Evaluating the Impact of a Constructivist Approach to Treemap Literacy

E.E. Firat[1], **C. Srinivas**[2], **C. Lang**[2], **B. Srinivas**[2], **R. S. Laramee**[3], **A. P. Joshi**[2]
[1]Cukurova University, [2]University of San Francisco, [3]University of Nottingham

***Abstract*—Constructivist learning is based on the principle that learners *construct* knowledge based on their prior knowledge and experiences. We explored the impact of a Constructivist approach to introduce students to the Treemaps visualization technique. We developed software that helps students understand Treemaps using a synchronized, multi-view, interactive Node-Link representation of the same data. While students in both groups - the ones who used the Node-Link diagram with the Treemaps and the ones who used only the interactive Treemaps demonstrated significant improvement in learning, students who only interacted with the Treemaps representation performed better on a variety of tasks related to reading and interpreting Treemaps.**

Treemaps are a popular technique that has emerged from the data visualization research community for representing hierarchical data [1].

Constructivism is a learning theory based on the concept that people actively "construct" their own knowledge, building on their prior knowledge and experiences as a foundation [2]. The background and previous knowledge of an individual impacts their ability to understand a given topic. This approach employs the pedagogical principle that students acquire knowledge more effectively when they actively engage with the subject matter.

Students study network diagrams and matrix-based representations to store and manipulate data at the high school level [3]. Node-link diagrams are popular visual designs and their associated literacy [4] has been studied in the data visualization community as well. Zoss *et al.* [4] defined network visualization literacy (NVL) as the ability to read, interpret, and visualize various types of networks. We built novel software that uses a Constructivist approach by taking advantage of the familiarity that students have with node-link diagrams to introduce them to Treemaps.

We evaluated the impact of the constructivist approach to teaching students about Treemaps with two groups of students in a Data Visualization course. One group of students was introduced to an interactive (constructivist) Node-Link + Treemaps (NLT) interface and another group of students was introduced to an interactive Treemaps-only (T) interface. The research question we were investigating was whether the Constructivist approach (NLT) helps students understand Treemaps concepts better than the *Treemaps-only* interface.

Theoretical Foundations

Constructivism

Constructivism describes the way that a student can make sense of new content, as well as how the content can be taught effectively. With Constructivism as an educational theory, teachers consider what students *know a priori* and enable their students to put their knowledge into practice.

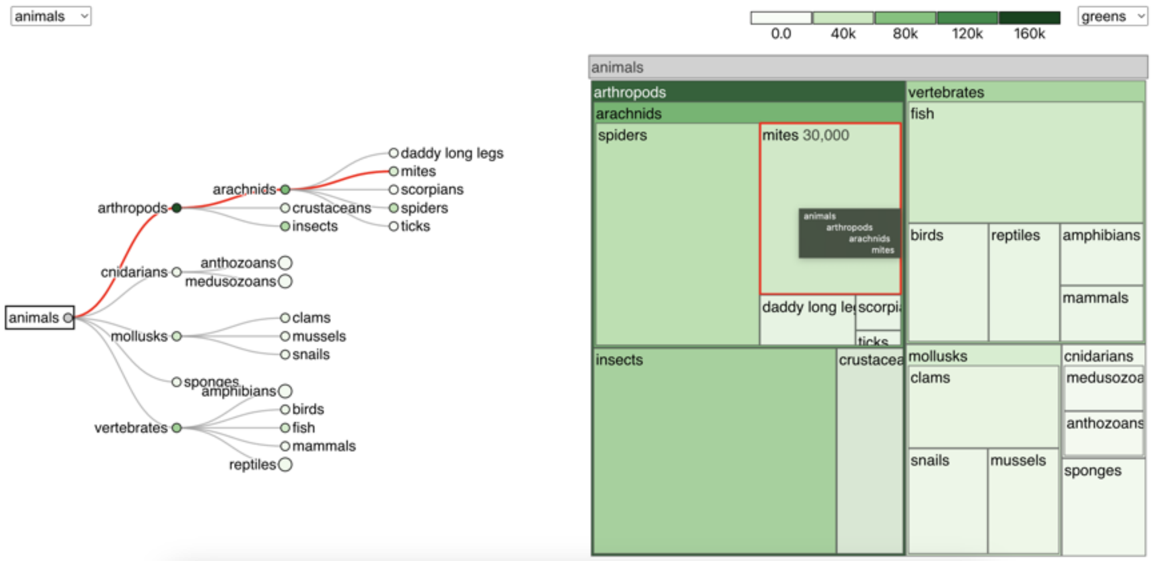

Figure 1: The Constructivist interface of the interactive software for teaching students about Treemaps. The software can be found at https://nodelink-treemap.github.io/.

Terminology

We make a distinction between **Constructivism** as an epistemology and **constructionist** strategies situated within that epistemology, as observed by Papert and Harel [2]. **Constructivism** assumes learners are individuals who derive meaning from the world based on their *existing knowledge* and their personal experiences, **Constructionism** represents the possibility for "facilitating meaning-making" for those individuals through *active learning* and the process of building a shareable result [2].

Related Work

Data visualization literacy (DVL) [5] can be broadly defined as the ability to read, make, and explain visual representations of data.

As we aim to increase the visualization literacy of general audiences, we need software/tests [6] that can be used to evaluate the current level of visualization literacy for an individual. Such community-approved tests can be used by researchers to identify baseline visualization literacy of their students. *Treemaps* are included as one of the 12 techniques in the Visualization Literacy Assessment Test (VLAT) [6].

Ruchikachorn and Mueller [7] propose a learning by analogy method that shows a step by step conversion between two visual designs to introduce a new visualization technique. Four visualization pair examples are used in the study to illustrate the concept. After interacting with the transitions, the students are better able to understand the unfamiliar visual designs.

Treemap evaluation has been a popular topic in the data visualization research field. Firat et al. [8] conducted an intervention that compared the use of slides and that of interactive software to teach students about treemaps. They found that interactive software was better at teaching students about treemaps than slides. In previous work, we evaluated the impact of just one software (constructivist) on student learning [9]. In this paper, we compare the constructivist software (NLT) with a Treemaps-only (T) interface to identify their impact on student learning.

Methodology

We developed novel software and evaluated its benefits to teach students in a Data Visualization course about Treemaps. We divided the students in the class into two groups - Treemaps only (T) group and Node-link + Treemaps (NLT) group. Students either interacted only with Treemaps (T) or interacted with a Node-Link + Treemaps (NLT) version of the software. Using a list of participant IDs, we assigned NLT group IDs to alternating students and T group IDs to the remaining students, ensuring an almost equal distribution between the two groups.

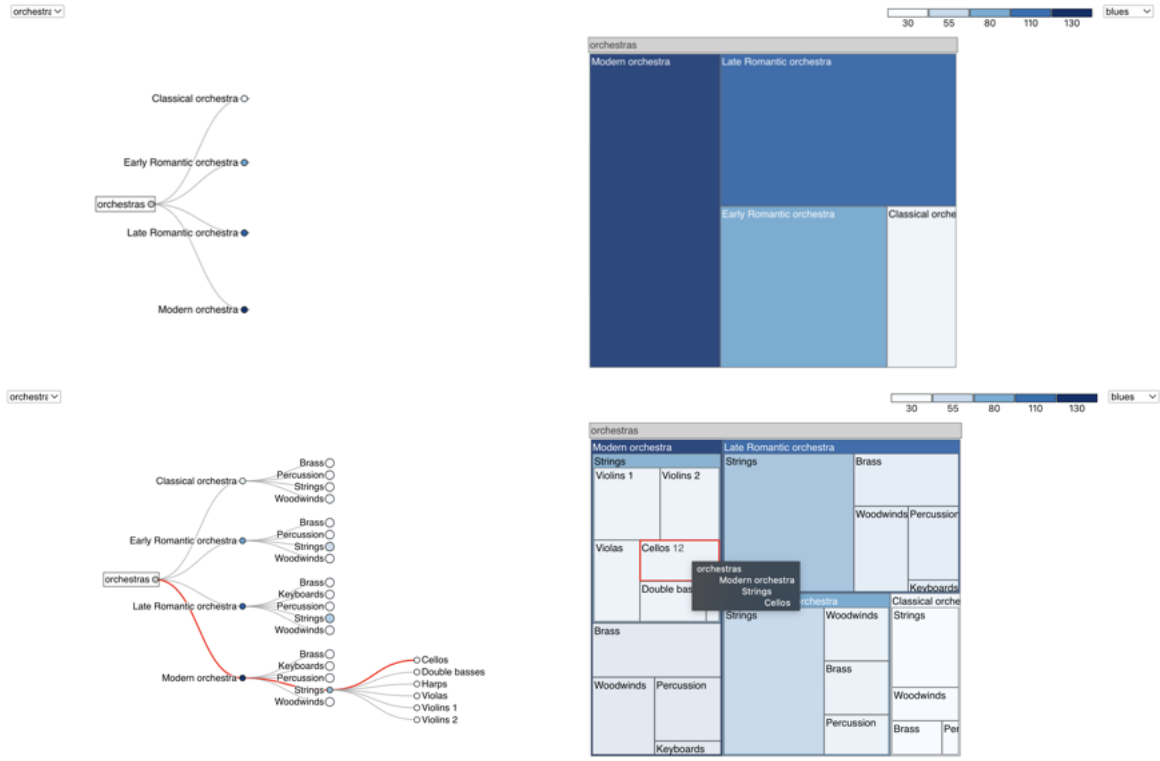

Figure 2: The two screenshots above show a snapshot of the before and after a node has been expanded and the way the color map is used to convey the numeracy of a node. The top figure shows the *Modern Orchestra* rectangle (top left) colored in dark blue, whereas on the bottom the *Modern Orchestra* rectangle is further expanded to reveal its various children.

Interactive Software

To help our students understand Treemaps, we built interactive, online software that allows them to synchronously examine node-link representation of the data along with a treemap representation of the data. The software contains two versions - the constructivist version contains both Node-link diagrams and Treemaps (NLT) and can be found at https://nodelink-treemap.github.io/, whereas the Treemaps-only (T) version can be found at https://nodelink-treemap.github.io/treemap. The source code for the software can be found at https://github.com/nodelink-treemap/nodelink-treemap.github.io. Figure 1 shows the interface of the program, where the path to a specific node in the hierarchy is highlighted in red in both representations (the node-link representation on the left and the treemap on the right). In the treemap view, we also show the viewer a tooltip that shows the hierarchy of the current node (which in this case is $animals \rightarrow arthropods \rightarrow arachnids \rightarrow mites$). A student can interact with the treemap by hovering over a rectangle to examine a node or by clicking on a node to zoom into its sub-tree. Figure 2 shows an example where a node (*Modern Orchestra*) is minimized in the left snapshot and expanded in the right snapshot. As the *Modern Orchestra* sector has a high aggregate value, the *Modern Orchestra* rectangle and corresponding circle in the node-link diagram are colored dark blue. On expanding the *Modern Orchestra* node in the node-link diagram, the child nodes are colored based on their aggregated/individual values (as shown in the bottom snapshot of Figure 2).

Evaluation

We conducted an intervention to introduce students to Treemaps using our Constructivism-based tool. The procedure for our educational intervention is as follows:

1) Students fill out a demographics survey and complete the color blindness check.
2) Students then fill out a modified mini-VLAT [10] as pre-assessment tool to evalu-

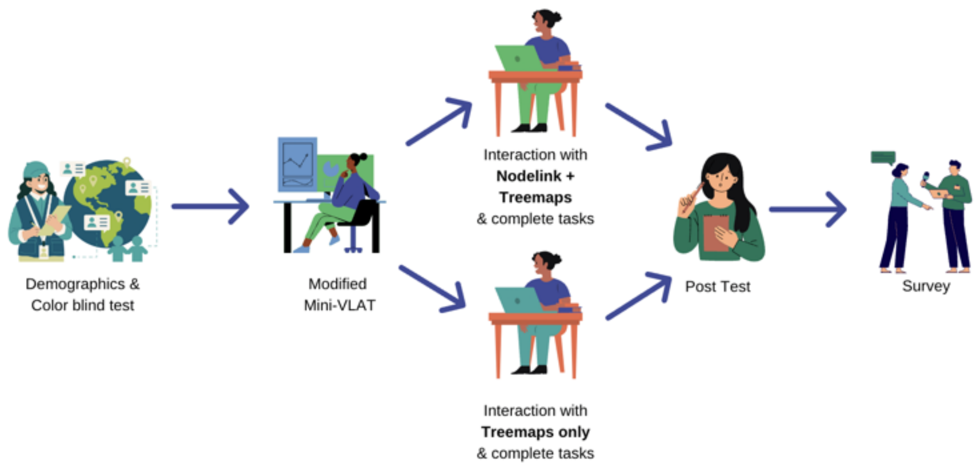

Figure 3: Procedure of the pedagogical intervention. After the demographics survey, we assessed the visualization literacy of all the students using the mini-VLAT. Students were then either asked to explore the data using the Node-Link Treemaps (NLT) interface or the interactive Treemaps (T) interface and answer some questions. This was followed by a post-test and a survey to assess student understanding of Treemaps and their feedback about the software.

ate their baseline visualization literacy. We added three more Treemaps-related questions to the mini-VLAT to assess student familiarity with Treemaps.

3) Students perform common tasks (expand, compare, hover, etc.) as they interact with the Treemaps software and answer specific questions about the data.
4) Students complete a post test (assessment) with 25 questions (to assess their understanding of Treemaps).
5) Student share their experience with the software by filling out a qualitative survey.

In this intervention, there were 69 students (65 undergraduate students, 4 graduate students). Out of those, 66 students were in the 18-24 age group, 3 students were in the 25-44 age group (24 females and 45 males). Two students were excluded from the study due to their lack of attention to the questions, as evidenced by their rapid response times and high error rates. Finally, we had 34 students in the NLT group and 33 students in the T group.

Interaction with Software and Exploratory Tasks

Students were asked to perform specific tasks to get experience with the visualization technique by using our software. Students in both the groups answered 16 questions as they explored the software and learned about Treemaps. The questions have been provided in the supplementary data at https://github.com/vis-graphics/treemaps-literacy/tree/main/surveys. Here is a set of representative tasks that the students performed:

- Locate the path for the species - ‘frogs’ using the software (Find a node)
- Which animal categories have similar populations in that specific category of species? (Compare number of children)
- Based on the color intensity and the size of the nodes in the arthropods family, the population of ’spiders’ is larger than that of ’mites’. (Compare number of children)
- To which category of species do ‘jellyfish’ belong? (Trace the hierarchy of a node)

Post Test

After the exploratory tasks with the software were performed, we assessed the understanding of Treemaps for students in both the groups. The assessment consisted of 25 questions: 10 questions required students to trace the hierarchical relationship of nodes for comparison, and 15 questions required students to compare the areas of rectangles in a Treemap. 6 of the questions were True/False questions, whereas the other 19 were multiple-choice questions. The post-intervention test questions are available in the supplementary material at https://github.com/vis-graphics/treemaps-literacy/blob/main/surveys/PostTest.pdf.

Survey Questions

After the post test, we asked the students the following few questions to solicit feedback about the impact of the software on their understanding of Treemaps and to ask for areas of improvement in the software.

1) Did the Treemaps software help you gain a better understanding of Treemaps? (Yes/No)
2) Why (or why not) do you think it was helpful?
3) Rate the efficacy of Treemaps software to visualize hierarchical data on a scale of 1 (not at all) - 7 (very much).
4) Why (or why not) do you think it was effective?
5) Do you recommend any improvements to the software?
6) What was your understanding of Treemaps **before** you interacted with the Treemaps software? Rank on a scale of 1 (not at all) to 7 (very much).
7) What was your understanding of Treemaps **after** you interacted with the Treemaps software? Rank on a scale of 1 (not at all) to 7 (very much).

Findings

We analyzed the answers of the students in both the groups and found that the majority of the students (76%) had not seen a Treemap before, but 58% of the students had seen a Node-link/Network diagram before. 83% of the students mentioned that they had no prior background in data visualization.

We will first examine the scores of the students in the modified mini-VLAT, then look at the scores on the exploratory tasks of students in both the groups, which is followed by examining the scores of the students on the post-test assessment, and a summary of the students' survey responses.

Mini-VLAT Scores of students in both groups

To understand and evaluate the visualization literacy of the students, we conducted the modified mini-VLAT [10] before the intervention and found that majority of the students in both the groups had a reasonably high score (see Figure 4).

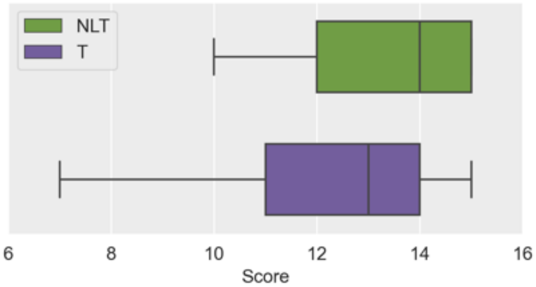

Figure 4: This chart shows the distribution of student scores (out of 16) on the modified Mini-VLAT [10] in the NLT group and the T group. The plot shows that while the students in the NLT group had a slightly higher median score on the mini-VLAT (NLT = 14, T = 13), their visualization literacy is comparable to the students in the T group. After conducting a *Welch's t-test* on the samples, we get a two-tailed p-value of 0.03.

Exploratory Task scores

The literacy assessment test was followed by exploratory interaction with the software for each group, accompanied by a set of 16 questions. The students are required to interact with the software (zoom, hover, drill-down) to answer the questions. Students in both the groups received an average of 15/16 questions correct, implying that they were able to use and interact with the software successfully. The 16 tasks and the analysis can be found in the supplementary material (https://github.com/vis-graphics/treemaps-literacy/)

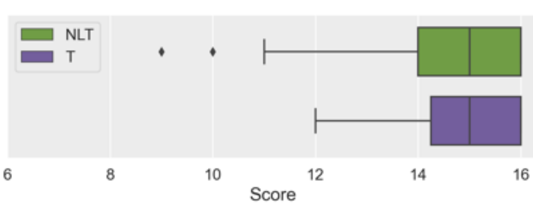

Figure 5: Students' scores on the *Exploratory Tasks* performed in the two groups. Participants in both the groups performed fairly similarly, with a few outliers for tasks following the software exploration phase. We conducted a *Welch's t-test* on the samples to get a two-tailed p-value of 0.30.

Post-Test Accuracy

We analyzed the performance of the students in the post-test, where they were asked to examine and interpret Treemaps. Figure 6 shows the distribution of the scores of students in both the groups. Students in the NLT group answered 81% of the questions accurately, whereas those in the Treemaps group answered 86% of the questions accurately. Even though the median scores in both the groups is the same at (22/25), the spread of the scores of the students in the NLT group is much wider than the spread of the scores in the Treemaps group.

On close examination of the student scores per question, as shown in Figure 7, we can see that the students in the Treemaps group outperformed the students in the Node-Link + Treemap group on majority of the questions. In some cases, such as Q5, Q16, and Q20, the difference is considerable.

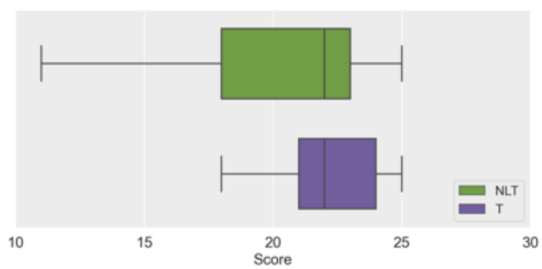

Figure 6: Post-Test score distribution for students in both the groups. Even though the median scores were the same, the distribution of the scores of the students in the T group (purple) skewed higher than that of the students in the NLT group (green). We conducted a *Welch's t-test* on the samples due to their unequal variances and obtained the p-value of 0.29.

Figure 8 shows Q5 where students were asked to compare the areas of two rectangles that were not next to each other. The accuracy of students in both the groups was low for this question, but students in the NLT group performed much worse than those in the T group. The low scores could be attributed to the large number of rectangles and small font size in the Treemap representation.

Survey data analysis

In the survey, students were asked to rate their understanding of Treemaps after having interacted with the software. Based on the self-assessment, we can see from Figure 9 that the students in both the groups had a significant improvement. Students

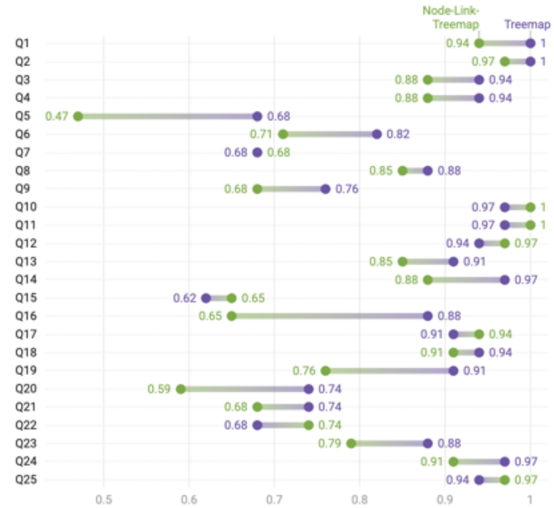

Figure 7: This plot shows the performance of students in both the groups on the post-test assessment.

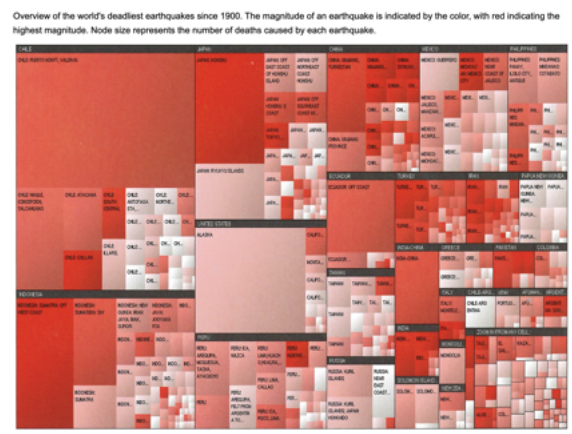

Figure 8: This shows the screenshot of the treemap in Q5 from the Post-Test where the accuracy of students in both the groups was low. The question was "Which two countries have the similar number of deaths?" and the answer options were (i) China-Mexico, (ii) Taiwan-India, (iii) Greece-Iran, (iv) Japan-United States, (v) Not sure.

in the NLT reported a higher overall improvement in their understanding than the students in the Treemaps only (T) group. This is encouraging evidence that the software is effective for teaching students about Treemaps.

We asked students to rate the overall efficacy of the software on a Likert scale (1-7). Figure 10 shows that a large majority of the students in both the groups rated the software as a 6 (very

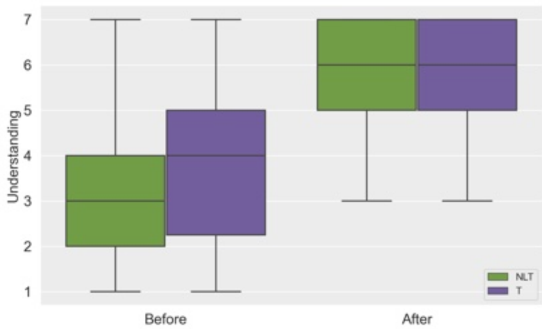

Figure 9: This shows the self-reported change in students' understanding of Treemaps using the software in both the groups. Participants in both groups reported an improvement in understanding.

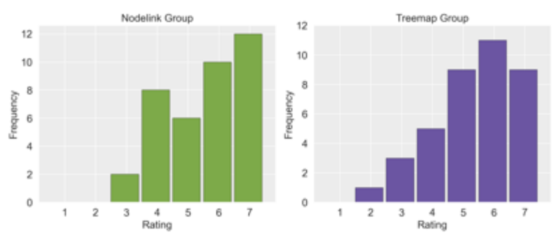

Figure 10: Students were asked to rate the efficacy of our Treemaps software to visualize hierarchical data on a scale of 1 (not at all effective) - 7 (extremely effective). Majority of the students in the both the groups rated it as effective (5), highly effective (6), or extremely effective (7).

effective) or 7 (extremely effective).

Students were asked to comment on the difficulty level, the amount they struggled with answering questions, how helpful the software was, and how effective they found it in terms of teaching them about Treemaps. We performed *sentiment analysis* on each comment using the Hugging Face Transformers Library. It utilizes AutoModel and AutoTokenizer modules to load a pre-trained BERT model. The analysis is based on the model's prediction of sentiment polarity. Majority of the students **in both the groups** said that they had very little to no difficulty in answering the assessment questions, they found that they did not have to struggle to answer the questions, they found the software to be quite helpful and effective in teaching them about Treemaps.

A student commented on the efficacy of the software by saying, "*You could clearly see how the area is divided to fit a square.*" Another student wrote "*... seeing the tree structure that creates the Treemap helped me understand the Treemaps better.*" A student who scored an 84% on the post-test assessment commented on the difficulty level - "*It was hard for me because I have never seen it before. It may be because it is the first time I saw it.*"

One of the students specifically commented on the Constructivist approach - "*I think the software allows me to expand or collapse the nodes and get a better understanding of Treemaps through the node-link map.*" Another student commented on the benefits of the interactive software - "It was helpful because having it be interactive means I can minimize the parts of the data that are distracting or I don't understand."

In terms of improvement to the software, students commented on providing more color scales in the software, bigger labels for the post-test questions, and four students specifically suggested adding a search bar to find nodes. One student wrote "A search bar could make it very easy to find an item that I do not know which category it belongs to. It could show the path of the item as a result."

All of our developed surveys, software (and its source code), data from the evaluation, and additional figures are available in the **supplementary** material at https://github.com/vis-graphics/treemaps-literacy.

Discussion

Based on our findings, we can answer our research question that while students in both the groups demonstrated significant learning improvements, students in the Treemaps only group (T) performed slightly better than the students in the Constructivist group ((NLT) on the majority of the assessment questions.

We were expecting students in the NLT group to gain a better understanding of Treemaps due to the linked, coordinates views of the synchronized node-link diagram. It may be that the students in the NLT group relied more on the "familiar" node-link diagrams when answering questions in the Exploratory Tasks phase (see Fig.3), whereas students in the Treemaps only (T) group had no choice but to use the interactive Treemap to answer questions. That may have led the students in the Treemaps only (T) group to obtain a better understanding of Treemaps.

Conclusion

We presented our findings on evaluating a Constructivist approach to teaching Treemaps to students. Based on conducting the intervention, we found that students in the Constructivist group *received lower scores* than the students that learned through *only* an interactive Treemap. Students in the both the groups reported an improved understanding of Treemaps and found the online tools to be effective to learning about Treemaps.

The results of our user-study on constructivism were surprising to us and did not support the original hypothesis we formulated at the beginning of the project. In general, we now believe the coordinated, linked views are not an optimal strategy to support visualization literacy. One theory as to explain why, is that a linked view like the node-link diagram actually acts as a distractor to the target view, in this case, a Treemap. Other views, such as a node-link view or a traditional tree hierarchy view, merely serve to steer the user's focus away from the visual design at hand. Although further studies are necessary to support this hypothesis, our previous work is consistent with this observation [8].

REFERENCES

1. B. Shneiderman, "Tree visualization with tree-maps: 2-d space-filling approach," *ACM Transactions on graphics (TOG)*, vol. 11, no. 1, pp. 92–99, 1992.
2. S. Papert and I. Harel, "Situating constructionism," *constructionism*, vol. 36, no. 2, pp. 1–11, 1991.
3. New York State, "Common core curriculum - Introduction to Networks," https://studylib.net/doc/7241596/precalculus-module-2--topic-a--lesson-1--teacher, 2021, [Online; accessed 23-October-2022].
4. A. Zoss, A. Maltese, S. M. Uzzo, and K. Börner, "Network visualization literacy: novel approaches to measurement and instruction," in *Network Science In Education*. Springer, 2018, pp. 169–187.
5. K. Börner, A. Bueckle, and M. Ginda, "Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments," *Proceedings of the National Academy of Sciences*, vol. 116, no. 6, pp. 1857–1864, 2019.
6. S. Lee, S.-H. Kim, and B. C. Kwon, "VLAT: Development of a visualization literacy assessment test," *IEEE transactions on visualization and computer graphics*, vol. 23, no. 1, pp. 551–560, 2016.
7. P. Ruchikachorn and K. Mueller, "Learning visualizations by analogy: Promoting visual literacy through visualization morphing," *IEEE Transactions on Visualization and Computer Graphics*, vol. 21, no. 9, pp. 1028–1044, 2015.
8. E. E. Firat, A. Denisova, and R. S. Laramee, "Treemap Literacy: A Classroom-Based Investigation," in *Eurographics 2020 - Education Papers*, M. Romero and B. Sousa Santos, Eds. The Eurographics Association, 2020.
9. E. E. Firat, C. Lang, B. Srinivas, I. Peng, R. S. Laramee, and A. Joshi, "A Constructivism-based Approach to Treemap Literacy in the Classroom," in *Eurographics 2023 - Education Papers*, A. Magana and J. Zara, Eds. The Eurographics Association, 2023.
10. S. Pandey and A. Ottley, "Mini-VLAT: A Short and Effective Measure of Visualization Literacy," *Computer Graphics Forum*, 2023.

Elif E. Firat is a lecturer in the Department of Computer Engineering at Cukurova University. Her research interests are in the field of information visualization and visualization literacy. Contact her at elifemelfirat@gmail.com.

Chandana Srinivas is an undergraduate student in the Department of Computer Science at the University of San Francisco. Contact her at csrinivas2@usfca.edu.

Colm Lang is an undergraduate student in the Department of Computer Science at the University of San Francisco. Contact him at cplang@usfca.edu.

Bhumika Srinivas is a graduate student in the Data Science Institute at the University of San Francisco. Contact her at bsrinivas@usfca.edu.

Robert S. Laramee is a Professor of Computer Science and Data Visualization at the University of Nottingham. The focus of his research is on information visualization, scientific visualization, and visual analytics. Contact him at Robert.Laramee@nottingham.ac.uk.

Alark P. Joshi is a Professor in the Department of Computer Science at the University of San Francisco. His research has focused on data visualization and computer science education. Contact him at apjoshi@usfca.edu.