A Study of Mental Maps in Immersive Network Visualization

Joseph Kotlarek* University of California, Davis
Oh-Hyun Kwon† University of California, Davis
Kwan-Liu Ma‡ University of California, Davis
Peter Eades§ University of Sydney
Andreas Kerren¶ Linnaeus University
Karsten Klein‖ University Konstanz
Falk Schreiber** University of Konstanz

2020 IEEE Pacific Visualization Symposium (PacificVis) 978-1-7281-5697-2/20/$31.00 ©2020 IEEE 10.1109/PacificVis48177.2020.4722

ABSTRACT

The visualization of a network influences the quality of the mental map that the viewer develops to understand the network. In this study, we investigate the effects of a 3D immersive visualization environment compared to a traditional 2D desktop environment on the comprehension of a network's structure. We compare the two visualization environments using three tasks—interpreting network structure, memorizing a set of nodes, and identifying the structural changes—commonly used for evaluating the quality of a mental map in network visualization. The results show that participants were able to interpret network structure more accurately when viewing the network in an immersive environment, particularly for larger networks. However, we found that 2D visualizations performed better than immersive visualization for tasks that required spatial memory.

Index Terms: Human-centered computing—Visualization—Visualization techniques—Graph drawings; Human-centered computing—Visualization—Empirical studies in visualization

1 INTRODUCTION

Networks are commonly used for representing relational data found in diverse domains, such as social networks, biological networks, communication networks, and power grids. A non-graphical representation of a network makes it very difficult for humans to understand its structure. Thus, network visualization is widely used to make sense of the structural information in the data. The visualization of a network helps to form a *mental map* of the data, which is the internal representation of the information inside the viewer's mind that is formed by looking at the visualization.

Most existing studies on mental maps in the context of network visualization have focused on dynamic network data [1–4, 27, 29–31]. More specifically, a significant body of the above research has investigated the effect of different layout methods on *mental map preservation*, which is often interpreted as *drawing stability* [1]: minimizing positional changes of nodes and edges of the network in the visualization between successive time periods. While mental maps in network visualization have frequently been studied in the context of dynamic network data, the quality of the mental map is also important for interactive visualization [1], as the interactions performed by the user often changes the visual representation of the data. In the context of this work, we define the quality of a mental map as that which helps the viewer form an accurate understanding of the structure, connectivity, and content of the network. The quality of the mental map is thus influenced by the visual representation of the network.

When designing interactive visualizations, one needs to consider various factors such as input devices and display environments. In recent years, new input and display technologies (e.g., large wall displays, table top displays with touch and pen input devices, and immersive 3D virtual/mixed reality environments) have become ubiquitous enough that researchers can consider these technologies for general visualization designs [12]. Immersive visualization and analytics is an emerging field that aims to leverage the many benefits of mixed reality technology to information visualization, such as *spatial immersion*, *multi-sensory presentation*, and *increased engagement* [12]. In the context of network visualization, immersive technologies have been used for facilitating novel visualization and interaction techniques [15, 19, 23, 24, 26].

In this study, we investigate the effects of an immersive 3D visualization environment on the quality of a mental map for comprehending network data. Specifically, we conduct a controlled user study to compare (1) a 3D immersive visualization environment with a head-mounted display and hand-tracker controllers and (2) a traditional 2D desktop environment with standard monitor and mouse. We study three tasks that are commonly used for evaluating the quality of a mental map in network visualization [4]: interpreting network structure, memorizing a set of nodes, and identifying the structural changes.

The results show that participants completed the tasks faster, and recalled changes to the network more accurately in the desktop environment. However, they interpreted network structures more accurately in the immersive environment, and reported the task as easier to complete in the immersive environment. An overwhelming majority of participants (85%) favored the immersive environment over the desktop environment. The results indicate that immersive environments mitigate confusion caused by edge crossings and node overlaps that are commonly found in 2D representations, while traditional 2D representations provide a better spatial overviews. We discuss the implications of these results for future network visualization designs in immersive environments.

2 RELATED WORK

Immersive environments, such as CAVEs [13] and head-mounted displays (HMDs), have been proven effective for various applications. Ball and North [5] have shown that high resolution tiled displays improve perception and navigation for visual tasks. Mania and Chalmers [25] have studied memory in immersive and non-immersive spaces and found that immersion significantly improved recall for simple memory tasks. Krokos et al. [21] have recently shown that recall can be improved with virtual reality (VR) through techniques such as the memory palace metaphor. Additionally, Kwon et al. [24] have shown that immersive network visualization has clear potential user performance on network interpretation tasks. While immersive technology may have been eschewed in the past, studies such as these in tandem with improvements in stereoscopic displays mark the clear entrance of immersive visualization into the

*e-mail: jtkotlarek@ucdavis.edu
†e-mail: kw@ucdavis.edu
‡e-mail: ma@cs.ucdavis.edu
§e-mail: peter.eades@sydney.edu.au
¶e-mail: andreas.kerren@lnu.se
‖e-mail: karsten.klein@uni-konstanz.de
**e-mail: falk.schreiber@uni-konstanz.de

2020 IEEE Pacific Visualization Symposium (PacificVis)
978-1-7281-5697-2/20/$31.00 ©2020 IEEE

realm of practicality [19].

While keyboard and mouse is the ubiquitous standard for desktop interaction, recent studies show intuitive interaction in VR environments can outperform traditional interaction systems, such as the work by Huang et al. [19] to create a VR gesture system for network visualization. This is no surprise, as Büschel et al. [7] state, intuitive and low-effort interaction is key to leveraging the benefits of immersion, and keeping users invested in the environment. With different methods of interaction, however, different methods of visualization are also in order.

There has been a significant amount of research adapting virtual reality and other immersive technology to applications in a wide breadth of fields, including biomedical imaging [33], scientific visualization [6, 9, 34], education [17, 32, 36, 38], and collaboration [8]. These studies primarily focused on the application in their respective fields. While immersive scientific visualization was quick to establish itself, the impact of immersion on abstract data visualization remains largely unexplored. Indeed, very few studies have considered the mental map in immersive network visualization. In the past year, however, more research into immersive visualizations with abstract data has been completed. Drogemuller et al. [11] evaluate navigation techniques for 3D network visualizations in virtual reality. Greffard et al. [15] introduced an immersive visualization designed to preserve the mental map. The work in this paper differs from the works above in that we instead investigate the impact of immersion on the mental map.

Mental maps are typically used to measure the quality of a dynamic network layout [2, 10, 27, 29, 30], and the importance of mental map preservation in dynamic layouts has been investigated by several studies [30, 31]. Previous work by Archambault and Purchase [3] investigates mental map preservation in a traditional (non-immersive) environment to show it can help users orientation with tasks such as location and path finding. Herman et al. [18] emphasize the importance of considering *predictability* which is also referred to as *preserving the mental map* in dynamic network layouts. As mentioned above, immersive environments also improve navigation and orientation [5, 8], but no study has yet combined these techniques.

3 Experiment

The main purpose of our study is investigating the difference in mental map qualities between two network visualization conditions: immersive 3D visualization and traditional 2D visualization. We designed a within-participant experiment: 2 *visualization conditions* $\times$ 3 *tasks* $\times$ 3 *networks*. The two visualization conditions are compared in terms of *task completion time* and *accuracy*. This section describes the considerations and design of the experiment.

3.1 Visualization Conditions

There are many different factors in designing a network visualization, such as the layout of the network, color scheme, and interaction techniques. To remove possible confounding factors, we need to focus on certain key factors that we wish to compare, while keeping other factors the same. In this study, we focus on the dimensionality (2D and 3D) of a visualization environment. For this, we designed the following two visualization conditions:

- ***2D***: This condition visualizes a network using a 2D layout in a traditional desktop display. In this condition, users use a mouse for interactions: highlight and select nodes, and navigate (pan and zoom) a network.
- ***3D***: This condition renders a network using a 3D layout in an immersive head-mounted display (HTC Vive Pro). Users use controllers for interactions, as shown Fig. 1. Room-scale tracking is used for tracking position and rotation of the HMD and controllers.

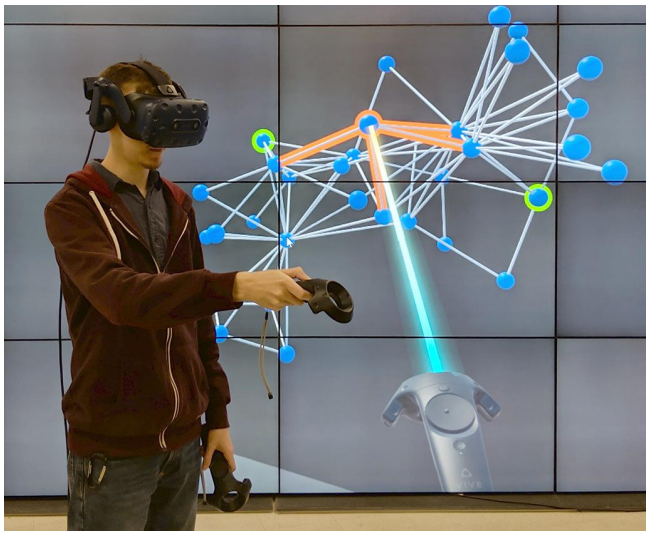


Figure 1: A user interacts with an immersive 3D network visualization using a head-mounted display and controllers. A virtual laser pointer is implemented for highlighting and selecting nodes in a network. In addition, users can scale and relocate the network visualization using pinch-to-zoom interaction with both controllers. The wall display behind the user mirrors the user's view. We compare this immersive network visualization with traditional desktop 2D network visualization to determine how it affects users' mental maps. This figure shows the *karate* network in a 3D layout, which was used in the training session.

3.1.1 Layout

The layout of a network is one of the most critical factors in designing a network visualization with a node-link diagram. While several layout methods are designed specifically for immersive network visualization (e.g., spherical layout methods [23, 24]), they often use additional design factors that are outside of the scope of this study (e.g., edge bundling). In addition, different layout methods can produce greatly varying layout results [14, 22]. Therefore, we use the same layout method (FM^3 layout method [16]) with the same parameter setting for both conditions. In both conditions, nodes are drawn as spheres and edges are rendered as thin cylinders with the same color scheme. The network visualizations used in the experiment are shown in Fig. 3 and the supplementary material. For *2D*, the network is initially laid out to utilize the majority of a 30-inch display, with an aspect ratio of 16:10. For *3D*, the network is initially scaled to fill a 1 m^3 cube and placed in front of the user in the room-scale area.

3.1.2 Interaction

Several interaction techniques are required to carry out the tasks in the experiment. For example, basic navigation techniques (e.g., zoom and pan) are needed to explore the structure of a network in the visualization. In addition, users need to select several nodes to complete the tasks (e.g., Fig. 2).

Most users are already familiar with the basic interaction techniques in a desktop environment using mouse and keyboard, such as click-to-select and scroll-to-zoom. Thus, it is important to design intuitive interaction techniques for the immersive environment to reduce the effect of interaction techniques for completing the tasks, which is not the main focus of this study.

Selection in a traditional 2D desktop display with keyboard and mouse is a well-established interaction, which most commonly implemented as a left-click with the mouse. However, designing an effective selection interaction for an immersive display environment is not as straightforward as the desktop environment. Since motion

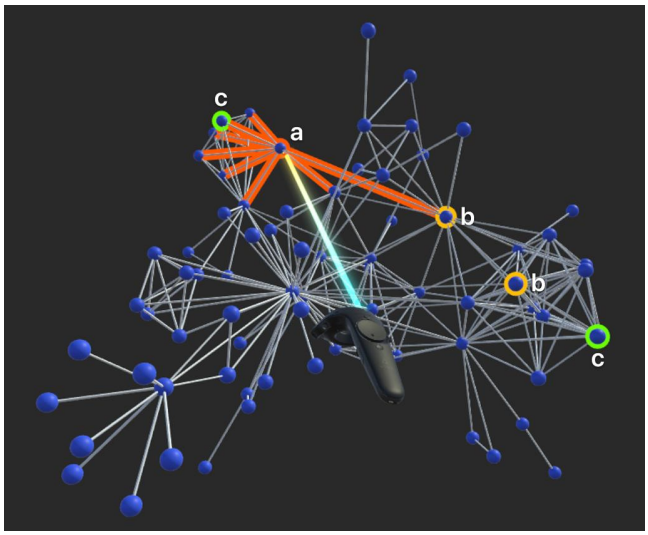


Figure 2: Highlighting and selecting nodes. (a) The orange node is being highlighted over by the user, ready to be selected. The edges connected to the highlighted node are also highlighted in orange while the user is pointing it with the controller. (b) The yellow nodes are the selected nodes by the user. (c) The light green nodes are pre-highlighted by the system to show task information to the user. This figure shows a *path* task. The goal is to find a shortest path between the light green nodes (c), where the nodes in the path needs to be selected (b). Participant can inspect the connected edges of a node by highlighting the node (a) with the virtual laser pointer. This figure shows the *lesmis* network in a 3D layout.

tracking is a common feature of commercially available stereoscopic displays, such as the HTC Vive and Oculus Rift, we leverage the ability to select objects by pointing with a tracked controller.

Users can select a node by first pointing the target node and then pressing a button. This virtual laser pointer technique is one of the most common ways to implement selection in virtual reality. We use this technique because of its intuitiveness and familiarity to most people with any amount of experience in VR or laser pointers.

Our virtual laser pointer implementation does have two minor improvements to make the experience better. The first is to reduce clutter in the visualization, the laser pointer is only visible while the user touches lightly on the touch-pad of the controller (Fig. 1) and a selection is counted when the user presses the touch-pad until an audible "click" is heard. The other improvement is an error tolerance for selecting distant objects. This helps significantly, as it is possible with our system to have nodes that take up less than half a degree of the user's field of view, making them very difficult to select with a simple ray casting. In our pilot study, we found that a cone of approximately two degrees at the apex made selection of distant objects much easier without causing the opposite problem of unintentionally selecting objects far from the pointer.

To match the ease of use to *scale* the network with the mouse scroll wheel in the traditional desktop environment, we implement a pinch-to-zoom technique for the immersive environment. To activate the pinch-to-zoom, the user must press a button on both tracked hand-held controllers, then stretch or contract the distance between the two controllers to scale the network by an equivalent factor. In addition, users can *move* the network by clicking and dragging in the desktop environment, and by dragging the controllers with the pinch-to-zoom button held in the immersive environment. This design is also motivated by intuitiveness and popularity. The pinch-to-zoom technique is widely used in many touch-based systems that most users are already familiar with, such as smartphones and tablets.

In the *2D* condition, *rotating* the network is not necessary for completing the tasks used in this study. However, the ability to look at the network from different directions is necessary with the 3D layout to address possible occlusion issues. Thus, while we exclude the rotation interaction from both conditions, we leverage the physical tracked space allowed by the immersive technology we use in the *3D* condition. By walking around the space, users can view the network from any angle, even though they cannot rotate the network itself.

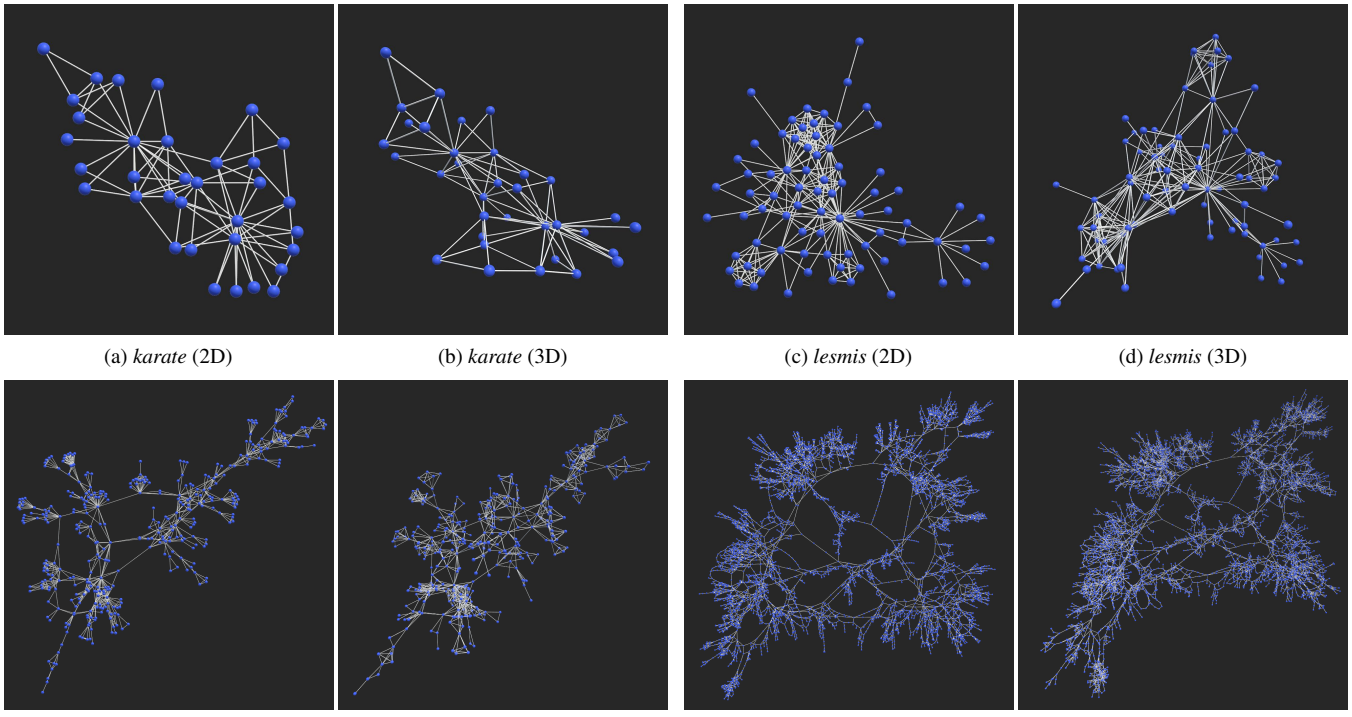
3.2 Tasks

Existing studies have measured the quality of the mental map in the user's mind by human-computer interaction experiments using various tasks. Archambault and Purchase [4] have grouped the tasks into three broad categories: *interpretation*, *memory*, or *change* tasks:

- An interpretation task asks a question that requires the user to understand the structure information of a network. An example task would be to ask the user about node degree or paths between two nodes in a network.
- A memory task requires the user to recall information about the network after viewing the visualization. For example, reconstructing the network on paper from the user's memory.
- A change task asks how the network changes; possible changes are changes in the degree or changes in the overall size.

Considering the above task categories by Archambault and Purchase [4], we design three tasks for the experiment:

- ***Path***: "Highlight nodes that form the shortest path from node A to node B." This is an *interpretation* task [4]. The participant is shown a network with two pre-highlighted nodes. Participants are asked to find a shortest path between the two pre-highlighted nodes by selecting a set of nodes that forms the shortest path. Participants can inspect the connected edges of a node by highlighting the node using the virtual laser pointer (Fig. 2 and Sect. 3.1.2). For each network, the same two nodes are pre-highlighted in both visualization conditions. While there can be multiple shortest paths, the participant only needs to find one.
- ***Memory***: "Select the five nodes that were previously highlighted." Participants are given 30 seconds to memorize five pre-highlighted target nodes in a network visualization. The visualization is then removed, and the participants are shown a blank screen for 10 seconds because immediately showing the network again was found to be too easy in our pilot study. After that, the network reappears with the same position and scale, but without highlighting the target nodes. The goal is to select the five target nodes that were previously highlighted (Fig. 4a–b). For each network, the same five nodes are assigned as the target nodes across all visualization conditions and participants. As we use the same layout method (FM^3 [16]) for both *2D* and *3D* conditions, the location of target nodes are relatively similar between the two conditions, as shown in Fig. 4a–b. This allows a fair comparison between the two conditions by reducing the effect of the target node location.
- ***Change***: "Select nodes that were not a part of the network." Participants are given 30 seconds to explore a modified network, where five target nodes and associated edges are removed from the full network. Then, participants are shown a blank screen for 1 second since immediately showing the new network was found to be too easy in our pilot study. After that, the full network is shown, where the five target nodes and corresponding edges are added back to the network. The goal is to identify the five target nodes (Fig. 4c–d). We removed the same five nodes from the full network to create the modified network for both visualization conditions and all participants.



(a) *karate* (2D) (b) *karate* (3D) (c) *lesmis* (2D) (d) *lesmis* (3D)

(e) *netsci* (2D) (f) *netsci* (3D) (g) *power* (2D) (h) *power* (3D)

Figure 3: The 2D and 3D layouts of the networks used in the experiment. All layouts are computed using FM^3 layout method [16]. The 2D layouts were shown in a 30-inch desktop display (2,560 $\times$ 1,600 px). The 3D layouts were rendered in a head-mounted display (2,880 $\times$ 1,600 px, 1,440 $\times$ 1,600 px per eye). To provide multiple views of a 3D layout from different viewpoints, the supplementary material includes videos of the 3D layouts. Since we used the same layout method (FM^3 [16]), both 2D and 3D layouts of a network show similar spatial proximity between nodes, while their dimensionalities are different. This allows a fair comparison between the two visualization conditions (*2D* and *3D*) by reducing the effect of spatial proximity between nodes.

> We use the layout of the full network for the modified network as well, i.e., we did not use a separate layout for the modified network. This design decision makes our experiment differ from existing experiments on the mental map quality of dynamic network visualization, where different layouts are used per time step of a dynamic network. However, it allows this study to focus on the effect of the two visualization conditions, not the drawing stability that is outside of the scope of this study.

The participants are given a written description of the objective at the start of each task. The duration of exploration and blank screen stages in the *memory* and *change* tasks are determined based on a pilot study. For the *path* task, we compute the accuracy of a selected path based on the ratio between the length of the selected path and the length of the shortest path(s). For the *memory* and *change* tasks, the accuracy of selected nodes by participants is computed based on the graph-theoretic distance from the correct target nodes (more detail in Sect. 4.2).

3.3 Networks

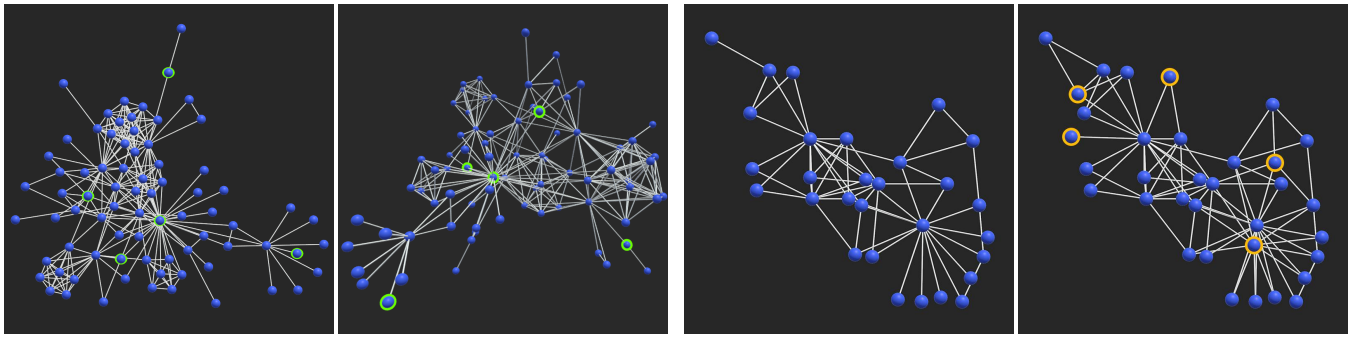
We use three different networks with varying sizes (i.e., number of nodes and edges) with one additional network that was used for the training session.

- ***Karate***: This is the well-known Zachary karate club network [37], which consists of 34 nodes and 78 edges. This network is used in the training session.
- ***Lesmis***: A co-occurrence network of the characters in Victor Hugo's novel *Les Misérables* [20], This network has 77 nodes and 254 edges.
- ***Netsci***: A co-authorship network in the field of network science [28]. We use the largest component of the full network, which consists of 379 nodes and 914 edges.
- ***Power***: The power grid of western states of the United States [35]. This network has 4,941 nodes and 6,594 edges.

3.4 Participants

We recruited 20 participants (11 female, 9 male) for our user study. The mean age of participants was 24.35, ranging from 18 to 34 years. Every participant was familiar with the concepts of virtual reality, and 16 participants had used a virtual reality device before this study. Of those 16, two participants said they had used virtual reality devices extensively. Additionally, every participant was familiar with the concept of a network, and 14 said they were familiar or experienced with network data structures. Four participants were in an undergraduate degree program, while the remaining 16 all possessed undergraduate degrees. 14 of the 16 were also pursuing postgraduate degrees.

Five participants had normal vision, while 15 had corrected vision. Of the 15, 13 wore glasses, and all but one were able to wear their glasses comfortably within the HTC Vive Pro HMD. The one participant that had to remove his/her glasses reported that he/she was still able to read text and see the nodes/edges clearly in the immersive environment. One participant had deuteranopia vision (self-reported), but reported that all colors used in the study were easily distinguishable.



(a) *memory* task with the *lesmis* network in the *2D* condition

(b) *memory* task with the *lesmis* network in the *3D* condition

(c) *change* task with the *karate* network in the *2D* condition (before change)

(d) *change* task with the *karate* network in the *2D* condition (after change)

Figure 4: Examples of *memory* and *change* tasks. For the *memory* task (a and b), participants explore the visualization for 30 seconds and are asked which nodes were highlighted. Since the same layout method (FM^3 [16]) was used for computing both 2D and 3D layouts, the location of target nodes are relatively similar in both 2D and 3D layouts. This allows a fair comparison between the two visualization conditions by reducing the effect of the target node location. For the *change* task, a modified network with five missing nodes (c) is first shown for 30 seconds. Then, participants are asked which nodes were missing when shown the full network (d) without any missing nodes. The highlighted nodes in (d) are the missing nodes in (c). We used the layout of the full network (d) for the modified network (c) as well since this study focuses on the effect of the two visualization conditions (*2D* and *3D*), not the drawing stability of the layout method we used.

3.5 Apparatus

Both visualization conditions are implemented as Unity3D applications. For the *3D* condition, participants use the HTC Vive Pro HMD and are instructed to stand in the middle of the room-scale environment at the start of each task, facing toward the area where the networks would be presented. The HTC Vive Pro HMD has a 2,880 $\times$ 1,600 px (1,440 $\times$ 1,600 px per eye) AMOLED display with a 90 Hz refresh rate. The immersive environment was driven by a desktop computer with an Intel i7 6900K CPU and dual (SLI enabled) NVIDIA GeForce GTX 1080 GPUs. The environment was consistently rendered at 90 frames per second. The tracked physical space for room-scale configuration measures 3.0 m by 3.1 m. The HTC Vive lighthouses are positioned 3.2 m off the ground. The virtual environment is a simple space with a floor indicating the boundary where users can walk. Participants use a 30-inch desktop display with 2,560 $\times$ 1,600 px for the *2D* condition.

3.6 Procedure

After ensuring the participants are aware of possible VR/HMD issues such as sickness or disorientation, we adjust the HMD fit and interpupillary distance (IPD) for each participant to provide optimal viewing conditions. Participants were allowed as much time as necessary to ensure the HMD was fit comfortably without issue. None of the participants experienced any issue with the HMD. Participants then completed a pre-study questionnaire, followed by training, the full experiment, and a post-study questionnaire.

Questionnaire. All participants answered a two-part questionnaire. The first part covered participant demographics, including age, gender, education, colorblindness, perceived spatial reasoning skills, and VR/visualization experience levels. The second part covered perceived task difficulty (7-point Likert scales) as well as free response questions about participant preference.

Training. Participants were allowed up to 10 minutes to familiarize themselves with the visualization conditions and their respective interfaces, in a guided tour of the features of both applications. Each participant received training before performing each task for the first time. This training consisted of performing the same set of tasks with the *karate* network, with the correct answer available at the end.

Procedure. The order of the tasks was the same for each participant: *path*, *memory*, and *change* tasks. The networks were ordered from smallest (*lesmis*) to largest (*power*) within each task. The order of the visualization conditions was counterbalanced such that half the participants started with *2D* and the other half started with *3D* to prevent learning effects. For example, a participant that starts with *2D* would perform tasks in order of *2D-path-lesmis*, *2D-path-netsci*, ..., *2D-memory-lesmis*, *2D-memory-netsci*, ..., then repeating everything with *3D*.

Participants were encouraged to take short rests between tasks, for as long as they needed. The visualization was fully reset between tasks. For the *3D* condition, the participants were instructed between tasks to re-orient themselves in the middle of the space and face the same direction so that movement in previous tasks does not affect the completion time or success rate of subsequent tasks. We informed the participants about their right to stop the experiment at any time.

3.7 Hypotheses

We expected the following results from the user study:

- H1: Task completion time will be shorter with traditional display (*2D*) than with immersive display (*3D*).
- H2: *3D* will outperform *2D* in terms of accuracy.
- H3: Participants will prefer *3D* over *2D* for exploring network data.

4 Results

The summary of results are shown in Fig. 5, where the colored bars show that the difference between *2D* and *3D* was statistically significant for the corresponding task and network. The gray bars in Fig. 5 show that the difference was not statistically significant. In addition, Fig. 6 shows the distribution of task difficulty ratings reported by participants.

Overall, participants performed faster with the *2D* condition than the *3D* condition for all three tasks, confirming H1. In terms of accuracy, participants performed better with the *3D* condition than the *2D* condition for the *path* task, as in H2. However, for the *memory* task, participants' responses were more accurate with the *2D* condition than with the *3D* condition, opposed to H2. Participants reported that the *path* task was easier in the *3D* condition than the *2D* condition. An overwhelming majority of participants (85%) favored the *3D* condition over the *2D* condition, as we expected in H3.

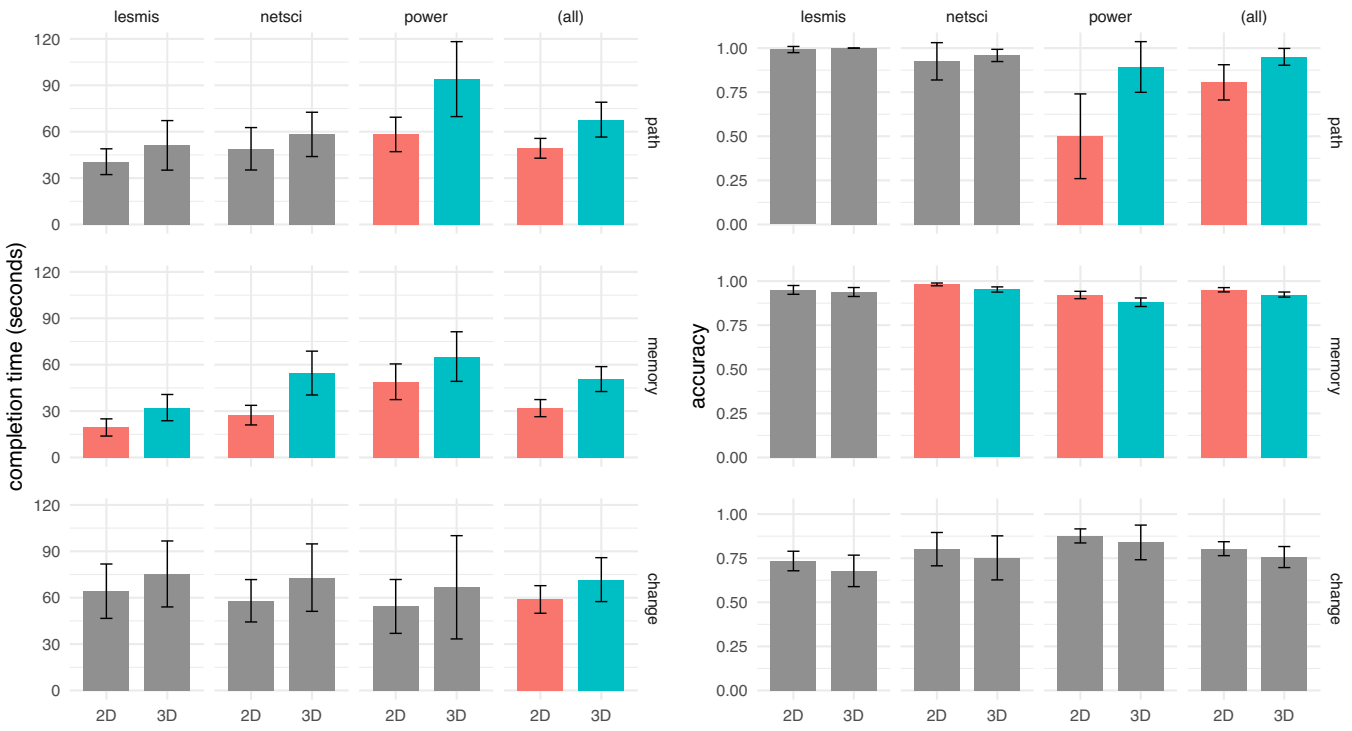


Figure 5: Summary of completion time (left) and accuracy (right) of the experiment. Each row of bar charts represents one of the three tasks (*path*, *memory*, and *change*). Each column corresponds to one of the three datasets (*lesmis*, *netsci*, and *power*) or all the datasets (*all*). Each bar represents the mean value of the corresponding responses. The error bars represent 95% confidence intervals. The colored bars show that the difference between the *2D* and *3D* conditions was statistically significant. The gray bars mean that there was no statistical significance in the difference between *2D* and *3D*.

4.1 Task Completion Time

Task completion time is measured with a timer built into the system created for this experiment. The timer starts when the user is able to start interacting with a network, and it ends when the user pressed the "continue" button to proceed to the next task. None of the tasks directly indicate that a user has chosen the correct answers, and so some users chose to be more meticulous than others when checking their work. As a result, there is high variance in the time required to complete all of the tasks. We compare each task and also network separately because the completion time varies depending on task and network.

Path task. On average for both conditions, participants completed the *path* task in 58.53s ($SD = 36.51$). They completed the *path* task faster with the *2D* condition, taking 49.26s ($SD = 24.77$), and slower with the *3D* condition, taking 67.81s ($SD = 43.60$). A paired t-test showed that this difference is statistically significant ($t(59) = 3.744$, $p = 0.0004$, 95% CI of the difference $= 8.637$ to 28.467). Cohen's d using the pooled standard deviation (SD_{pool}) suggested that the effect size is medium ($d = 0.523$).

When we analyzed the completion time of *path* task for each network, there was a significant difference only for the *power* network ($t(19) = 3.359$, $p = 0.003$, 95% CI of the difference $= 13.487$ to 58.076), where participants took 58.22s ($SD = 23.79$) with *2D* but took 94.00s ($SD = 51.80$) with the *3D* condition. Cohen's d using SD_{pool} showed that the effect size is large ($d = 0.888$). We did not observe a significant difference between *2D* and *3D* conditions for *lesmis* ($p = 0.092$) and *netsci* ($p = 0.229$).

Memory task. On average for both conditions, participants completed the *memory* task in 41.31s ($SD = 28.25$). They completed the *memory* task faster with the *2D* condition, taking 31.92s ($SD = 21.42$), and slower with the *3D* condition, taking 50.70s ($SD = 31.18$), which is shown by t-test to be statistically significant ($t(59) = 5.847$, $p = 2.318 \times 10^{-7}$, 95% CI of the difference $= 12.356$ to 25.206). Cohen's d using SD_{pool} suggested that the effect size is medium to large ($d = 0.702$).

The differences between the *2D* and *3D* conditions for each network were also significant. For the *lesmis* network, participants completed the *memory* task significantly faster ($t(19) = 4.887$, $p = 0.0001$, 95% CI of the difference = 7.339 to 18.333) with the *2D* condition (mean = 19.41s $SD = 11.88$) than with the *3D* condition (mean = 32.25s, $SD = 18.12$). Cohen's d using SD_{pool} showed that the effect size is large ($d = 0.838$). For the *netsci* network, participants completed the *memory* task significantly faster with the *2D* condition than with the *3D* condition, having mean completion times of 27.38s ($SD = 13.54$) and 54.59s ($SD = 30.30$), respectively ($t(19) = 3.959$, $p = 0.0008$, 95% CI of the difference $= 12.824$ to 41.596). Cohen's d using SD_{pool} showed that the effect size is large ($d = 1.160$). Participants also completed the *memory* task significantly faster ($t(19) = 2.714$, $p = 0.014$, 95% CI of the difference $= 3.729$ to 28.856) with the *2D* condition than with the *3D* condition for the *power* network, having mean completion times of 48.96s ($SD = 24.65$) and 65.26s ($SD = 34.26$), respectively. Cohen's d using SD_{pool} suggested that the effect size is medium ($d = 0.546$).

Change task. On average for both conditions, participants completed the *change* task in 65.28s ($SD = 46.16$). They completed the *change* task faster with the *2D* condition (mean = 58.86s $SD =$ 34.54) than with the *3D* condition (mean = 71.69, $SD = 54.97$). A paired t-test shows that this difference is statistically significant ($t(59) = 2.397$, $p = 0.020$, 95% CI of the difference $= 2.118$ to 23.546. However, Cohen's d using SD_{pool} showed that the effect size is small ($d = 0.280$). In addition, there was no significant

difference between the *2D* and *3D* conditions when we analyzed the completion time of the *change* task for each network individually.

4.2 Accuracy

We define the accuracy of a response according to each task.

Path task. For the *path* task, the accuracy (Equation 1) is based on the ratio between the length of the selected path and the length of the shortest path(s). The participants were asked to find the shortest path between two pre-highlighted nodes, and so the correct answer would be any complete path that is the same length as the shortest path between those two nodes. Thus, we use the shortest path length as a fraction of the participant's selected path when the participant makes a valid path, and an accuracy of zero is used for invalid paths (e.g., disconnected nodes), shown in the following equation.

$$\text{accuracy} = \frac{\text{the length of the shortest path}}{\text{the length of the selected path}} \tag{1}$$

With this definition, the best accuracy score a participant can receive is 1, and any path that is valid but contains extra nodes will be between 0 and 1, approaching 0 as more redundant nodes are added.

Overall, participants succeeded in making a valid path in 89.17% of all responses, and that path was shortest in 80.83% of all responses. Participants performed significantly better in the *3D* condition with the mean accuracy of 0.950 out of 1.0 ($SD = 0.185$), compared to the mean accuracy in the *2D* condition, at 0.806 ($SD = 0.387$). A paired t-test confirms that this difference is statistically significant ($t(59) = 2.820$, $p = 0.0065$, 95% CI of the difference $= 0.042$ to 0.248). Cohen's d using SD_{pool} suggested that the effect size is small to medium ($d = 0.478$).

When we analyzed the accuracy for the *path* task per network separately, there was a significant difference only for the *power* network ($t(19) = 2.970$, $p = 0.008$, 95% CI of the difference $=$ 0.116 to 0.670), where the mean accuracy was 0.500 ($SD = 0.513$) with *2D* but was 0.893 ($SD = 0.307$) with the *3D* condition. Cohen's d using SD_{pool} suggested that the effect size is small to large ($d =$ 0.929). We did not observe a significant difference between *2D* and *3D* conditions for *lesmis* ($p = 0.330$) and *netsci* ($p = 0.479$).

Memory task. If we only consider whether a selected node is a target node or not, it is not possible to measure how close the selected node is to a target node. Thus, we define the accuracy of a response for the *memory* task based on the graph-theoretic distance between a selected node by participants and the closest target node:

$$\text{accuracy} = 1 - \frac{\text{the distance from the selected node}}{\text{the diameter of the network}} \tag{2}$$

We normalize the distance by the diameter of a network, which is the maximum distance between any pair of nodes in the network. Using this definition, the accuracy of a correctly selected node is 1, and the accuracy of an incorrectly selected node will be between 0 and 1. We use the mean of this accuracy score of participant's responses. Some participants provided fewer than five responses (i.e., the number of target nodes). In that case, we considered the accuracy of the remaining responses is 0 as they could not remember some target nodes at all. For example, if a participant exactly identified three target nodes (accuracy of 1) but did not provide any more responses (accuracy of 0), the accuracy of the responses by this participant is 0.6. A single target node can be the closest one to multiple selected nodes, which can lead to incorrect accuracy scores. However, there were no responses that had this issue.

Although Euclidean distance is a possible option for measuring the accuracy of a response instead of graph-theoretic distance, the physical sizes of the *2D* and *3D* conditions are quite different. Also, due to different dimensionalities of 2D and 3D layouts of a network, it is not clear to define a normalization scheme for comparing Euclidean distances in the two visualization conditions. Thus, we use the graph-theoretic distance for measuring the accuracy of response.

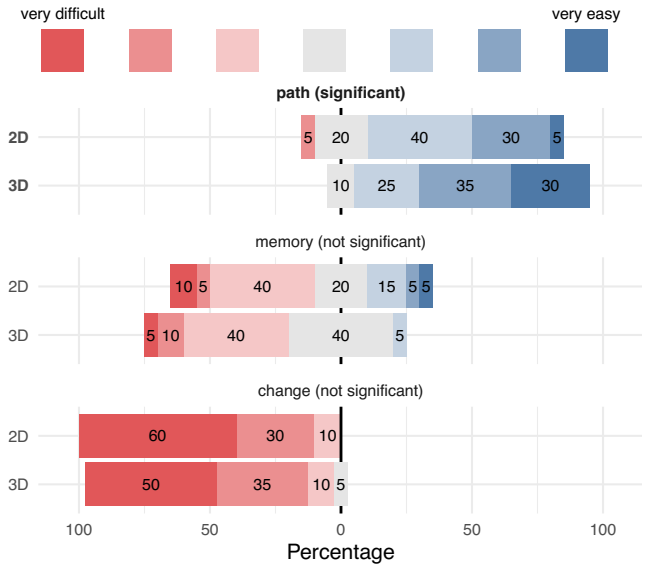


Figure 6: User responses to questions about the difficulty of each task and visualization condition.

For all conditions and networks, the mean accuracy of responses was 0.937 ($SD = 0.053$) for the *memory* tasks, where participants exactly identified 62.67% of the target nodes. The accuracy of responses were higher in the *2D* condition (mean $= 0.951$, $SD = 0.048$) than the *3D* condition (mean $= 0.923$, $SD = 0.056$). A paired t-test revealed that the difference is statistically significant ($t(59) = 4.2086$, $p = 8.88 \times 10^{-5}$, 95% CI of the difference $= 0.014$ to 0.040). Cohen's d using SD_{pool} showed that the effect size is medium ($d = 0.527$).

Comparing *memory* task accuracy per network separately shows that the differences in mean accuracy between the *2D* and *3D* conditions were significant for the *netsci* and *power* networks, but not for the *lesmis* network ($p = 0.316$). For the *netsci* network, the mean accuracy of responses with the *2D* condition was 0.981 ($SD = 0.017$) while the mean accuracy with *3D* was 0.952 ($SD = 0.031$), and the difference was statistically significant confirmed by a paired t-test ($t(19) = 3.940$, $p = 0.0009$, 95% CI of the difference $= 0.014$ to 0.044). Cohen's d using SD_{pool} showed that the effect size is large ($d = 1.146$). For the *power* network, a paired t-test showed that the difference of the mean accuracy between *2D* (mean $= 0.921$, $SD = 0.045$) and *3D* (mean $= 0.880$, $SD = 0.052$) conditions was statistically significant ($t(19) = 3.058$, $p = 0.0065$, 95% CI of the difference $= 0.013$ to 0.069). Cohen's d using SD_{pool} suggested that the effect size is large ($d = 0.852$).

Change task. We use the accuracy definition of the *memory* task (Equation 2) for the *change* task as well. Overall, the mean accuracy for all networks and conditions were 0.780 ($SD = 0.196$). The mean accuracy was higher in the *2D* condition (mean $= 0.804$, $SD = 0.153$) than the *3D* condition (mean $= 0.756$, $SD = 0.231$). However, the difference was not statistically significant ($p = 0.458$). Moreover, there were no significant differences between *2D* and *3D* conditions when we analyzed for each network individually.

4.3 User Feedback

Participants were asked to complete post-study questionnaires that provide more insight into what each person experienced between the two visualization conditions.

We asked participants to rate the perceived difficulty of each task with each condition on a 7-point Likert scale, where 1 meant "very difficult" and 7 meant "very easy." The overall distribution of the responses is shown in Fig. 6. Participants reported that the *path* task was easier using the *3D* condition (median = 6, IQR = 2) than *2D* (median = 5, IQR = 1.25). A Wilcoxon signed-rank test (a non-parametric test equivalent to the paired t-test) showed that the difference was statistically significant ($V = 11$, $p = 0.007$). However, participants reported that the *2D* and *3D* conditions are not different in terms of difficult for *memory* ($V = 78$, $p = 0.306$) and *change* ($V = 25$, $p = 0.478$) tasks.

It is clear, however, that the *path* task was the easiest task (median = 5.5, IQR = 1), the *change* task was the hardest (median = 1, IQR = 1), while the *memory* task was somewhere in between (median = 3, IQR = 1). A Friedman test (non-parametric alternative to one-way repeated measures ANOVA) showed that the difference is statistically significant ($\chi^2(2) = 71.563$, $p = 2.886 \times 10^{-16}$)

85% of the participants (17 out of 20) reported that they preferred *3D* over *2D*, showing a clear majority. To the question of what they liked about either *2D* or *3D*, the following lists some of the participants' most common and insightful responses, starting with most the frequent sentiments:

- *"3D makes it easier to grasp the structure."*
- *"3D networks are easier and/or more interesting to explore"*
- *"Edge occlusion was much easier to resolve within 3D than 2D space."*
- *"Node selection is too much effort in 3D*
- *"2D layout is better for quick overview and memory"*

5 Discussion

The results suggest that one visualization condition did not consistently outperform the other. Although participants took more time to complete tasks with the immersive 3D visualization, they interpreted the structure of a network more accurately (*path* task). The traditional 2D desktop visualization environment was better for the *memory* task, where participants performed better in terms of both task completion time and accuracy than the immersive 3D condition. In addition, the two visualization conditions performed similarity in the *change* task except for the overall task completion time.

Finding the shortest path was significantly easier and more accurate with the immersive 3D visualization than the traditional 2D desktop environment. A significant portion of this can be attributed to the nature of 3D layouts having much fewer *edge crossings*. Since the 3D layouts used in the immersive space were almost completely devoid of any crossed edges, it was never the case that a user would be unable to tell whether the path they had highlighted was complete or not, as each edge was easy to follow through the environment from start to finish. Even with unavoidable occlusions in 3D, participants could separate the edges by simply moving their head to change the viewpoint. Several participants commented on this:

- *"Easy to see the links between the nodes (3D)"*
- *"Better sense of the topological structure (3D)"*
- *"3D makes it easier to grasp more of the structure of the graph"*
- *"It looked like they overlapped and it was harder to tell where the nodes were on the 2D display."*
- *"It felt easier to see and understand the structure of the graph in 3D."*
- *"VR version takes advantage of depth to make graph clearer."*
- *"Edge overlapping was a lot easier to resolve in VR than in 2D space"*

In the case of the traditional 2D environment, and especially with the largest network (*power*), it was common for edges to not only cross other edges, but also travel through nodes, making it unclear whether it was a single edge that went through the node, or two edges that happened to connect to the same node at opposite sides. If not for the edge-highlighting that was built into the system, it would be obvious that a 3D layout should do much better and any other factors would likely be insignificant. However, there is likely more at play here. The edge-highlighting allowed participants to see very precisely if two nodes were connected simply by hovering over it, and all participants were instructed in the use of this feature. In spite of this, participants with the *2D* condition still marked incomplete paths and complete but non-optimal paths more frequently than participants with the *3D* condition. In short, the immersive experience helped participants complete the path finding task more accurately beyond the 3D nature of the layout. This could be evidence of the natural formation of a 3D mental map that helped participants understand the structure of a network and connections between nodes better than could happen in the 2D traditional display environment.

Participants performed slower with the immersive 3D environment in most regards compared to the traditional 2D setup. Part of this may have been due to participants' lack of familiarity with the controls and virtual environment. In addition, the ability to show a better overview of the *2D* condition seems to affect the overall completion time. Some comments from participants are related to these findings:

- *"Selecting a node is much easier in 2D"*
- *"Selection of nodes was easier with a mouse than the immersive 3D controller due to a mouse being more precise and not prone to unintended movements."*
- *"For very precise applications, I might find it useful to use the VR display in conjunction with traditional mouse+keyboard controls."*
- *"The desktop gave me more familiarity in terms of viewing the graph and how to maneuver it."*
- *"A graph can fit within a window size well in 2D"*
- *"2D is easy to remember the positions"*
- *"It is easier to have an overview with the desktop"*

Since the *2D* condition did not require any rotation, we also exclude the rotation interaction in the *3D* condition to remove a possible confounding factor. In addition, many participants in the pilot study were disoriented after a few rotations of the network. Thus, we let users view the network from different angles by walking around the space rather than directly rotating the network. However, the observations on participants' movement during the experiment and the feedback from participants suggest that the physical movements required for finding good viewpoints might lead to taking more time in the immersive environment. Two participants specifically commented that rotating the network visualization with the controllers will be helpful:

- *"I felt that having the ability to rotate the graph in VR would have been helpful."*
- *"Add rotation to VR especially."*

However, these two participants reported that they are moderately or highly experienced with virtual reality technologies. Investigating the effect of spatial awareness in immersive visualization can be an interesting research topic for future study.

In addition, participants in the immersive environment took more time not necessarily because they didn't find answer as quickly, but because they often spent extra time exploring the network or having fun with the visualization before continuing to the next task, something they did not do with the traditional visualization conditions. Here are several related comments from participants:

- *"It's fun to zoom in and zoom out in VR settings"*

- *"3D was more fun and it is easier to visualize where the nodes were."*
- *"I liked the 3D version because it allowed me to step inside the graph and see the nodes hidden in the back."*
- *"3D makes it easier to explore."*

The mental maps formed through this experiment seem to be limited due to the short time frame of all the tasks. Since participants were given only 30 seconds to scrutinize each visualization (*memory* and *change* tasks), they only had time to construct a relatively immature mental map. This mental map was too fragile for immersive space, based on the overall lower accuracy in immersive 3D for the *memory* task. Participants mentioned on several occasions that after having changed their perspective in the immersive environment, they lost their track of the locations of target nodes in the *memory* task. The difficulty of the *change* task seems to have contributed to some unexpected results. Participants answered overwhelmingly that this task was the hardest. While all participants finished the entire set of trials, a few participants gave up on completing the task for the largest network (*power*), by either selecting randomly or selecting less than five nodes. We believe the results can be different if more time was allowed for explore and memorize a network for both *memory* and *change* tasks.

Another way to reduce the difficulty of the *change* task is allowing users to see both structures (before and after adding nodes), similar to the experiments of drawing stability in dynamic network visualization techniques. However, this could make the task too easy, which is what we found during our pilot study. In dynamic network visualization, different layouts are used for each time step, while we used a single layout for both steps as our focus was not the drawing stability. This leads to another research question: would immersive environments be helpful for dynamic network visualization?

To achieve high internal validity, we have carefully designed the two visualization conditions by (1) keeping the factors that are not the focus of the study the same in both conditions (i.e., the same layout algorithm, interaction paradigm, and color scheme) and (2) counterbalancing the order of visualization conditions. In addition, the user's performance is measured using well-defined tasks and measurements to have high construct validity.

However, our study has limitations in terms of external validity (generalizability). Since we considered the entire visualization environment as a single condition, our study only shows the effect of the combination of multiple factors of each environment, but not the effect of individual factors. Thus, the results can vary, for example, with different display and interaction devices. Moreover, different data and tasks can lead to different results. Further studies are needed to evaluate various factors of network visualization.

6 Conclusion

Literature on the mental map in network visualization is substantial but sparsely populated when it comes to the immersive visualization design, despite the rapid development of immersive technologies. This work investigates the mental maps of viewers under different visualization conditions. The immersive 3D network visualization proved to be more helpful for interpreting the structure of a network. It suggests that immersive technology may allow deeper and more intuitive understanding of abstract data structures, though it is more carefully considered and must be precisely designed. The traditional 2D visualization has proven superior for spatial memory as they provide a better overview of the entire network.

To further investigate the use of immersive technology in understanding abstract data structures, such as networks, the next step is to delve deeper into the individual processes for observing and understanding immersive environments with longer and more detailed experiments. These are processes such as how users mentally store relative positions and directions, and how they are able to turn that into an understanding of the data presented. Investigating the mental capacity of humans in immersive environments will help us design better systems that leverage human perception more effectively.

Acknowledgments

This research is sponsored in part by the U.S. National Science Foundation through grants IIS-1741536 and IIS-1528203.

References

[1] D. Archambault, K. Klein, and K. Misue. Reimagining the Mental Map and Drawing Stability. Technical Report 2018-12, National Institute of Informatics Shonan Meeting, 2018.

[2] D. Archambault, H. Purchase, and B. Pinaud. Animation, Small Multiples, and the Effect of Mental Map Preservation in Dynamic Graphs. *IEEE Transactions on visualization and Computer Graphics*, 17(4):539–552, 2011.

[3] D. Archambault and H. C. Purchase. Mental Map Preservation Helps User Orientation in Dynamic Graphs. In *International Symposium on Graph Drawing*, pages 475–486. Springer, 2012.

[4] D. Archambault and H. C. Purchase. The "Map" in the Mental Map: Experimental Results in Dynamic Graph Drawing. *International Journal of Human-Computer Studies*, 71(11):1044–1055, 2013.

[5] R. Ball and C. North. Effects of Tiled High-Resolution Display on Basic Visualization and Navigation Tasks. In *Proc. ACM CHI '05 Extended Abstracts on Human Factors in Computing Systems*, pages 1196–1199, 2005.

[6] S. Bryson. Virtual Reality in Scientific Visualization. *Commun. ACM*, 39(5):62–71, 1996.

[7] W. Büschel, J. Chen, R. Dachselt, S. Drucker, T. Dwyer, C. Görg, T. Isenberg, A. Kerren, C. North, and W. Stuerzlinger. Interaction for Immersive Analytics. In *Immersive Analytics*, pages 95–138. Springer, 2018.

[8] M. Cordeil, T. Dwyer, K. Klein, B. Laha, K. Marriott, and B. H. Thomas. Immersive Collaborative Analysis of Network Connectivity: CAVE-style or Head-Mounted Display? *IEEE Transactions on visualization and Computer Graphics*, 23(1):441–450, 2017.

[9] M. de Ridder, Y. Jung, R. Huang, J. Kim, and D. D. Feng. Exploration of Virtual and Augmented Reality for Visual Analytics and 3D Volume Rendering of Functional Magnetic Resonance Imaging (fMRI) Data. In *Proc. International Symposium on Big Data Visual Analytics*, pages 1–8. IEEE, 2015.

[10] S. Diehl, C. Görg, and A. Kerren. Preserving the Mental Map using Foresighted Layout. In D. S. Ebert, J. M. Favre, and R. Peikert, editors, *Data Visualization 2001*, pages 175–184, Vienna, 2001. Springer Vienna.

[11] A. Drogemuller, A. Cunningham, J. Walsh, M. Cordeil, W. Ross, and B. Thomas. Evaluating Navigation Techniques for 3D Graph Visualizations in Virtual Reality. In *Proc. International Symposium on Big Data Visual and Immersive Analytics*, pages 1–10. IEEE, 2018.

[12] T. Dwyer, K. Marriott, T. Isenberg, K. Klein, N. Riche, F. Schreiber, W. Stuerzlinger, and B. H. Thomas. Immersive Analytics: An Introduction. In *Immersive Analytics*, pages 1–23. Springer, 2018.

[13] A. Febretti, A. Nishimoto, T. Thigpen, J. Talandis, L. Long, J. D. Pirtle, T. Peterka, A. Verlo, M. Brown, D. Plepys, D. Sandin, L. Renambot, A. Johnson, and J. Leigh. CAVE2: A Hybrid Reality Environment for Immersive Simulation and Information Analysis. In *Proc. Engineering Reality of Virtual Reality*, 2013.

[14] H. Gibson, J. Faith, and P. Vickers. A Survey of Two-dimensional Graph Layout Techniques for Information Visualization. *Information Visualization*, 12(3–4):324–357, 2013.

[15] N. Greffard, F. Picarougne, and P. Kuntz. Immersive dynamic visualization of interactions in a social network. In *Challenges at the Interface of Data Analysis, Computer Science, and Optimization*, pages 255–262. Springer, 2012.

[16] S. Hachul and M. Jünger. Drawing Large Graphs with a Potential-Field-Based Multilevel Algorithm. In *Proc. Graph Drawing*, pages 285–295, 2004.

[17] C. R. Hall, R. J. Stiles, and C. D. Horwitz. Virtual Reality for Training: Evaluating Knowledge Retention. In *Proc. IEEE Virtual Reality Annual International Symposium*, pages 184–189, 1998.

[18] I. Herman, G. Melançon, and M. S. Marshall. Graph Visualization and Navigation in Information Visualization: A Survey. *IEEE Transactions on visualization and Computer Graphics*, 6(1):24–43, 2000.

[19] Y.-J. Huang, T. Fujiwara, Y.-X. Lin, W.-C. Lin, and K.-L. Ma. A Gesture System for Graph Visualization in Virtual Reality Environments. In *Proc. IEEE Pacific Visualization Symposium*, pages 41–45. IEEE, 2017.

[20] D. E. Knuth. *The Stanford GraphBase: A Platform for Combinatorial Computing*. Addison-Wesley, 1993.

[21] E. Krokos, C. Plaisant, and A. Varshney. Virtual Memory Palaces: Immersion Aids Recall. *Virtual Reality*, pages 1–15, 2018.

[22] O.-H. Kwon and K.-L. Ma. A Deep Generative Model for Graph Layout. *IEEE Transactions on visualization and Computer Graphics*, 26(1):665–675, 2020.

[23] O.-H. Kwon, C. Muelder, K. Lee, and K.-L. Ma. Spherical Layout and Rendering Methods for Immersive Graph Visualization. In *Proc. IEEE Pacific Visualization Symposium*, pages 63–67, 2015.

[24] O.-H. Kwon, C. Muelder, K. Lee, and K.-L. Ma. A Study of Layout, Rendering, and Interaction Methods for Immersive Graph Visualization. *IEEE Transactions on visualization and Computer Graphics*, 22(7):1802–1815, 2016.

[25] K. Mania and A. Chalmers. The Effects of Levels of Immersion on Memory and Presence in Virtual Environments: A Reality Centered Approach. *CyberPsychology & Behavior*, 4(2):247–264, 2001.

[26] K. Marriott, J. Chen, M. Hlawatsch, T. Itoh, M. A. Nacenta, G. Reina, and W. Stuerzlinger. *Immersive Analytics: Time to Reconsider the Value of 3D for Information Visualisation*, pages 25–55. Springer International Publishing, 2018.

[27] K. Misue, P. Eades, W. Lai, and K. Sugiyama. Layout Adjustment and the Mental Map. *Journal of Visual Languages & Computing*, 6(2):183–210, 1995.

[28] M. E. J. Newman. Finding Community Structure in Networks Using the Eigenvectors of Matrices. *Physical Review E*, 74:036104, 2006.

[29] S. C. North. Incremental Layout in DynaDAG. In *Graph Drawing*, volume 1027 of *LNCS*, pages 409–418. Springer, Heidelberg, 1996.

[30] H. C. Purchase, E. Hoggan, and C. Görg. How Important Is the "Mental Map"? —An Empirical Investigation of a Dynamic Graph Layout Algorithm. In *Graph Drawing*, volume 4372 of *LNCS*, pages 184–195. Springer, Heidelberg, 2007.

[31] H. C. Purchase and A. Samra. Extremes are better: Investigating mental map preservation in dynamic graphs. In *Diagrammatic Representation and Inference*, volume 5223 of *LNCS*, pages 60–73. Springer, Heidelberg, 2008.

[32] I. Radu, E. Southgate, F. Ortega, and S. Smith. Summary: IEEE Virtual Reality Workshop on K-12 Embodied Learning through Virtual & Augmented Reality (KELVAR). In *Virtual Reality Workshop on K-12 Embodied Learning through Virtual & Augmented Reality (KELVAR 2017)*, pages 1–2. IEEE, 2017.

[33] W. Usher, P. Klacansky, F. Federer, P.-T. Bremer, A. Knoll, J. Yarch, A. Angelucci, and V. Pascucci. A Virtual Reality Visualization Tool for Neuron Tracing. *IEEE Transactions on visualization and Computer Graphics*, 24(1):994–1003, 2018.

[34] A. Van Dam, A. S. Forsberg, D. H. Laidlaw, J. J. LaViola, and R. M. Simpson. Immersive VR for Scientific Visualization: A Progress Report. *IEEE Computer Graphics and Applications*, 20(6):26–52, 2000.

[35] D. J. Watts and S. H. Strogatz. Collective Dynamics of 'Small-World' Networks. *Nature*, 393:440–442, 1998.

[36] C. D. Wickens. Virtual Reality and Education. In *Proc. IEEE International Conference on Systems, Man, and Cybernetics*, pages 842–847, Oct 1992.

[37] W. Zachary. An Information Flow Model for Conflict and Fission in Small Groups. *Journal of Anthropological Research*, 33:452–473, 1977.

[38] M. Zyda. From Visual Simulation to Virtual Reality to Games. *Computer*, 38(9):25–32, 2005.