct variables (the differences may represent, for example, size, shape or color lightness) to display two different geographical phenomena, such as soil depth and moisture. In the present study, each of the two phenomena had three levels of intensity (low, medium and high), which provided a total of six options in the map legend (Fig 1, EN1 left). From the map legend, the map users were required to create a mental representation of nine possible combinations (Fig 1, EN2 center). Intrinsic bivariate visualizations apply visual variables which are visually inseparable (typically, the visual variables include hue, color lightness and opacity), resulting in a map legend comprising nine combinations (Fig 1, IN right). In this latter case, the map legend was identical to the mental representation of all the possible combinations. Although, color lightness, hue and opacity are considered to interact with each other in the psychology of perception [27–29], cartography regards them as mutually independent entities [30]. Therefore, in cartography, these parameters are used as independent visual variables.
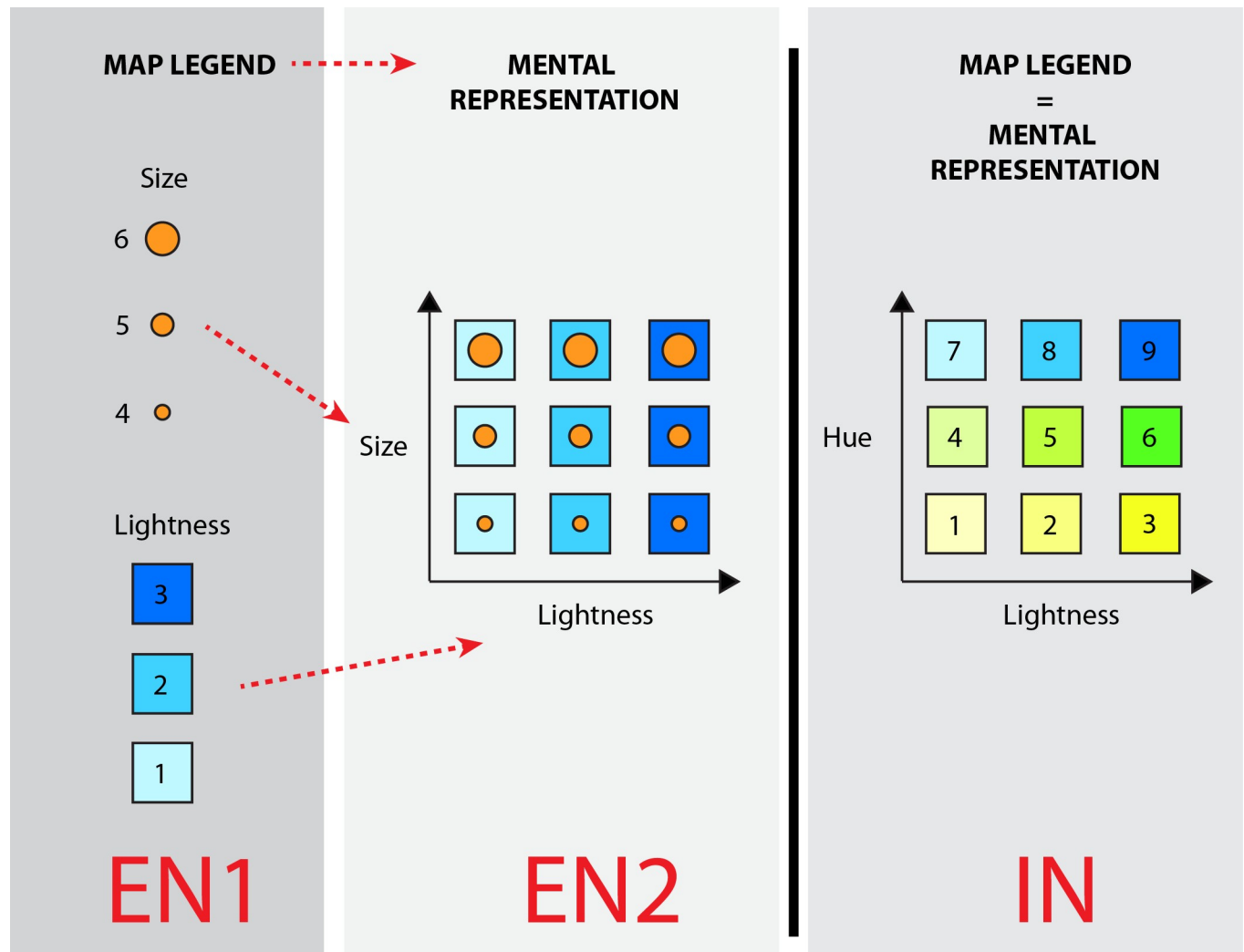
Fig 1. Examples of extrinsic and intrinsic bivariate encoding of geographic variables. EN1: extrinsic (separable) encoding of variables according to size and color lightness; EN2: mental representation of extrinsic visualization (all the possibilities); IN: intrinsic (inseparable) encoding of variables according to hue and lightness.

Each visualization type can be expected to induce a different cognitive and perceptual load on the user [31–33]. The differences in cognitive processing relate to selective attention theory, which specifies that only a limited number of elements can be processed at one time [34]. The perception aspect can be explained according to pre-attentive visual processing theory [35, 36]. Some visual elements, designated pre-attentive, can be detected in a single glance and thereby serve as the central components of a visualization. In map reading, pre-attentive elements can aid in identifying boundaries and detecting the presence or absence of other elements; for example, size (an extrinsic variable) is pre-attentive, while lightness (an intrinsic variable) cannot be considered a pre-attentive feature. Since the processing of extrinsic and intrinsic visual elements is not only based on perception but involves a broader cognitive context, it appears reasonable to assume that the situation will be more complex when both extrinsic and intrinsic visual variables are employed.

Bivariate mapping and the use of various visual variable combinations have been the subject of numerous research studies [e.g., 21, 37–42]. Elmer conducted an extensive comparison of

Fig 2. Examples of extrinsic visualization (left) and intrinsic visualization (right) used in the study. Legends for both extrinsic and intrinsic visualization are also given. Areas with identical values are depicted with two different encoding systems to enable a visual comparison of the differences between each visualization.

visual variable combinations [43]. Kunz studied the use of bivariate visualization methods (extrinsic and intrinsic) to produce visualizations of natural hazards (avalanches) and the levels of uncertainty in the presented data (avalanche hazard prediction) [44]. Šašinka et al. investigated the differences in processing intrinsic and extrinsic visualizations, focusing mainly on cognitive style and map reading skills [45]. The results of the study (and of related eye-tracking studies) revealed significant differences between extrinsic and intrinsic visualizations associated with both map-reading performance and task processing. Although the aim of the study was not to compare methods of visualization, the results showed that a group of laypersons (psychology students) worked more effectively with the extrinsic method, while participants with better map reading skills (cartography students) demonstrated better performance using intrinsic visualization.

However, the authors noted an important limitation in their study, consisting in a relatively sophisticated topic (avalanche risk and its uncertainty) which the participants (especially psychology students) may have found difficult to understand.

The present study was designed based on the results of the above studies. The stimulus material included two bivariate maps with the same content; an example is shown in Fig 2. We selected soil depth and soil moisture as suitable phenomena for depiction since they are considered generally comprehensible and quantifiable. Information gathered from volunteers informed our selection of the topic during the experiment design process. It was critical that participants intuitively understood the relationship between the visual variables and the depicted phenomena in the legend's design. We followed the principle of cultural metaphors and applied a visual representation of the data to match the metaphors which aid conceptual thinking [6, 46]. We used a combination of color and size for extrinsic visualizations and different colors for intrinsic visualizations. Fig 2 illustrates examples.

Methods

To compare the performance of working with maps which use different cartographic visualizations, we conducted an experiment with two tests (one for each of the two selected methods of visualization); for more details of the research design, see Fig 5. We applied a combination of confirmatory and exploratory data analysis methods [45, 47, 48].

The confirmatory analysis tested our hypotheses on the differences between extrinsic and intrinsic visualizations in map reading performance. The data collected were response time and correctness (as investigated by Elmer [43]). Several evaluation methods and concepts allow the measurement of user performance with an information system [49]. The most common parameters are effectiveness and efficiency. According to ISO 9241–11 [50], effectiveness is defined as the "accuracy and completeness with which users achieve specified goals", and efficiency corresponds to the necessary resources (e.g., time) to achieve a desired result. We

calculated effectiveness as the rate of correctness and efficiency as the task completion time [51]. The aim of the exploratory analysis of the eye-tracking data [52] was to gain deeper insight into the differences between the visualizations at the level of individual elements and to employ eye-tracking as a means of collecting objective data [53–55]. Eye-tracking is a valuable tool for studying eye behavior which occurs during map reading since it provides objective measurement of the visual strategies employed by map readers. The review article from Krassanakis and Cybulski [56] provides an overview of existing eye-tracking studies which have appeared in cartographic research over the last decade. The review showed that cartographers used eye tracking mainly in the evaluation of cartographic symbolization and design principles.

Map and items design

The task layout was identical in both tests (Figs 3 and 4): instructions for the tasks were displayed in the upper area of the screen, the map legend was at the right, and the visual field of the map was in the center. The lower area of the screen displayed a button bar with four possible selections for the correct answer. The participants selected an area which satisfied the given condition (e.g., "Find the area with low soil moisture."). In subtest A (Fig 3), the marked areas covered four square units; in subtest B, the marked area only covered one square unit (Fig 4). To answer the questions, participants were required to click on the correct button. Only one correct answer was possible.

We generated the visualization using ArcMap (version 10.7) using the color schemas from ColorBrewer 2.0 [57]. The extrinsic visualization used three circle sizes (6, 10 and 14 pts; #deebf7) to indicate soil moisture, and three color classes (#fee8c8, #fdbb84 and #e34a33) to indicate soil depth. The colors were selected to suit a realistic representation of the phenomena as they occurred in reality, such as blue for moisture and brown for soil depth. Three colors were used to create the intrinsic visualization (A: #e0f3db, #a8ddb5, #43a2ca; B: #e0ecf4, #9ebcda, #8856a7; C: #fee8c8, #fdbb84, #e34a33). A brown color scheme was used to indicate dry areas, and green-blue was used to indicate wet areas. Soil depth was indicated using a geographical principle, darker shades representing greater depth. All colors had a transparency of 40% to allow the base map to be visible. To create the base map, OpenStreetMap data was used [58].

Procedure

The study was designed to illustrate the effect of various types of task. As mentioned in the introduction, we evaluated the maps / visualizations according to their purpose. We therefore designed the study to depict two types of phenomena. In the first scenario, the aim was to answer a question which related to only one variable (either soil moisture or soil depth). In the second scenario, participants were required to think about both phenomena in parallel. We assumed that the extrinsic method would be more suitable for an isolated assessment of phenomena because of its properties (both variables are presented separately through different visual qualities). The intrinsic method, however, is relatively more suitable for tasks which involve a unified search. Another reason for diversifying the task types (division into subtests 1 and 2) was to produce greater informative value and reliability in the achieved results. If performance of the extrinsic method possessed greater stability for each of the types in all tasks, the assumption that this method produces better results from the examined lay population would be more strongly supported.

The test involved a total of 30 items. The first parts of each subtest (A.1 and B.1) contained six items which focused on a single phenomenon (Fig 3). The items covered six possible
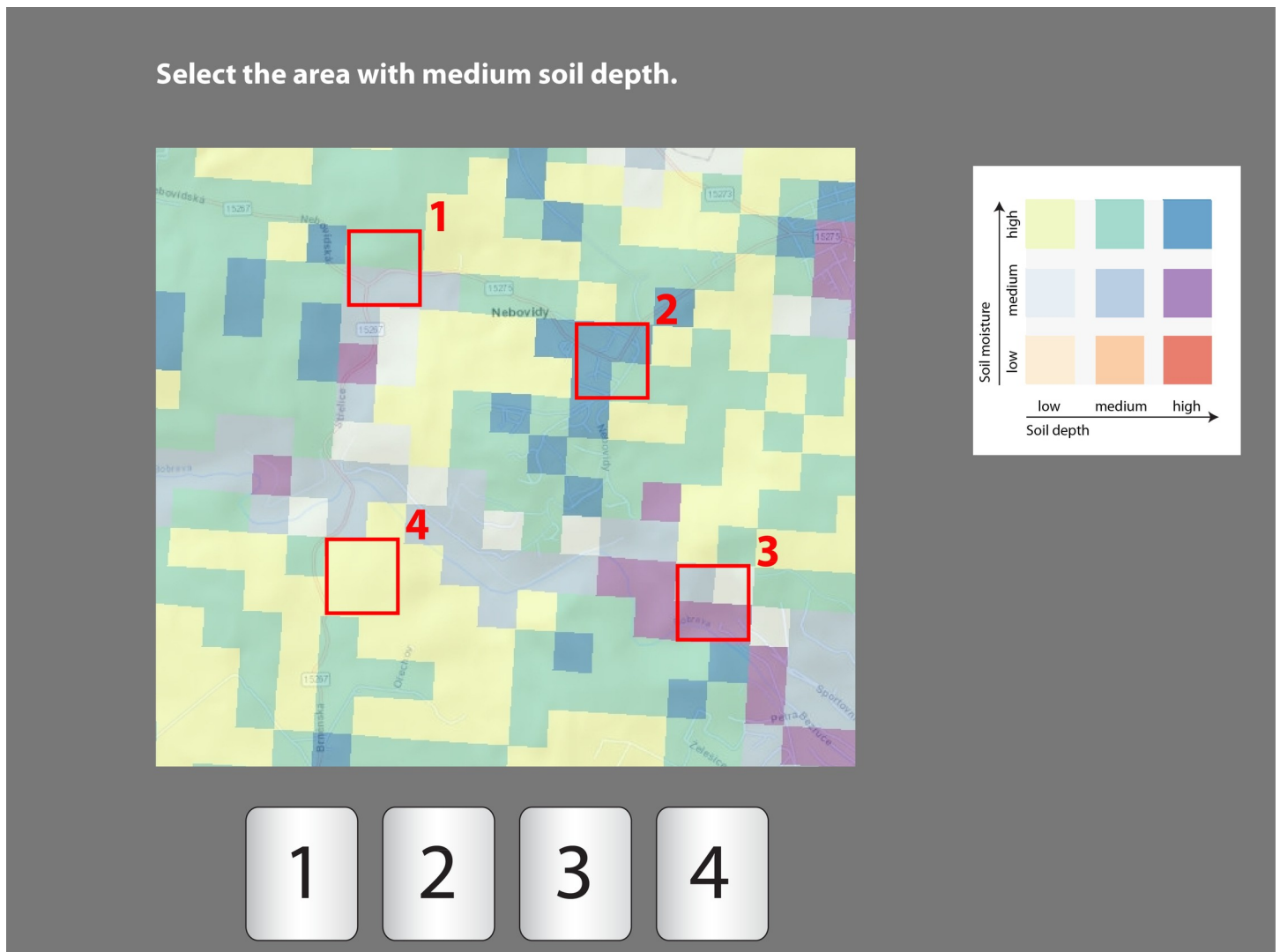
Fig 3. Example of an intrinsic visualization item (subtest A–part A.1.). The task was to select the area which contained "medium soil depth"; the correct answer was area No. 1.

https://doi.org/10.1371/journal.pone.0250164.g003

options: low, medium and high soil moisture, and low, medium and high soil depth. In the second parts of each subtest (A.2 and B.2), participants were asked questions about the two phenomena in each item item (Fig 4), with both A.2 and B.2 covering nine options (low moisture and low soil depth, low moisture and medium soil depth, etc.). We employed a between-subject design (Fig 5) to eliminate the effect of interference caused by experience with the given type of task.

Each participant completed an informed consent form, received a financial reward and was randomly allocated to one of the two research groups. Before the experiment, they were informed about the expected duration of the tests and given the opportunity to ask the experimenter questions. The instructions required the participants to work without interruption during the assessed part of the experiment. No participant required any additional explanation, and during a brief follow-up inquiry, no participant reported any problem in comprehending the content of the tasks. The participants received feedback on the correctness of their
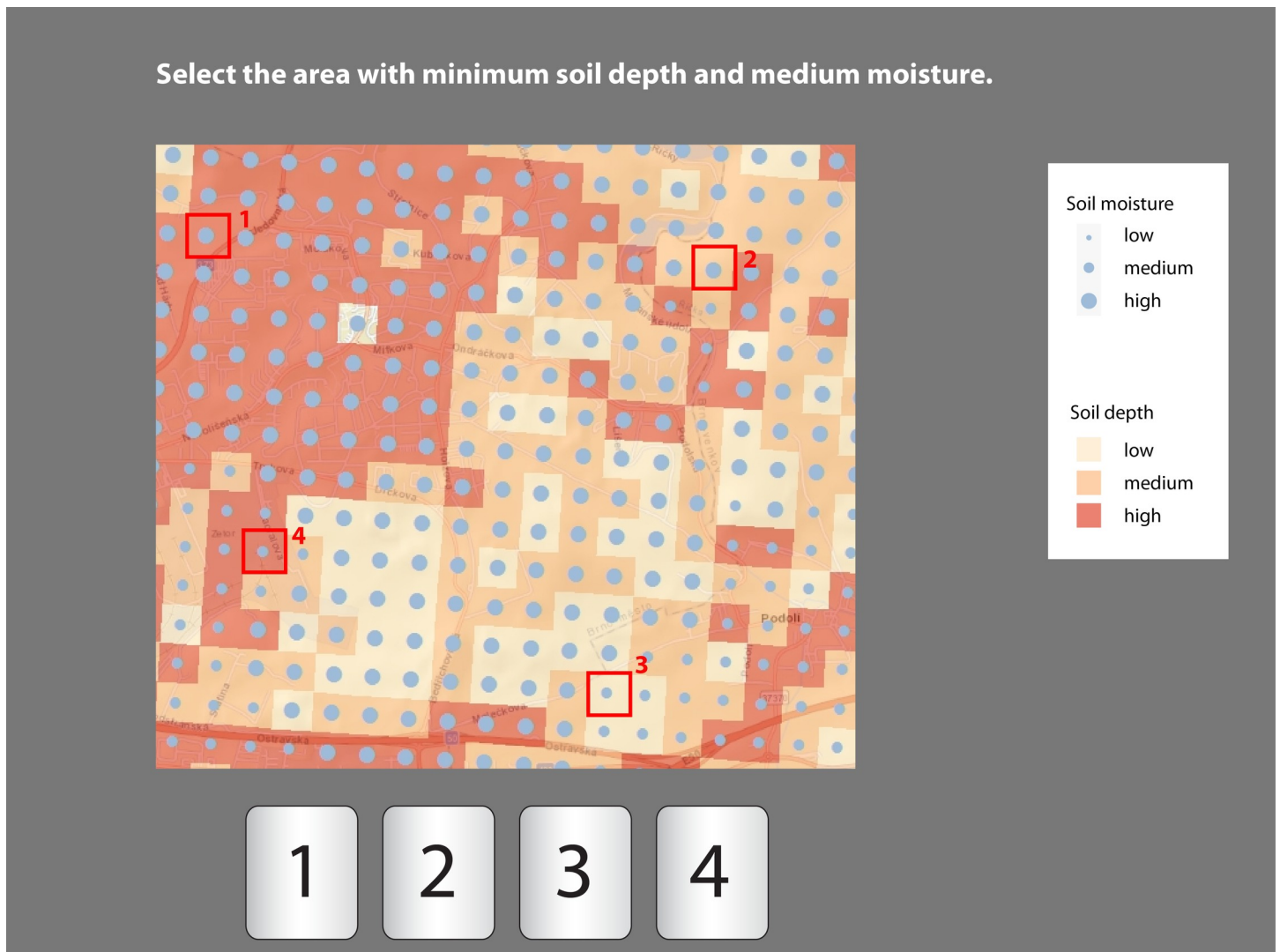
Fig 4. Example of an extrinsic visualization item (subtest B—part B.2). The task was to select the area which best satisfied the conditions of "low soil depth" and "medium soil moisture"; the correct answer was area No. 3.

responses for the two sample items (one sample item was presented at the beginning of each subtest, A and B). No feedback was given during the assessed component of the tests. Each test item was preceded with a fixation cross displayed for 500 ms in the same position each time in the upper area of the screen.

Apparatus

The test was administered using a DELL Precision M4800 notebook with a 22″, 60 Hz AOC E2260P external monitor. The resolution was set at 4:3 (1024 x 768) to correspond exactly to the stimuli (Figs 3 and 4). The participants used a mouse to select their answers. The experimenter was present throughout the experiment to monitor its course. Mounted to the monitor was a remote SMI RED-m eye-tracker with a sampling rate of 60 Hz to collect eye-tracking data. Eye-tracking data collection, calibration and validation was done using the SMI Experiment Center 3.7 software. The calibration procedure was only considered satisfactory when
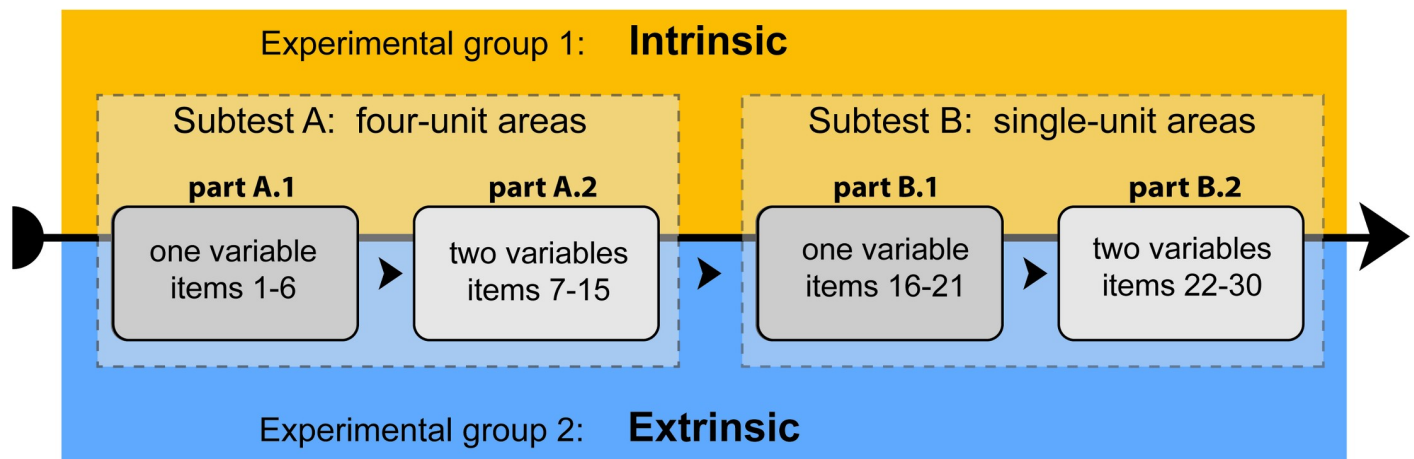
Fig 5. Between-subject experimental design (subtests A and B; parts A.1, A.2, B.1 and B.2). Both independent experimental groups, Intrinsic and Extrinsic, performed the test in exactly the same manner. The order of all items was constant for both groups, and all participants.

the values returned by the eye-tracker were within 0.5˚. The experiment was administered using the Hypothesis software tool [45, 53] (a web-based tool used in research and psychological diagnostics [59]). The behavioral raw data were exported from Hypothesis in ".xlsx" format and then processed using *R* (version 4.0.0) with the "rstatix" [60], "rcompanion" [61] and "multicon" [62] packages. Because of the relatively small sample size, we incorporated several specific procedures in our analyses. First, we used non-parametric statistical tests (i.e., Wilcoxon's rank-sum test for independent samples and Wilcoxon's signed-rank test for paired samples), which do not require Gaussian data distribution and can process potential outliers. Second, we reported not only the related effect sizes (i.e., rank-biserial correlation; *r*) but also their 95% confidence intervals (*CIs*), which were computed on the basis of 10,000 bootstraps. Third, we computed 95% *CIs* for the descriptive statistics of means and medians. This step gave us deeper insight into the obtained results, especially with respect to the small sample size since *CIs* tend to be very wide in small samples, and therefore for reliability, any potential significant differences should not be permitted to overlap. Eye-tracking data were imported into the OGAMA 5.0 software and paired with the behavioral data via HypOgama [53]. The fixations were calculated using the I-DT model with the parameters set to the following values (as recommended by Popelka et al. [53]): maximum distance = 20 px, minimum number of samples = 5; "do not merge consecutive fixations".

Participants

The Research Ethics Committee of Masaryk University approved this project (No.: 0257/ 2018). Participants were recruited via social networks and each signed an informed consent form. They received a financial reward (approx. 8 euros) for participation in the study.

The research sample was composed of 31 students (8 males and 23 females), aged between 19 and 28 (m = 21.8, med = 21). The sample was randomly divided into an "intrinsic" group and an "extrinsic" group (block randomization was used). The former (intrinsic) group consisted of 15 students (2 males and 13 females; m = 21.4). The extrinsic group consisted of 16 students, 6 males and 10 females (m = 22.3). All the participants were students of social sciences and humanities (Faculty of Arts or Faculty of Social Studies) at Masaryk University. Students of geography and related fields were excluded from the study.

The eye-tracking part of the study yielded 23 datasets; for the remainder of the participants (8), no data were recorded during the session for technical reasons. The data were from 4 males and 19 females, aged between 19 and 28 (m = 22.22, med = 22). The "extrinsic" group was composed of 12 students; the "intrinsic" group consisted of 11 students.

After completing the experiment, we performed a quality check of the eye-tracking data. The total data loss was 2.65% for the extrinsic group and 4.1% for the intrinsic group. All items with a dropout rate of above 10% were excluded from the analysis: this was 22 data points (out of a total 330 data points) in the case of the intrinsic method and 6 data points (out of a total 360 data points) in the case of the extrinsic method. No participant was excluded completely (because of a high dropout rate throughout the test).

Results

We used several metrics which employ extrinsic and intrinsic methods of visualization to evaluate the differences between the groups in participant performance. We examined both behavioral (correctness, response time) and eye-tracking (dwell time, direct saccades) metrics. Details of the metrics calculations are specified in the respective section of the Results chapter. Non-parametric statistics were used to calculate the differences between and within the groups. A Wilcoxon rank-sum test was applied to compare independent groups (i.e., extrinsic vs. intrinsic); a Wilcoxon signed-rank test for dependent samples was used to compare performance between subtests. Effect size (r) was calculated for all results to determine the size of the differences [63].

A post-hoc sensitivity analysis of the differences between two independent means according to G*Power [64] ($1\text{-}\beta = 0.80$, $\alpha = 0.05$, $n1 = 16$, $n2 = 15$, two-tailed) showed that with the given sample, we would only be able to detect medium to large effect sizes with differences between the two groups greater than a standard deviation of 1 (non-centrality parameter $\delta = 2.899$, critical $t = 2.8987$, $df = 29$, $d = 1.042$). We therefore did not interpret any results with small effect sizes.

Split-half reliability coefficients performed on two random halves and adjusted with the Spearman-Brown prophecy formula were also calculated for each subtest. The results indicated that all the task subtests were reliable (mean of the split-half correlations for A1 = 0.847, A2 = 0.835, B1 = 0.889, and B2 = 0.736).

Correctness

Response correctness was one of the key parameters observed in the map-related tasks. Using the Wilcoxon rank-sum test, we compared the overall correctness of the responses related to the extrinsic and intrinsic groups. The extrinsic visualization showed a significantly higher overall correctness (N = 16, 96.3%) than the intrinsic visualization (N = 15, 90.0%), with a moderate effect size ($Z = 173$, $p = 0.031$, $r = 0.390$ [95% CI: .059, .656]). The results are charted in Fig 6.

We also investigated incorrect responses to explore the error rate at the level of individual items (i.e., the distractors selected). Particular attention was given to items with a significant difference between the two visualizations, namely items No. 1, 2, 9, 21 and 28 (Fig 7). In the case of all items with the exception of No. 2, the intrinsic method was associated with higher error rates (item No. 2 showed a reverse scenario). A plausible explanation of the above phenomenon was identified only with respect to item No. 21 (Fig 7). The item required the participants to select the area with the lowest soil depth. The correct answer was unit No. 1 (lowest soil depth/highest moisture). In the intrinsic visualization, participants tended to select unit No. 3 (medium soil depth/medium soil moisture), which can likely be explained by unit No. 3 being surrounded by a darker color and thus appearing lighter (see [65–67]) and could
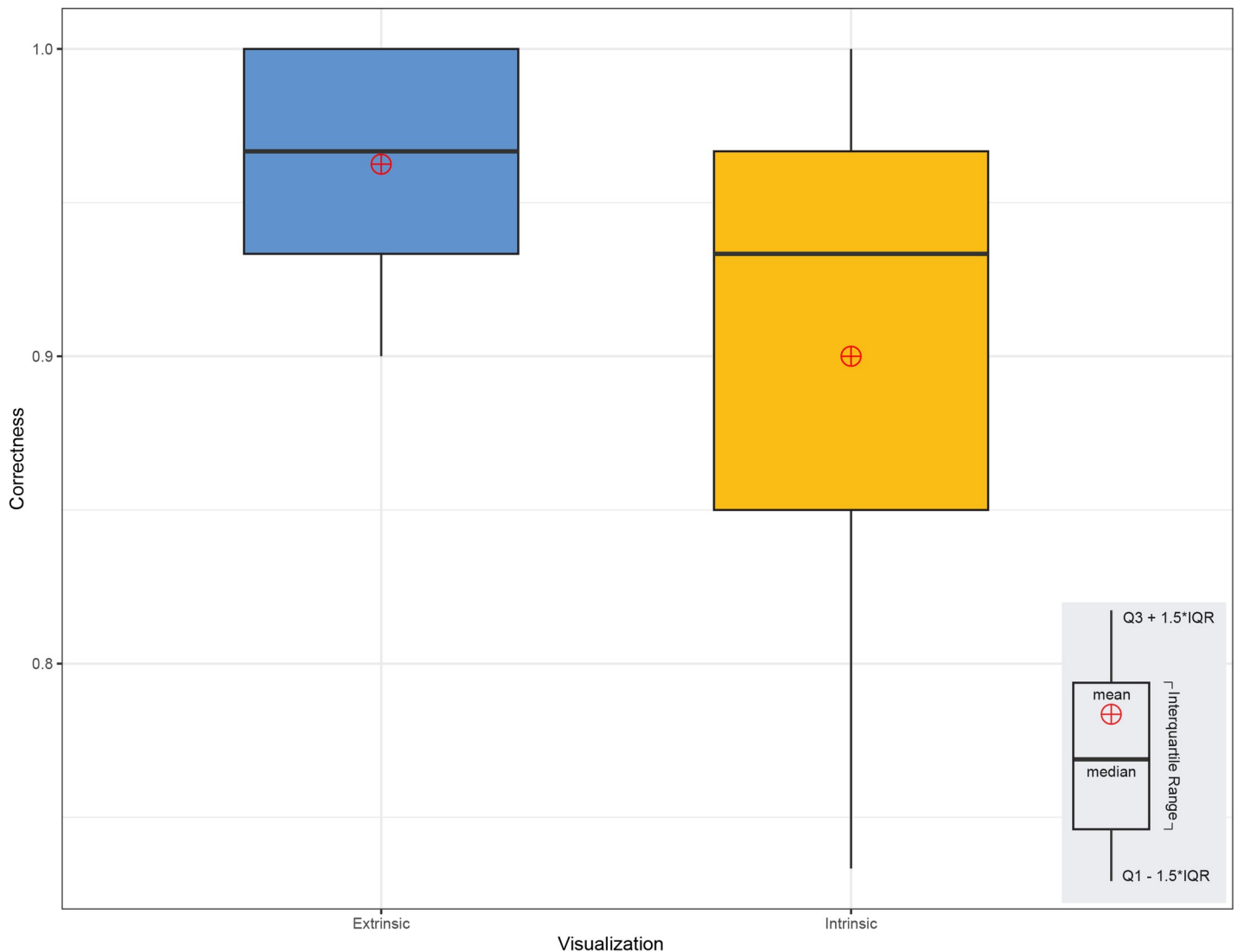
Fig 6. Response correctness for the entire test. Correctness was calculated as a ratio of the number of correct answers to the number of all answers.

therefore have been misinterpreted as the neighboring value (lowest soil depth/medium soil moisture). We identified no other trends.

Response time

For a comparison of processing speeds (response times; RTs), we applied the Wilcoxon rank-sum test. For each subtest, we performed a separate univariate outlier analysis. The analysis revealed three cases of extremely long and irregular response times (over 20,000 ms, different participants) and were excluded from further analysis. However, reaction times are usually distributed ex-Gaussian and demonstrate a rapid rise on the left and have a long positive tail on the right [68, 69]; the traditional outlier detections (e.g., ± 2 SD or 1.5 IQR) are therefore not recommended [70] since these extreme values should not be understood as outliers. Hence, we decided to keep the remainder of the outliers and applied non-parametric statistical analyses
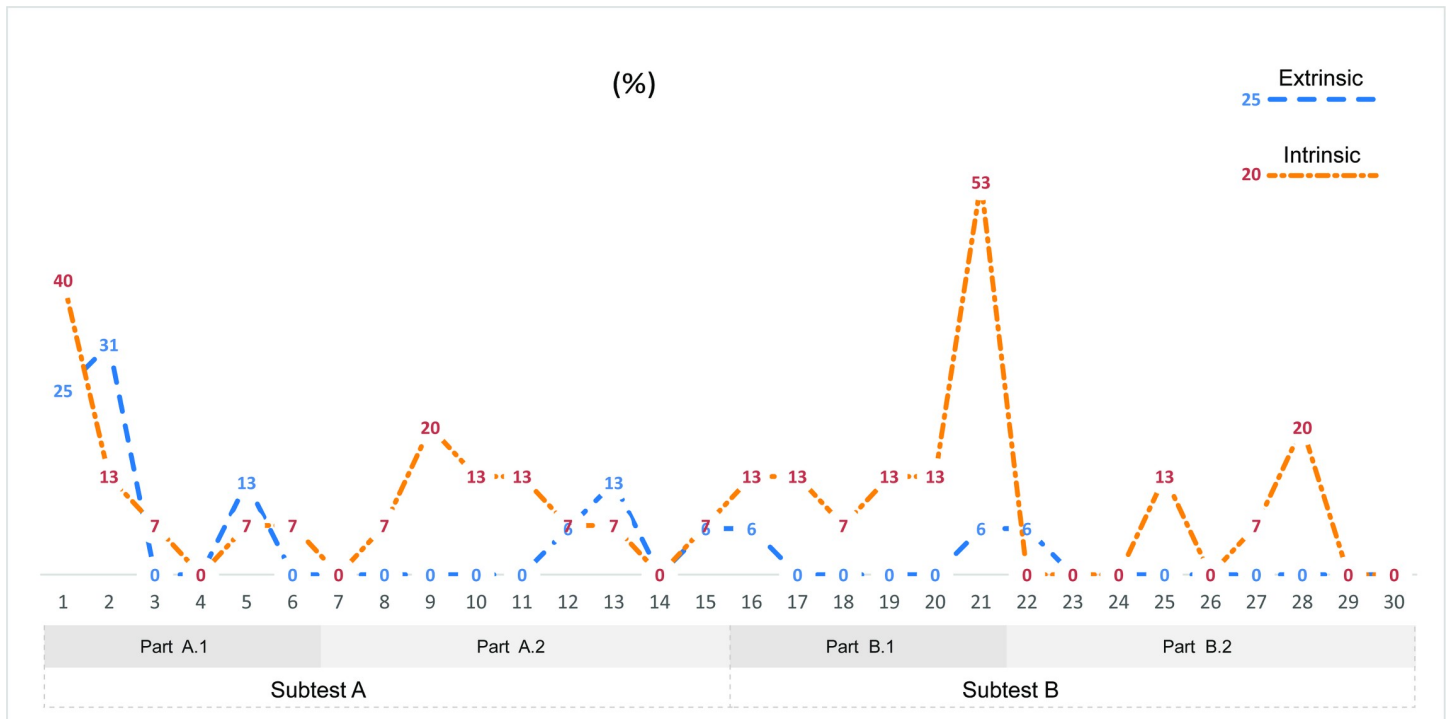
Fig 7. Error rate per item. INT–red/yellow, EXT–blue. Particular attention was given to items with a significant difference between the two visualizations (1, 2, 9, 21, 28). The error rate was calculated as a percentage of incorrect answers of all answers.

instead. The response time analysis therefore covered both correct and incorrect answers. The total response time was significantly less for the extrinsic method (N = 16, median = 6.494 ms [95% CI: 5536, 8726]) than for the intrinsic method (N = 15, median = 10.217 ms [95% CI: 9588, 11597]), with a large effect size (Z = 12, $p < 0.001$, $r = -0.767$ [95% CI: -0.844, -0.607]; see Fig 8 and Table 1).

The same pattern was observed in a comparison of the RTs of individual subtests. The extrinsic stimuli consistently indicated lower RTs than the intrinsic stimuli. We identified the largest differences between visualizations in parts A1 and B1; the differences between visualizations in parts A2 and B2 were moderate. All the differences, with the exception of those related to A2, were significant at a significance level of 5% (Table 1). All the differences, with the exception of those related to A2 and B2, yielded large effect sizes; we also observed large gaps in the upper bounds in the confidence intervals of the extrinsic group and the lower bounds of the confidence intervals in the intrinsic group, suggesting that the obtained statistically significant results were reliable.

At the individual subtest levels (A1, A2, B1, B2), we examined the differences between the test items with one and two variables using the Wilcoxon signed-rank test. An exploration of response times at the subtest level revealed an interesting pattern (Fig 9). In the extrinsic "A" levels, A2 (two variables) resulted in significantly longer response times than A1 (one variable), with a large effect size (Z = 7, $p < 0.001$, $r = -0.789$ [95% CI: -0.880, -0.558]). We observed a similar effect in relation to the extrinsic "B" levels, where B2 showed significantly longer response times than B1 (Z = 0, $p < 0.001$, $r = -0.880$ [95% CI: -0.882, -0.879]). However, we noted an inverse pattern in relation to the intrinsic visualizations, where A1 (one variable) received significantly longer response times than A2 (large effect size; Z = 68, $p = 0.021$, $r = 0.655$ [95% CI: 0.227, 0.886]), and similarly, B1 resulted in significantly longer response
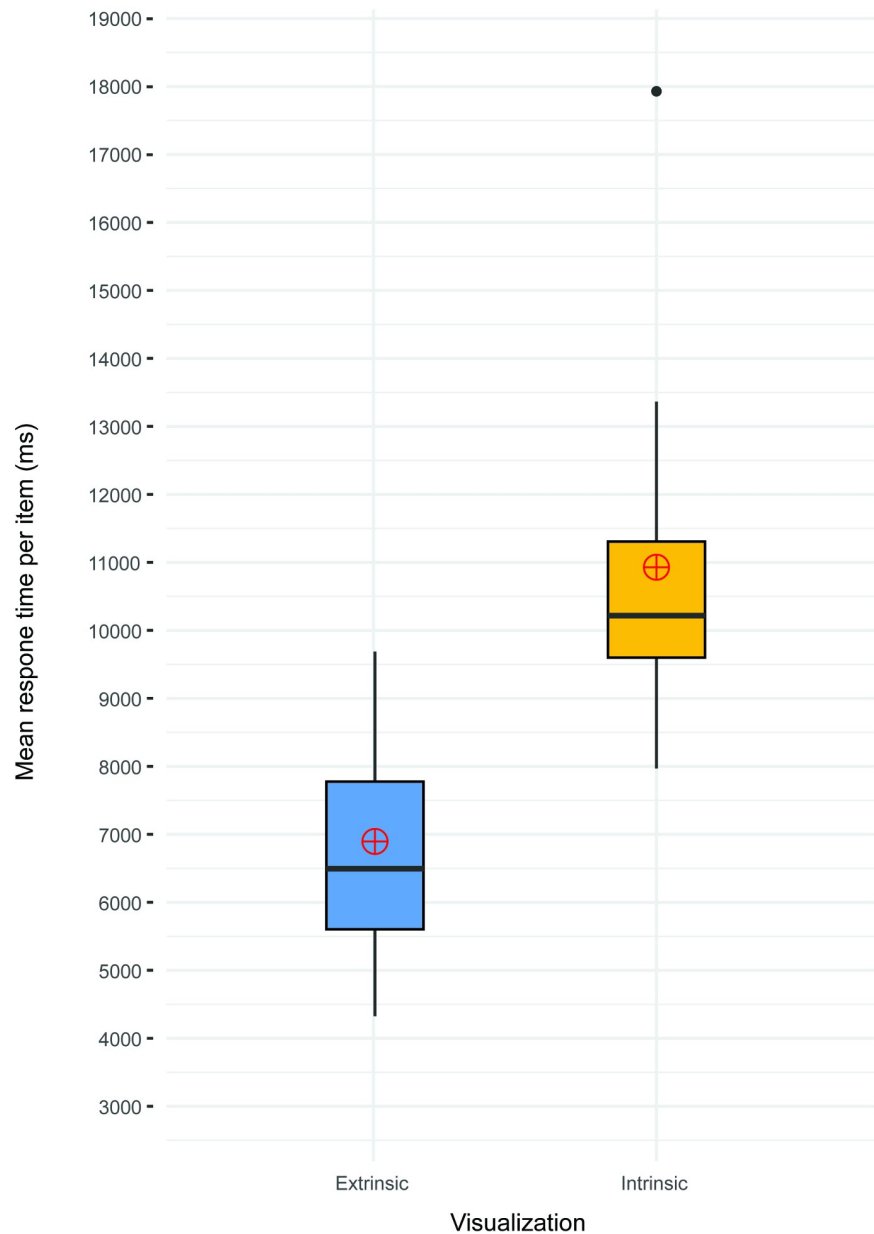
Fig 8. Mean response time per extrinsic/intrinsic visualizations (calculated from the response times to all extrinsic/intrinsic items for all participants).

times than B2 ($Z = 115$, $p < 0.001$, $r = 0.806$ [95% CI: 0.589, 0.883]). In a comparison of the effect sizes for both the extrinsic and intrinsic visualizations, we can see that the effect sizes of the differences between one and two variables were greater in the extrinsic group. It can therefore be assumed that extrinsic visualization is more efficient when a single variable is applied, while intrinsic visualization is more suitable for two variables.

In addition to the above, we performed a response time comparison at the item level. For most items, extrinsic visualization resulted in shorter response times than intrinsic visualization. The opposite was true for only four items, intrinsic visualization only inducing slightly

Table 1. Response times for the individual subtest parts (ms).

| part	Extrinsic				Intrinsic				Wilcoxon rank-sum test		
	mean [95% CI]	sd	median [95% CI]	iqr	mean [95% CI]	sd	median [95% CI]	iqr	Z	p-value	Effect size r [95% CI]
A1	6622 [5571, 7673]	1972	6406 [5094, 7286]	2007	13634 [12009, 15260]	2934	14917 [10800, 16056]	3436	5	$p < 0.001$	-0.779 [-0.839, -0.650] large
A2	8609 [7272, 9946]	2509	7827 [6749, 10230]	3066	10102 [8944, 11259]	2090	10184 [7912, 11915]	3237	71	0.093	-0.305 [-0.607, 0.043] moderate
B1	4945 [4266, 5623]	1273	4762 [3992, 6041]	1781	10798 [8928, 12668]	3377	9825 [8275, 14162]	4416	3	$p < 0.001$	0.831 [-0.853, -0.735] large
B2	6666 [5861, 7470]	1509	6638 [5132, 7708]	2159	8192 [6996, 9388]	2159	7738 [6542, 9604]	2359	68	0.041	0.369 [-0.643, -0.036] moderate
whole test	6896 [6035, 7756]	1615	6494 [2172, 5536]	2172	10929 [9576, 12281]	2442	10217 [9588, 11597]	1705	12	$p < 0.001$	-0.767 [-0.844, -0.607] large

https://doi.org/10.1371/journal.pone.0250164.t001
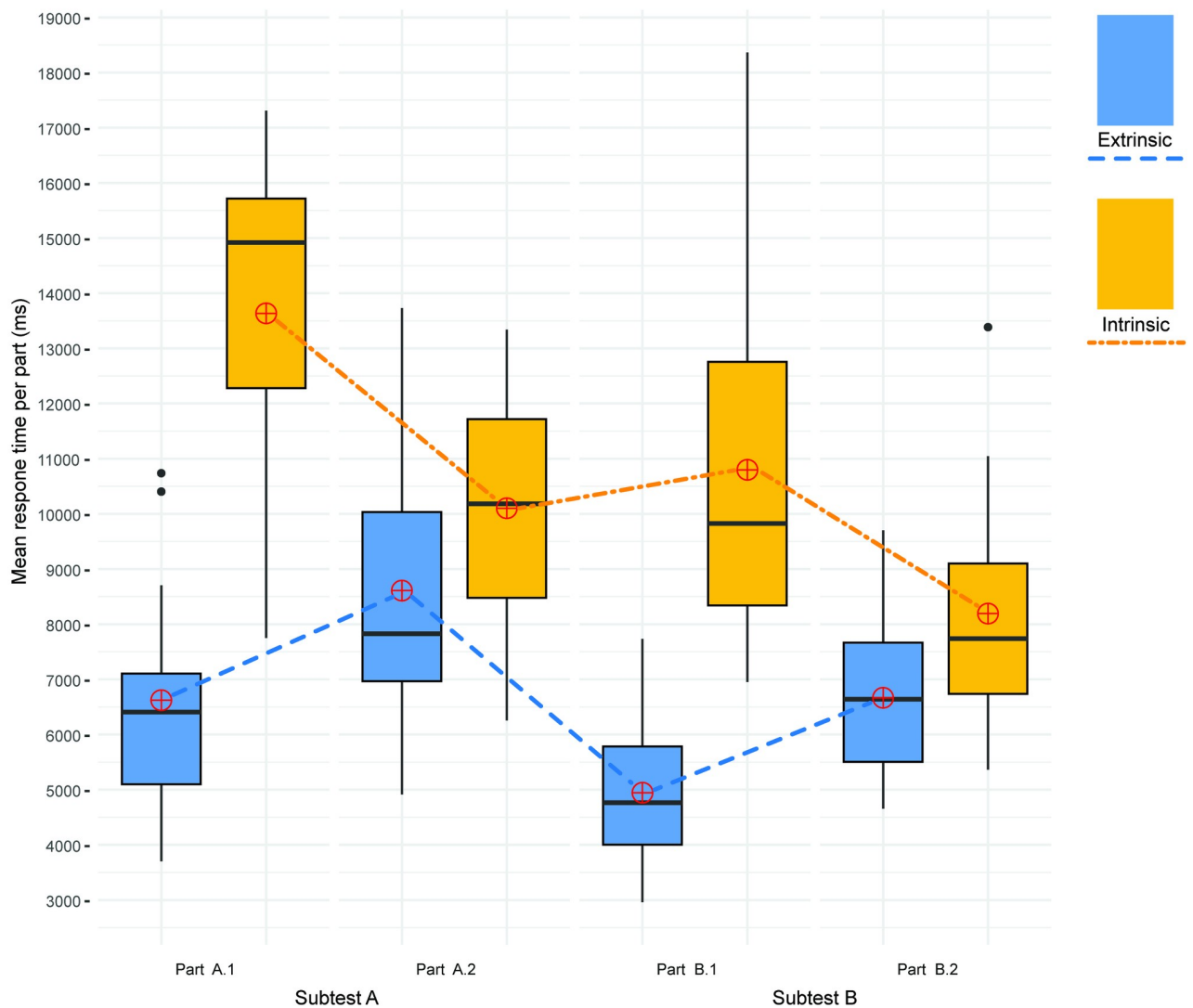
Fig 9. Mean response time (ms) per item (calculated for the individual subtest levels).

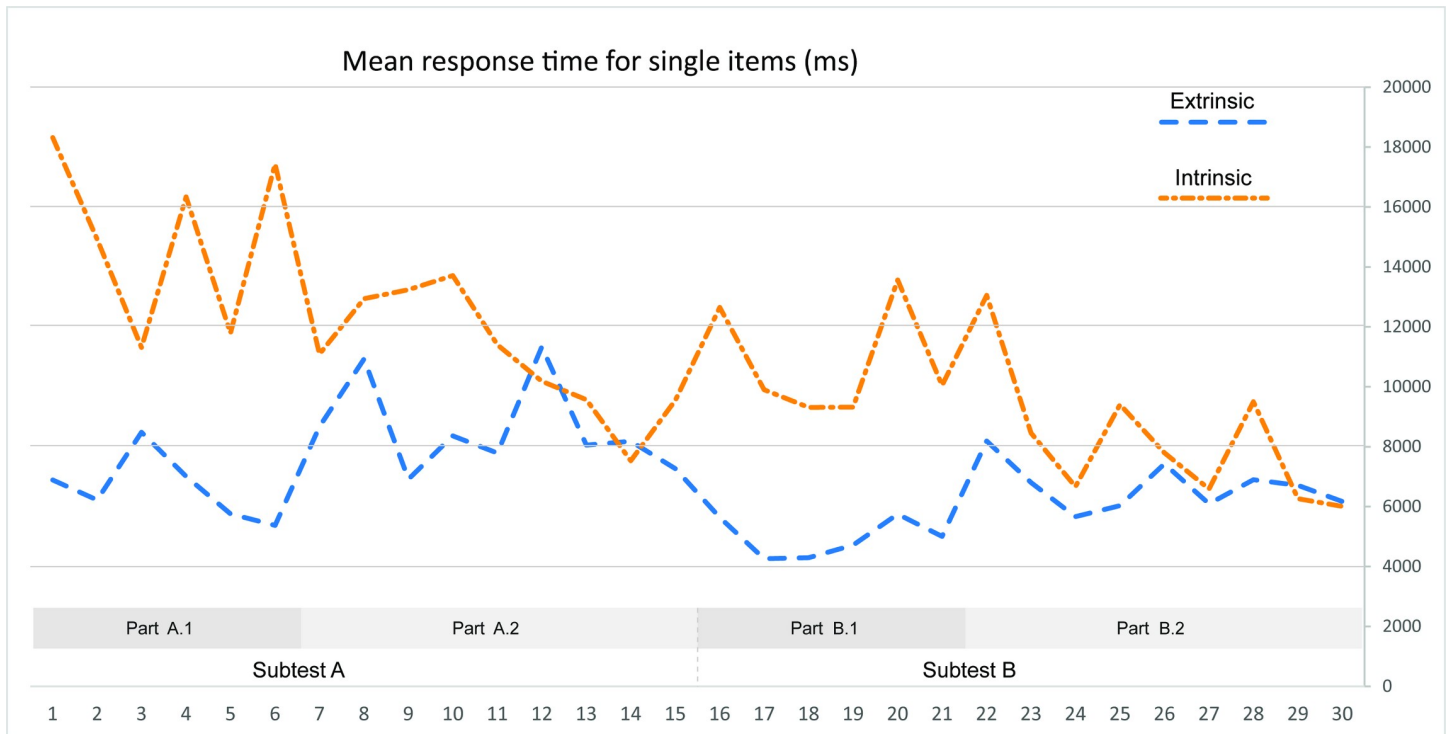https://doi.org/10.1371/journal.pone.0250164.g009

Fig 10. Mean response times (ms) per individual items for extrinsic and intrinsic visualization (all items).

shorter response times (Fig 10). The differences were significant for most items. An analysis at the item-level also revealed two other interesting phenomena: the first consisted in significant variability across the items observed, even within the individual subtests. This variability reflected the complex nature of maps as research stimuli. The difficulty of a test item depended on an array of interacting factors, including the type of correct answer, the distractors selected and the visualized territory. The second observed phenomenon was that the obtained performance curves associated with both visualization types did not overlap, meaning that the difficulty of the test items varied depending on the type of visualization. In other words, the items that were relatively simple to solve in combination with intrinsic visualization were more difficult with extrinsic visualization, and vice versa.

Eye-tracking analysis

For the purposes of the eye-tracking analysis, the stimuli were divided into three key Areas of Interest (AOI): instructions (the textual component), map legend and map. The analysis consisted in a comparison of the dwell times related to the AOI of the individual items (Fig 10 and Table 2). We were also curious about a comparison of the total dwell times for the extrinsic

Table 2. Summary of AOI dwell times (ms).

AOI	Extrinsic				Intrinsic				Wilcoxon rank-sum test		
	mean [95% *CI*]	sd	median [95% *CI*]	iqr	mean [95% *CI*]	sd	median [95% *CI*]	iqr	Z	*p*-value	Effect size *r* [95% *CI*]
Instructions	2514 [2172, 2856]	539	2570 [2193, 2703]	470	2566 [1807, 3325]	1130	2350 [1579, 4252]	1366	74	0.6505	-0.103 [-0.338, 0.53]
Map Legend	216 [84, 348]	208	180 [62, 246]	127	3740 [2921, 4559]	1219	3675 [2411, 4864]	1841	*0*	*p < 0.001*	*-0.847 [-0.851, -0.749]*
Map	3773 [3048, 4497]	1141	3592 [2683, 5117]	1785	3769 [2895, 4642]	1300	3188 [2771, 5847]	1354	70	0.833	- 0.051 [-0.363, 0.49]
All AOI	6503 [5549, 7457]	1502	5970 [5052, 8326]	2712	10075 [8350, 11800]	2567	9142 [8014, 13853]	2192	*10*	*p < 0.001*	*-0.719 [-0.842, -0.475]*

(N = 12) and intrinsic (N = 11) groups (see S1 File). The results showed significant differences in total dwell times, the extrinsic visualization indicating shorter dwell times with a large effect size. A closer examination revealed that the differences were caused by map legend dwell times. The "extrinsic" group displayed significantly shorter dwell times on the map legend than the intrinsic group, with a large effect size and also with a very large gap between the upper bounds of the confidence intervals of the extrinsic group and the lower bounds of the confidence intervals of the intrinsic group. No significant differences were observed in the dwell times during the instructions.

We also visually inspected the oculomotor data in this study at both the item and subtest levels. The times spent on AOI were converted into percentages. The graphs in Figs 11 and 12 show the ratio of time spent on the instructions, map and map legend. We can observe that at the beginning of the experiment, the participants in the extrinsic group needed approximately 10% of the total time-on-task to decode the map's legend; as their experience increased, the time needed to decode the map legend decreased to as little as zero for some items. The "intrinsic" group, by contrast, initially spent about 40% of the time exploring the map legend, with the percentage decreasing with experience, although it remained relatively high (30%).

A comparison of direct saccades (transitions) between the map legend and the visual field of the map reveal a pattern similar to that described for dwell times. Four AOI were defined (instructions, map, map legend, button bar), and a matrix of transitions between the AOI for each item was generated. Fig 13 displays the ratio of the direct map-to-legend/legend-to-map transitions to the total number of transitions between the AOI. It is clear from the graph that the "extrinsic" group only made use of the map legend at the beginning of the experiment; later, direct saccades occurred less. The "intrinsic" group, by contrast, made use of the legend throughout the tasks, with the number of repeated map-to-legend transitions being higher for the tasks with a single variable (A1 and B1).

Discussion

The results of the present study showed that the intrinsic visualization employed was significantly less effective and efficient than extrinsic visualization. In the case of intrinsic visualization, the participants needed significantly more time to solve the tasks and simultaneously produced more errors.

Nevertheless, the response time differences between the two visualization methods were less pronounced when two variables were considered (soil moisture and soil depth). This levelling was caused by the increase in the time needed to solve the tasks with two variables in both extrinsic subtests (A and B). The effect was not observed with the intrinsic visualization. The above finding is in accordance with the studies performed by Nelson [39] and Elmer [43]. We emphasize that the findings and differences between the visualizations can be generalized only with regard to the population on which the research was conducted. It is a lay population with a basic level of map skills and who may also achieve higher education in humanities and social sciences. Conversely, as the results of the study [6] suggest, a population with a higher level of map literacy may prefer the intrinsic method in certain tasks. Another potentially significant change which affects how we work with maps is the type of formal education or the cultural background of users [71–73].

An exploratory analysis of eye-tracking data provided a deeper insight into the above results. Dwell time analysis showed that both groups spent comparable time on the instructions and the map; the reason for longer response times of the "intrinsic" group consisted in the time needed to decode the map legend. While the "extrinsic" group took only a fraction of the total dwell time to interpret the map legend, in the case of the "intrinsic" group, it was over
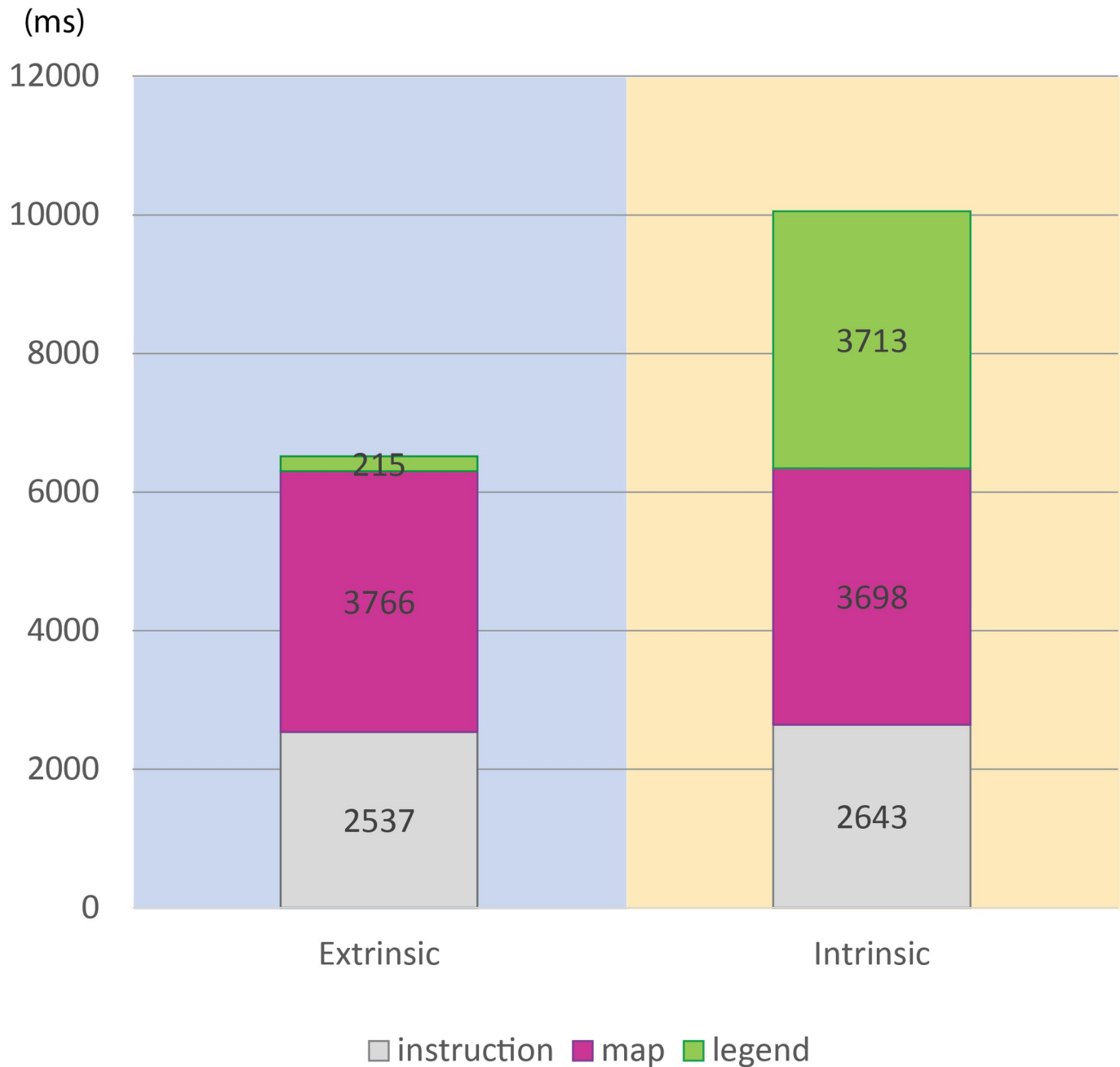
Fig 11. Mean AOI dwell time per extrinsic/intrinsic group (ms).

a third of the total time-on-task. An analysis at the item level revealed yet another tendency: at the beginning of the experiment, the participants in the extrinsic group needed approximately 10% of the total time-on-task to interpret the map legend; as their experience increased, this time decreased to as little as zero for some items. The "intrinsic" group initially spent about 40% of the time decoding the map legend, and although this percentage decreased with experience, it remained as high as 30%. The above results appear to indicate that the map legend of an intrinsic visualization is so complex and essential that it needs to be referred to throughout
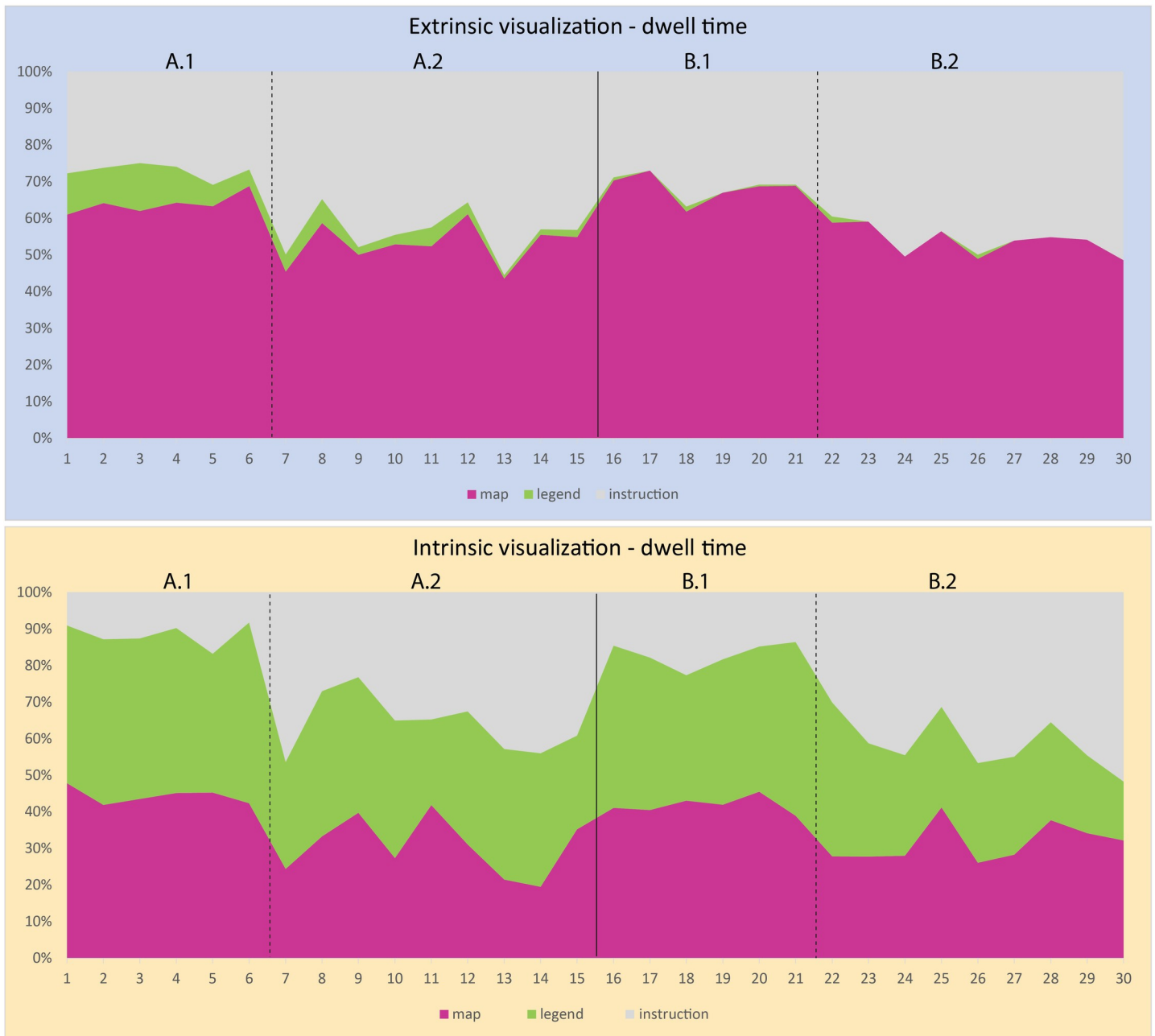
Fig 12. Dwell time on AOI (%) for the extrinsic visualization (top) and intrinsic visualization (bottom). Extrinsic visualization–proportion of dwell time at AOI in single items; (top); Intrinsic visualization–proportion of dwell time at AOI in single items (bottom).

the task. The same conclusion could be drawn from an analysis of direct saccades between the defined AOI.

A comparison of the performance of the "extrinsic" and "intrinsic" verified the greater effectiveness and efficiency of extrinsic visualization. The results also showed that the type of task (i.e., whether it concerned a single variable or two variables) had a definitive effect on performance, which is in accordance with the statement [e.g., 7, 6] that the performance of map work partly depends on whether the given type of visualization is suitable for the task at hand.
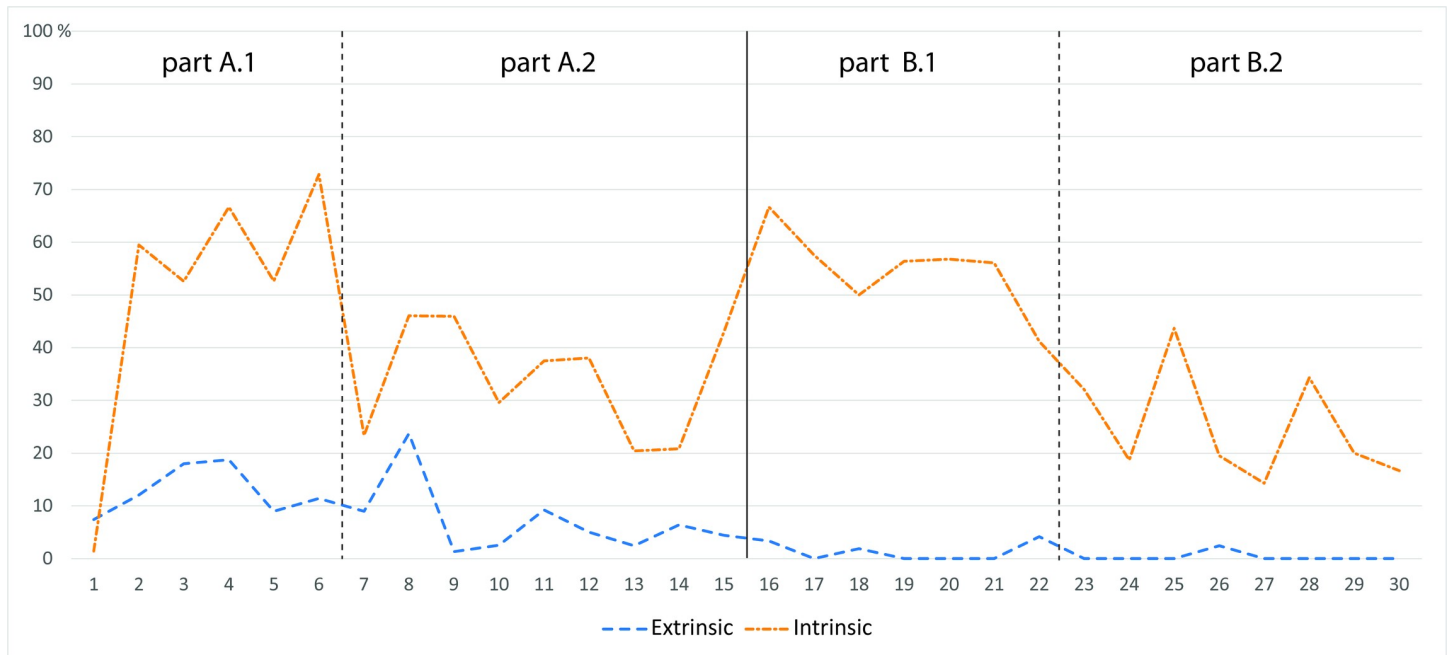
Fig 13. Ratio of the direct map-to-legend/legend-to-map saccades to the total number of direct saccades (%) between the defined AOI (instructions, legend, map, button bar).

If we want to understand the effects of different forms of visualizations during the process of cartographic communication, we must first understand the underlying cognitive processes [17–20]. A particular type of task may require the activation of specific cognitive processes which are appropriate to a particular visualization type. Anderson [74] emphasized that visual representations differ not only in the coding system, but, importantly, in the cognitive processes they evoke.

The results of the present study indicate that in the case of extrinsic visualization, the map user first perceives and processes both visually distinct variables consecutively, subsequently "putting them together" in their working memory when solving the task. Intrinsic visualization, by contrast, requires only one variable to be kept in working memory at any moment when the task concerns two variables (soil moisture and depth). When the task involves a single variable, however, the user must first decode the map legend and keep all three levels of the variable in their working memory. In the above, the results confirm our assumption that the cartographic visualization must be selected according to the type of task or operation to be performed with the particular map.

Our study is not without limitations. One of the limitations was the small sample size and resulting low power in the statistical tests. Low power may lead to an increase of the risk that the existing differences in performance will be falsely not detected as statistically significant. However, we took several (mostly statistical) precautions to prevent the misinterpretation of our data. We conducted a post-hoc sensitivity analysis which suggested that with a given sample size, medium to large effect sizes could be acceptably interpreted (see the first section of the Results chapter), whereas results with small effect sizes would be inconclusive. Rigorous statistical procedures which allow the interpretation of results on smaller sample sizes were also employed in the study (including bootstrapped confidence intervals for means, medians and effect sizes). Regarding the research sample's composition, we attempted to form a sample which was as homogenous as possible (age, level of education, field of study, experience with

maps, etc.) and randomly added participants to the extrinsic/intrinsic groups to obtain an equally balanced sample size for each experimental condition (block randomization) and to reduce potentially confounding effects.

Furthermore, the sample size in our study does not deviate from the standard practice of the field of research in question. King [52] pointed out that many studies work with the relatively small samples given by the high requirements for laboratory equipment. Cognitive cartography surely is one of the fields in which certain studies have contributed significantly to increasing knowledge, regardless of their sample sizes [75–78].

However, the size of the research sample and its composition (European university students of humanities and social sciences with common map literacy skills) permitted us to generalize the conclusions for similar populations. Further research with this method on different samples is required to expand the results of the present study and explore how different population characteristics (map literacy, level and type of formal education) affect the preference for specific types of visualization.

Conclusion

The performed confirmatory analysis verified the superiority of extrinsic visualization in the case of a population of individuals with higher formal education in humanities and social sciences, both in terms of effectiveness and efficiency. Complementary exploratory analysis of eye tracking data suggests that the reason is the character of the intrinsic map legend, which demands greater cognitive resources from map readers during processing. The present study's findings are significant not only for basic research in visualization and cognitive processes but also in their implications for cartographic practices. Even despite a relatively small sample size, the results were statistically significant, but also, importantly, very large effects were discovered. The extrinsic method can be considered convincingly proved as a more suitable visualization type for the given types of task and the lay population.

The results of the present study also raise the question of whether a higher level of efficiency and effectiveness would be maintained with the extrinsic method even if the target population was composed of individuals with high levels of map literacy and different formal education, and whether cultural background plays a role. The specific character of a cartographic visualization sets a significant limit on empirical testing. Even a relatively minor change in partial parameters (e.g., the absolute size of the circles in the case of the extrinsic method, or using a different color scheme in case of the intrinsic method) may affect, for example, the processing speed of visual search or the memorability of the legend, and consequently result in a difference in overall performance. Therefore, to maintain research rationality, it seems reasonable to conduct more partial studies with relatively smaller samples while comparing a wider variability in the applied visualizations and their partial modifications. That is also our objective for future research: if the trends we uncovered are confirmed in studies which examine modified legends, the findings may then be generalized to a wider population, and the principle of the varying methods which were applied can be verified as the cause of the differences in processing. Changes in visualization parameters may also explain the revealed differences at the level of individual items.

Supporting information

S1 File.
(DOCX)

S1 Fig.
(TIF)

Author Contributions

Conceptualization: Čeněk Šašinka, Zdeněk Stachoň.

Data curation: Zdeněk Stachoň, Jiří Čeněk, Alžběta Šašinková.

Formal analysis: Čeněk Šašinka, Zdeněk Stachoň, Jiří Čeněk, David Lacko.

Funding acquisition: Čeněk Šašinka.

Methodology: Čeněk Šašinka.

Resources: Čeněk Šašinka, Jiří Čeněk, Alžběta Šašinková.

Software: Pavel Ugwitz.

Supervision: Čeněk Šašinka, Zdeněk Stachoň.

Validation: Jiří Čeněk, Alžběta Šašinková, Stanislav Popelka, David Lacko.

Visualization: Zdeněk Stachoň.

Writing – original draft: Čeněk Šašinka, Zdeněk Stachoň, Alžběta Šašinková, David Lacko.

Writing – review & editing: Čeněk Šašinka, Zdeněk Stachoň, Jiří Čeněk, Stanislav Popelka, Pavel Ugwitz, David Lacko.

References

1. Montello DR. Cognitive map-design research in the twentieth century: theoretical and empirical approaches. Cartogr Geogr Inf Sci. 2002; 29: 283–304. https://doi.org/10.1559/152304002782008503

2. Montello DR., Freundschuh S. Cognition of geographic information. In Usery LY, McMaster RB, editors. A research agenda for geographic information science, Florida: CRC Press; 2004. pp. 61–91. PMID: 15531985

3. Kolacny A. Cartographic information: A fundamental concept and term in modern cartography. Cartogr J. 1969; 6: 47–49. https://doi.org/10.1179/caj.1969.6.1.47

4. Staněk K, Friedmannová L, Kubíček P, Konečný M. Selected issues of cartographic communication optimization for emergency centers. Int J Digit Earth. 2010; 3: 316–339. https://doi.org/10.1080/17538947.2010.484511

5. Konečný M, Kubíček P., Stachoň Z., Šašinka Č. The usability of selected base maps for crises management: users' perspectives. Appl. Geomat. 2011; 3: 189–198. https://doi.org/10.1007/s12518-011-0053-1

6. Šašinka Č, Stachoň Z, Kubíček P, Tamm S, Matas A, Kukaňová M. The Impact of Global/Local Bias on Task-Solving in Map-Related Tasks Employing Extrinsic and Intrinsic Visualization of Risk Uncertainty Maps. Cartogr J. 2019; 56: 175–191. https://doi.org/10.1080/00087041.2017.1414018

7. Lokka I, Çöltekin A. Simulating Navigation with Virtual 3D Geovisualizations—A focus on memory related factors. Int Arch Photogramm Remote Sens Spatial Inf Sci. 2016; XLI-B2: 671–673. https://doi.org/10.5194/isprs-archives-XLI-B2-671-2016

8. 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. https://doi.org/10.1080/13658816.2011.642801

9. 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. https://doi.org/10.1080/15230406.2013.860255

10. Roth RE. Cartographic Interaction Primitives: Framework and Synthesis. Cartogr J. 2012; 49: 376–395. https://doi.org/10.1179/1743277412Y.0000000019

11. Rautenbach V, Coetzee S, Çöltekin C. Development and evaluation of a specialized task taxonomy for spatial planning–A map literacy experiment with topographic maps. ISPRS J Photogramm Remote Sens. 2017; 127: 16–26. https://doi.org/10.1016/j.isprsjprs.2016.06.013

12. Innes L. Maths for map users. Proceedings of the 21st International Cartographic Conference. 2003; pp. 727–738. Available: https://icaci.org/files/documents/ICC_proceedings/ICC2003/Papers/096.pdf.

13. Rinner C, Ferber S. The effects of map reading expertise and map type on eye movements in map comparison tasks. Abstract and poster presentation at the Conference on Spatial Information Theory. 2005; pp. 14–18. Available: https://digital.library.ryerson.ca/islandora/object/RULA%3A79.

14. 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: 134–15. https://doi.org/10.1080/15230406.2015.1021713

15. Koç H, Demir S. Developing valid and reliable map literacy scale. Rev Int Geograph Educ Online. 2014; 4: 120–137.

16. Clarke D. Are you functionally map literate? Proceedings of the 21st International Cartographic Conference. 2003; pp. 10–16. Available: http://lazarus.elte.hu/cet/publications/088.pdf.

17. Lobben AK. Tasks, Strategies, and Cognitive Processes Associated with Navigational Map Reading: A Review Perspective. Prof Geogr. 2014; 56: 270–281. https://doi.org/10.1111/j.0033-0124.2004.05602010.x

18. Fabrikant SI, Lobben AK. Introduction: Cognitive Issues in Geographic Information Visualization. Cartographica The International Journal for Geographic Information and Geovisualization. 2009; 44:139–143. https://doi.org/10.3138/carto.44.3.139

19. Slocum TA, Blok C, Jiang B, Koussoulakou A, Montello DR, Fuhrmann S, Hedley NR. Cognitive and usability issues in geovisualization. Cartogr Geogr Inf Sci. 2001; 28, 61–75. https://doi.org/10.1559/152304001782173998

20. MacEachren AM, Kraak M-J. Research Challenges in Geovisualization. Cartogr Geogr Inf Sci. 2001; 28: 3–12. https://doi.org/10.1559/152304001782173970

21. Olson JM. Spectrally Encoded Two-Variable Maps. Ann Am Assoc Geogr. 1981; 71: 259–276. https://doi.org/10.1111/j.1467-8306.1981.tb01352.x

22. Larkin JH, Simon HA. Why a diagram is (sometimes) worth ten thousand words. Cogn Sci. 1987; 11: 65–99. https://doi.org/10.1016/S0364-0213(87)80026-5

23. MacEachren AM. How Maps Work: Representation, Visualization, and Design. 1st ed. New York: Guilford Press; 1995.

24. Robinson AH, Morrison JL, Muehrcke PC, Kimerling AJ, Guptill SC. Elements of Cartography. 6th ed. New York: John Wiley & Sons; 2009.

25. Slocum TA, McMaster RB, Kessler FC, Howard HH. Thematic cartography and geovisualization. 3rd ed. New jersey: Pearson Prentice Hall; 2008.

26. Gershon N. Visualization of an imperfect world. IEEE Comput Graph. 1998; 18: 43–45. https://doi.org/10.1109/38.689662

27. D'Zmura M. Color in visual search. Vision Res. 1991; 31: 951–966. https://doi.org/10.1016/0042-6989(91)90203-h PMID: 1858326

28. Lindsey DT, Brown AM, Reijnen E, Rich AN, Kuzmova YI, Wolfe JM. Color channels, not color appearance or color categories, guide visual search for desaturated color targets. Psychol Sci. 2010; 21: 1208–1214. https://doi.org/10.1177/0956797610379861 PMID: 20713637

29. Itti L, Kof CH. A saliency-based search mechanism for overt and covert shifts of visual attention. Vision Res. 2000; 40: 1489–1506. https://doi.org/10.1016/s0042-6989(99)00163-7 PMID: 10788654

30. Bertin J. Sémiologie graphique. 2nd ed. Paris: Gauthier-Villars; 1973.

31. Lavie N., Tsal Y. Perceptual load as a major determinant of the locus of selection in visual attention. Percept Psychophys. 1994; 56: 183–197. https://doi.org/10.3758/bf03213897 PMID: 7971119

32. Granholm E, Asarnow RF, Sarkin AJ, Dykes KL. Pupillary responses index cognitive resource limitations. Psychophysiology. 1996; 33: 457–461. https://doi.org/10.1111/j.1469-8986.1996.tb01071.x PMID: 8753946

33. Brünken R, Steinbacher S, Plass JL, Leutner D. Assessment of cognitive load in multimedia learning using dual-task methodology. Exp Psychol. 2002; 49: 109–119. https://doi.org/10.1027//1618-3169.49.2.109 PMID: 12053529

34. Nelson ES. Using Selective Attention Theory to Design Bivariate Point Symbols. Cartogr Perspect. 1999; 32: 6–28. https://doi.org/10.14714/CP32.625

35. Treisman AM, Gelade G. A feature-integration theory of attention. Cogn Psychol. 1980; 12: 97–136. https://doi.org/10.1016/0010-0285(80)90005-5 PMID: 7351125

36. Wolfe JM, Cave KR, Franzel SL. Guided search: an alternative to the feature integration model for visual search. J Exp Psychol Hum Percept Perform. 1989; 15: 419–433. https://doi.org/10.1037//0096-1523.15.3.419 PMID: 2527952

37. MacEachren AM. Visualizing Uncertain Information. Cartogr Perspect. 1992; 13: 10–19. https://doi.org/10.14714/CP13.1000

38. Nelson ES. Designing Effective Bivariate Symbols: The Influence of Perceptual Grouping Processes. Cartogr Geogr Inf Sci. 2000; 27: 261–278. https://doi.org/10.1559/152304000783547786

39. Nelson ES. The Impact of Bivariate Symbol Design on Task Performance in a Map Setting. Cartographica. 2002; 37: 61–78. https://doi.org/10.3138/V743-K505-5510-66Q5

40. Kubíček P, Šašinka Č, Stachoň Z, Štěrba Z, Apeltaur J, Urbánek T. Cartographic Design and Usability of Visual Variables for Linear Features. Cartogr J. 2017; 54: 91–102. https://doi.org/10.1080/00087041.2016.1168141

41. Kubíček P, Šašinka Č, Stachoň Z. Uncertainty Visualization Testing. Proceedings of the 4th conference on Cartography and GIS. 2012; pp. 247–256. Available: https://www.researchgate.net/publication/292155643_UNCERTAINTY_VISUALIZATION_TESTING.

42. Brus J, Kučera M, Popelka S. Intuitiveness of geospatial uncertainty visualizations: a user study on point symbols. Geografie. 2019; 124:163–85. https://doi.org/10.37040/geografie2019124020163

43. Elmer M. Symbol Considerations for Bivariate Thematic Maps. M.Sc. Thesis, University of Wisconsin–Madison. 2012. Available from: http://resources.maphugger.com/melmer_webedition.pdf.

44. Kunz M, Grêt-Regamey A, Jurni L. Visualizing natural hazard data and uncertainties—Customization through a web-based cartographic information system. Int arch photogramm remote sens spat inf sci. 2010; 38: 1–7. Available from: https://www.isprs.org/proceedings/XXXVIII/part4/files/Kunz.pdf.

45. Lakoff G., Johnson M. The metaphorical structure of the human conceptual system. Cognitive Sci. 2018; 4: 195–208.

46. Harold J., Lorenzoni I., Shipley T. F., Coventry K. R. Cognitive and psychological science insights to improve climate change data visualization. In Nature Climate Change. 2016; 6(12): 1080–1089. https://doi.org/10.1038/nclimate3162

47. Tukey JW. Exploratory data analysis. Reading: Addison-Wesley; 1977.

48. Behrens JT. Principles and procedures of exploratory data analysis. Psychol Methods. 1997; 2: 131–160. https://doi.org/10.1037/1082-989X.2.2.131

49. Madan A, Kumar S. (2012). Usability evaluation methods: a literature review. Int. J. Eng. Sci. Technol. 2012; 4: 590–599.

50. ISO 9241–11:2018 Ergonomics of human-system interaction—Part 11: Usability: Definitions and concepts. 2018 [cited 19 November 2020]. In: ISO Online Browsing Platform [Internet]. Available: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en.

51. Quesenbery W. The Five Dimensions of Usability. In Albers MJ, Mazur MB, editors. Content and Complexity: Information Design in Technical Communication. New York: Routlege; 2013. pp. 81–102.

52. King AJ, Bol N, Cummins RG, John KK. Improving Visual Behavior Research in Communication Science: An Overview, Review, and Reporting Recommendations for Using Eye-Tracking Methods. Commun Methods Meas. 2019; 13: 149–177. https://doi.org/10.1080/19312458.2018.1558194

53. Popelka S, Stachoň Z, Šašinka Č, Doležalová J. EyeTribe Tracker Data Accuracy Evaluation and Its Interconnection with Hypothesis Software for Cartographic Purposes. Comput Intell Neurosci. 2016; 9172506. https://doi.org/10.1155/2016/9172506 PMID: 27087805

54. Brychtova A, Coltekin A. An Empirical User Study for Measuring the Influence of Colour Distance and Font Size in Map Reading Using Eye Tracking. Cartogr J. 2016: 202–212. https://doi.org/10.1179/1743277414Y.0000000103

55. Keil J, Edler D, Kuchinke L, Dickmann F. Effects of visual map complexity on the attentional processing of landmarks. PLoS One. 2020; 15: e0229575. https://doi.org/10.1371/journal.pone.0229575 PMID: 32119712

56. Krassanakis V, Cybulski P. A review on eye movement analysis in map reading process: the status of the last decade. Geodesy Cartogr. 2019; 68: 191–209. https://doi.org/10.24425/gac.2019.126088

57. Brewer CA, Hatchard GW, Harrower MA. ColorBrewer in Print: A Catalog of Color Schemes for Maps. Cartogr Geogr Inf Sci. 2003; 30: 5–32. https://doi.org/10.1559/152304003100010929

58. OpenStreetMap contributors. Planet dump [Data file from 20180404]. 2018. Available: https://planet.openstreetmap.org.

59. Šašinka Č, Morong K, Stachoň Z. The Hypothesis Platform: An Online Tool for Experimental Research into Work with Maps and Behavior in Electronic Environments. ISPRS International Journal of Geo-Information. 2017; 6(12):407. https://doi.org/10.3390/ijgi6120407

60. Tomczak M, Tomczak E. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. Trends Sport Sci. 2014; 1: 19–25.

61. Mangiafico S. rcompanion: Functions to Support Extension Education Program Evaluation. R package version 2.3.25. 2020. Available: https://cran.r-project.org/package=rcompanion.

62. Sherman R. multicon: Multivariate Constructs. R package version 1.6. 2015. Available: https://cran.r-project.org/package=multicon.

63. Rosenthal R. Parametric measures of effect size. In Cooper H, Hedges LV, editors. The handbook of research synthesis. New York: Russell Sage Foundation; 1994. pp. 231–244.

64. Faul F, Erdfelder E, Buchner A, Lang A.-G. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behav Res Methods. 2009; 41: 1149–1160. https://doi.org/10.3758/BRM.41.4.1149 PMID: 19897823

65. Brewer CA. Review of Colour Terms and Simultaneous Contrast Research for Cartography. Cartographica. 1992; 29: 20–30. https://doi.org/10.3138/80ML-3K54-0204-6172

66. Nothdurf HC. Salience from feature contrast: additivity across dimensions. Vision Res. 2000; 40: 1183–1201. https://doi.org/10.1016/s0042-6989(00)00031-6 PMID: 10788635

67. Choudhury AMR. 5—Unusual visual phenomena and colour blindness. In: Choudhury AMR. Principles of Colour and Appearance Measurement. Object Appearance, Colour Perception and Instrumental Measurement. 1st ed. Cambridge: Woodhead Publishing; 2014. pp. 185–220. https://doi.org/10.1533/9780857099242.185

68. Van Zandt T. Analysis of Response Time Distributions. In Pashler H, Wixted J, editors. Stevens' Handbook of Experimental Psychology. New Jersey: John Wiley & Sons Inc.; 2002. pp. 461–516.

69. Whelan R. Effective analysis of reaction time data. Psychol Rec. 2008; 58: 475–482. https://doi.org/10.1007/BF03395630

70. Ratcliff R. Methods for dealing with reaction time outliers. Psychol Bull. 1993; 114: 510–532. https://doi.org/10.1037/0033-2909.114.3.510 PMID: 8272468

71. Chang K-T, Antes JR. Sex and Cultural Differences in Map Reading. American Cartographer. 1987; 14: 29–42, https://doi.org/10.1559/152304087783875345

72. Lacko D, Šašinka Č, Čeněk J, Stachoň Z, Lu W-L. Cross-Cultural Differences in Cognitive Style, Individualism/Collectivism and Map Reading between Central European and East Asian University Students. Stud Psychol (Bratisl). 2020; 62: 23–43. https://doi.org/10.31577/sp.2020.01.789

73. Stachoň Z, Šašinka Č, Čeněk J, Štěrba Z, Angsuesser S, Fabrikant SI, et al. Cross-cultural differences in figure–ground perception of cartographic stimuli. Cartogr Geogr Inf Sci. 2019; 46: 82–94. https://doi.org/10.1080/15230406.2018.1470575

74. Anderson JR. Representational Types: A Tricode Proposal. Technical Report 82–1. Washington, D.C.: Office of Naval Research, 1982. Available from: https://apps.dtic.mil/dtic/tr/fulltext/u2/a116887.pdf

75. Alaçam Ö, Dalcı M. A Usability Study of WebMaps with Eye Tracking Tool: The Effects of Iconic Representation of Information. Human-Computer Interaction New Trends Lecture Notes in Computer Science. 2009; 12–21. https://doi.org/10.1007/978-3-642-02574-7_2

76. Fuchs S, Spachinger K, Dorner W, Rochman J, Serrhini K. Evaluating cartographic design in flood risk mapping. Environmental Hazards. 2009; 8: 52–70. https://doi.org/10.3763/ehaz.2009.0007

77. Çöltekin A, Brychtová A, Griffin AL, Robinson AC, Imhof M, Pettit C. Perceptual complexity of soil-landscape maps: a user evaluation of color organization in legend designs using eye tracking. International Journal of Digital Earth. 2016; 10: 560–581. https://doi.org/10.1080/17538947.2016.1234007

78. Çöltekin A, Heil B, Garlandini S, Fabrikant SI. Evaluating the Effectiveness of Interactive Map Interface Designs: A Case Study Integrating Usability Metrics with Eye-Movement Analysis. Cartography and Geographic Information Science. 2009; 36: 5–17. https://doi.org/10.1559/152304009787340197