Pacific Graphics (2023)
R. Chaine, Z. Deng, and M. H. Kim (Editors)

Poster

Revisiting Visualization Evaluation Using EEG and Visualization Literacy Assessment Test

Soobin Yim Chanyoung Jung Chanyoung Yoon Sangbong Yoo Seongwon Choi and Yun Jang†

Sejong University

Abstract

Using EEG signals, also known as Electroencephalogram, can provide a quantitative measure of human cognitive load, making it an effective tool for evaluating visualization. However, the suitability of EEG for visualization evaluation has not been verified in previous studies. This paper investigates the feasibility of utilizing EEG data in visualization evaluation by comparing previous experiments. We trained and estimated individual CNN models for each subject using the EEG data. Our study demonstrates that EEG-based visualization evaluation provides a more feasible estimate of the difficulties experienced by subjects during the visualization task compared to previous studies that used accuracy and response time.

CCS Concepts

• ***Human-centered computing*** → *Visualization design and evaluation methods;*

1. Introduction

Electroencephalogram (EEG) signals can offer a quantitative assessment of human cognitive workload, rendering it a valuable tool for assessing visualizations [APM*11, NSJ*20, YYYJ23, PBJ*18]. Nonetheless, prior studies that assessed visualizations using EEG did not juxtapose their mental workload estimations with established research on visualizations. Consequently, it is impossible to affirm the suitability of their proposed approach for evaluating visualizations through EEG.

This paper explores the potential for employing EEG data in the assessment of visualization evaluation through a comparison of previous visualization research experiments. [LKK16]. We verify the feasibility of EEG-based visualization evaluation by comparing the estimations of individual CNN models, item difficulty index, and VLAT [LKK16]. Our study shows that utilizing EEG-based assessment for visualization evaluation offers a more practical means of gauging the challenges encountered by participants during the visualization task, in contrast to prior studies that relied on accuracy and response time as metrics.

2. Visualization Experiment

This experiment was designed as a benchmark for the Visualization Literacy Assessment Test (VLAT) [LKK16]. VLAT was proposed to measure visualization literacy.Visualization tasks in VLAT are composed of varying difficulty levels, making it suitable for measuring cognitive load induced by visualizations of varying complexity.

† Corresponding author, e-mail: jangy@sejong.edu

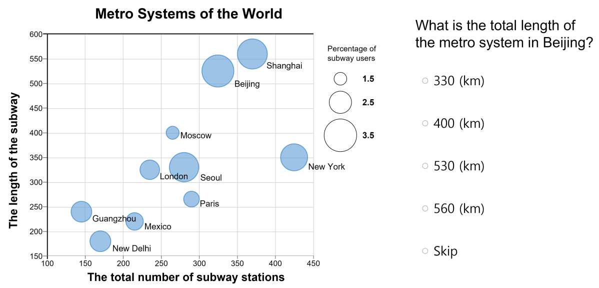

Figure 1: *Visualization in ID 47.*

2.1. Experiment Design

Since the body movement can introduce noise in the EEG data, we did not include any interaction techniques within the visualizations. Six subjects aged between 24 to 35 years were recruited for the study. There were 53 potential test items, including 35 four-option multiple-choice items, 3 three-option multiple-choice items, and 15 true-false items. Figure 1 is an example of the visualization task of our experiment. EEG data were collected from the subjects while they were conducting the tasks.

2.2. EEG Data Collection

We used the Emotiv Epoc FLEX to collect EEG data using saline sensors. This device has 32 electrodes. The EEG data were collected only during task performance, excluding rest periods. The collected data consists of timestamp, raw data by channel, band

DOI: 10.2312/pg.20231289

Table 1: *Comparison of item difficulty index between our study and Lee et al. [LKK16] and its impact on mental workload estimation. The black number indicates the Item ID in VLAT, and the red number indicates the number of items classified as the Item Difficulty Index.*

Item Difficulty Index	VLAT's Item Difficulty Index	The Estimated Mental Workload in the Model				Count
		Low	Low and Mid	Mid	High	
Easy	Easy	1, 4, 17, 21, 23, 25, 42	7, 38, 56	2, 6, 20, 32, 44, 57, 61	-	17
	Moderate	35	8, 22, 27, 33, 34, 59	3, 5, 18, 19, 28, 29, 48, 52, 54	-	16
	Hard	11, 15, 31, 47, 49	10, 16, 37, 41, 55	9, 36, 60	-	13
Moderate	Moderate	14	-	12, 51	-	3
	Hard	45, 53	40, 46	-	-	4
Count		17	15	21	0	53

power data by channel. The collected band power data consists of theta (4∼8 Hz), alpha (8∼12 Hz), low-beta (12∼16 Hz), high-beta (16∼25 Hz), and gamma (25∼45 Hz).

2.3. Item Difficulty Index

The item difficulty index is a metric representing the ratio of subjects who answered an item correctly, and its value ranges from 0 to 1.0 [TCTH91]. The task classification based on the index is as follows: items with $P_i > 0.85$ are considered easy, those with $0.5 < P_i \leq 0.85$ are considered moderate, and those with $P_i \leq 0.5$ are considered hard [LKK16].

3. Model

EEG data are min-max normalized and randomly sampled to prevent overfitting. The individual mental workload estimation CNN model is trained to incorporate cognitive differences that differ from person to person in the experiment. We compare the mental workload estimated by the model. The model achieved a training accuracy ranging from 96.73% to 99.72% and a test accuracy ranging from 88.54% to 91.54%.

4. Finding

Table 1 compares the difficulty of visualizations measured by Lee et al. [LKK16] with those measured in this work. Every number within the table represents an item ID, while the numerical value within the count indicates the number of tasks associated with a specific difficulty level. Based on the item difficulty index, Lee et al. [LKK16] classified problems into 17 hard items, 19 moderate items, and 17 easy items. According to the behavioral data we collected, we classified problems into 0 hard items, 7 moderate items, and 46 easy items. Also, our model classified problems into 17 low items, 15 low and moderate items, and 21 moderate items. This difference is interpreted as the EEG-based model being able to capture complex cognitive mechanisms not seen in the item difficulty index, which is calculated by a simple formula. Therefore, an EEG-based evaluation is valuable for visualization because EEG data contain the mental workload the subjects experience.

5. Conclusions

We verify the validity of EEG utilization through data-driven comparison by benchmarking the experiments of previous visualization studies [LKK16]. We found that, depending on the subjects, there were differences in quantitative metrics, such as the item difficulty index. The EEG-based visualization evaluation classifies the mental workload of the subjects, including all the mental workloads experienced during the tasks, independent of the percentage of correct answers. Therefore, the EEG-based method is suitable for estimating the mental workload in visualization evaluation, which has not been captured in previous studies. In the future, we will investigate whether EEG can be applied to visualization evaluation in various visualizations.

Acknowledgments

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2022-0-00305, Development of automatic data semantic information composition/expression technology based on augmented analysis for diagnosing industrial data status and maximizing improvement, 50%) and (No.2021-0-00469, Development of heterogeneous fusion data detection and tracking technologies, 50%)

References

[APM*11] ANDERSON E. W., POTTER K. C., MATZEN L. E., SHEPHERD J. F., PRESTON G. A., SILVA C. T.: A user study of visualization effectiveness using eeg and cognitive load. *Computer graphics forum 30*, 3 (2011), 791–800. 1

[LKK16] LEE S., KIM S.-H., KWON B. C.: Vlat: Development of a visualization literacy assessment test. *IEEE transactions on visualization and computer graphics 23*, 1 (2016), 551–560. 1, 2

[NSJ*20] NUAMAH J. K., SEONG Y., JIANG S., PARK E., MOUNTJOY D.: Evaluating effectiveness of information visualizations using cognitive fit theory: a neuroergonomics approach. *Applied Ergonomics 88* (2020), 103173. 1

[PBJ*18] PAREKH V., BILALPUR M., JAWAHAR C., KUMAR S., WINKLER S., SUBRAMANIAN R.: Investigating the generalizability of eeg-based cognitive load estimation across visualizations. In *Proceedings of the 20th International Conference on Multimodal Interaction: Adjunct* (2018), pp. 1–5. 1

[TCTH91] THORNDIKE R. M., CUNNINGHAM G. K., THORNDIKE R. L., HAGEN E. P.: *Measurement and evaluation in psychology and education*. Macmillan Publishing Co, Inc, 1991. 2

[YYYJ23] YIM S., YOON C., YOO S., JANG Y.: A mental workload estimation model for visualization using eeg. In *The proceedings of the 56th Hawaii International Conference on System Sciences (HICSS)* (2023). 1