Uncovering Best Practices in Immersive Space to Think

Kylie Davidson* Virginia Tech
Lee Lisle† Virginia Tech
Ibrahim A. Tahmid‡ Virginia Tech
Kirsten Whitley US Department of Defense
Chris North§ Virginia Tech
Doug A. Bowman¶ Virginia Tech

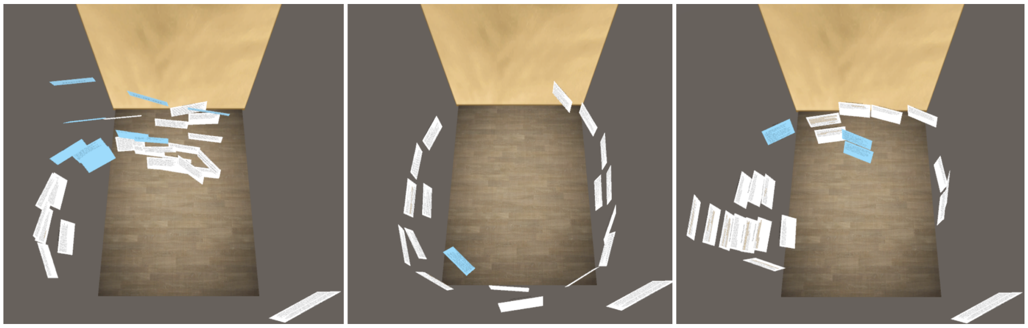

Figure 1: A top down view into the final spatial layout of a participant. From left to right, session 1, session 2, and session 3. We can see distinct differences in the use of space and deeper organizational features used. The document located in the bottom right corner is the analysis prompt.

ABSTRACT

As immersive analytics research becomes more popular, user studies have been aimed at evaluating the strategies and layouts of users' sensemaking during a single focused analysis task. However, approaches to sensemaking strategies and layouts are likely to change as users become more familiar/proficient with the immersive analytics tool. In our work, we build upon an existing immersive analytics approach–Immersive Space to Think–to understand how schemas and strategies for sensemaking change across multiple analysis tasks. We conducted a user study with 14 participants who completed three different sensemaking tasks during three separate sessions. We found significant differences in the use of space and strategies for sensemaking across these sessions and correlations between participants' strategies and the quality of their sensemaking. Using these findings, we propose guidelines for effective analysis approaches within immersive analytics systems for document-based sensemaking.

Index Terms: Human-centered computing—Visualization—Human computer interaction (HCI); Human-centered computing—Visualization—Virtual reality Human-centered computing—Visualization—Information visualization Human-centered computing—Visualization—Sensemaking

1 INTRODUCTION

Immersive analytics is an emerging research field that combines visual analytics, data visualization, and virtual/augmented reality. It focuses on using immersive human-computer interfaces to support analytic reasoning and knowledge generation (sensemaking) processes [10, 40]. Sensemaking involves extracting information, generating hypotheses, and formulating findings to make sense of the world around us. It is crucial in various domains like journalism [11], research, and intelligence analysis [36].

Visual analytics research emerged to assist sensemaking by offering interactive visual interfaces [44]. As immersive technology becomes more accessible, immersive analytics systems have been developed to enhance sensemaking. While previous work has focused on understanding how immersive analytics systems can support sensemaking using qualitative or quantitative data, it fails to explore the persistence of initial strategies and spatial layouts across multiple sessions. Furthermore, it overlooks strategies' evolution and refinement as users become more familiar with the tool and its effective usage. Therefore, there is a gap in understanding how learning to use immersive analytics systems influences overall analysis strategies and whether refined strategies lead to more effective sensemaking. This work aims to fill this gap.

We conducted a multi-session user study where participants completed three separate sensemaking tasks. The goal of this study was to provide users the ability to try different strategies for sensemaking as they became more familiar with the tool, as well as become more comfortable completing analysis in an immersive environment, in order to more deeply understand the evolution of their strategies and the effects of strategy change on task performance. We found that our participants used different spatial layouts over the different sensemaking sessions and utilized different strategies across the sessions. We also found correlations between certain patterns of user interaction and the quality of sensemaking analysis and that the overall organization of the sensemaking schemas increased over time.

The key contributions of this work are as follows. 1) An understanding of how learning and familiarity changes the sensemaking process in immersive analytics systems and how that influences the schemas and strategies of users. 2) Guidance on using immersive analytics systems for more effective analysis. 3) New data analysis methodologies for understanding sensemaking in immersive analytics prototypes.

*e-mail: kyliedavidson@vt.edu
†e-mail: llisle@vt.edu
‡e-mail: iatahmid@vt.edu
§e-mail: north@cs.vt.edu
¶e-mail: dbowman@vt.edu

2 Related Work

2.1 Sensemaking

Sensemaking is a complex process that involves taking various streams of unstructured data and forming hypotheses and theories. Pirolli & Card defined the sensemaking loop for intelligence analysts as having two sub-loops, the **foraging loop** and the **sensemaking loop** [36]. In the foraging loop, analysts gather data into a "Shoebox" and begin to triage the data for relevance. In this loop, they "Search & Filter" information and "Read & Extract" essential elements of information. In the sensemaking loop, analysts form schemas or structures within the data and develop hypotheses. During this stage, analysts "Schematize," "Build cases," and "Search for Relations/Evidence." The sensemaking loop ends with a presentation stage where they formalize the findings into a presentation such as an intelligence briefing or a written report. The knowledge generation process is bi-directional and iterative, where analysts work up and down the stages of the process as needed while completing sensemaking.

The sensemaking task is difficult, complex, and time-consuming, especially when moving towards the presentation stage of the sensemaking task [36]. Significant research has been devoted to assisting with both sub-loops of the sensemaking task [8, 15, 41]. Visual and immersive analytics are two approaches to assisting with sensemaking, and we review them in the following section.

2.2 Visual and Immersive Analytics

Visual analytics (VA) is the process of knowledge generation through interactive visual interfaces [44]. Going beyond information visualization, these interfaces should provide the ability to synthesize large amounts of information, detect expected or unexpected pieces of information, and provide timely insight into data [23] through human-in-the-loop interactions [37]. Overall, these systems should provide a deeper understanding during the sensemaking process. Some VA systems include StarSpire, JigSAW, and Space to Think [1, 7, 41], which have all been developed in an attempt to aid in this knowledge generation process.

With virtual and augmented reality technologies becoming more readily available, Immersive Analytics (IA) has built on the ideas of VA to provide new opportunities to support the knowledge generation process. IA is defined as the science of using immersive human-computer interfaces for analytic reasoning [10]. These interfaces should be designed to assist in the knowledge generation process through the use of abstract data visualizations while providing the benefits of embodied interaction [14, 40]. Immersive technologies can provide expansive space, depth cues, limited distractions, increased information bandwidth, increased engagement, and better spatial orientation [2, 6, 32]. Together these benefits from immersive technologies provide new opportunities to assist in the sensemaking process.

One approach to IA systems is providing quantitative data visualization opportunities. One example is ImAxes [12], which provided embodied interaction for creating visualizations within IA systems. Using IA technologies for quantitative data exploration and analysis provides new opportunities for natural user interactions with data [12, 46].

Another approach to IA systems is providing users with "space to think" [1] during analytic tasks through large tracked areas for offloading cognition into the environment around them. There has been much work in this area spanning exploration of how participants use these types of spaces [25, 26], understanding the high- and low-level organization of these spaces [13, 29, 30, 38], cognitive load within these systems [17], and techniques for assisting users with organizing content [42]. While this research area is growing, there has been limited work on understanding how these systems may be used long-term or by professionals. Batch et al. [5] conducted a study that studied sensemaking by professional users with the ImAxes system providing insights into how IA system usage changes over extended periods.

Additionally, Davidson et al. explored multi-session sensemaking in a system called Immersive Space to Think [13]. In this study, novice users were tasked to complete a single sensemaking task over the course of three separate sessions. They found that changes to strategies and spatial layouts in this immersive space occurred over time. While this study provided many insights into a more realistic single-sensemaking task, this study did not allow participants to learn from previous sessions when beginning a new sensemaking task. Our work aims to fill this gap by evaluating how learning familiarity and cognitive biases affect IA tool usage over multiple uses and how more refined usage strategies might improve the effectiveness of sensemaking with these tools.

2.3 Immersive Analytics System Usage and Learning

Immersive analytics systems provide many unique opportunities for data exploration, sensemaking, and decision making. However, in most traditional IA user studies, participants are trained on how to use a system and then are asked to complete a sensemaking task with that system [17, 25, 29]. While this provides initial knowledge on system usage, it does not allow participants to become familiar with the tools, try multiple strategies, and become comfortable completing sensemaking with them. As Norman pointed out, natural user interfaces are not always intuitive, and users may struggle with these interactions at first [34].

Additionally, biases such as anchoring bias [21] and mere exposure bias [47] may provide challenges for users as they may feel more comfortable sticking to the first pieces of information presented to them or have a tendency to prefer familiar tools/strategies. When users are presented with an IA tool for the first time, they may struggle with interacting and understanding the new opportunities that these tools present. In this work, we aim to understand how learning and familiarity affect sensemaking. By allowing participants to complete three separate sensemaking tasks on three separate days to become familiar with an IA prototype, they could try multiple strategies and schemas and use different features within the prototype during their sensemaking task.

3 Immersive Space to Think Approach

The immersive analytics prototype used in this work builds upon the existing Immersive Space to Think (IST) approach [4, 13, 25, 26]. The IST approach uses a VR-based system, which provides a large-tracked area for completing a text-based sensemaking task. A top-down view of our prototype can be seen in Figure 1 with an example of how the documents can be organized during a sensemaking task. The text documents for sensemaking are loaded in a randomized order onto a virtual bulletin board. The documents size is 0.5x0.3 (m) which supports about one paragraph of text without the need for scrolling. During the sensemaking task, users can remove the documents from this bulletin board and place them anywhere within the 3D tracked space during their analysis. To support sensemaking, IST also provides users with a set of features to aid in their analysis. Our IST prototype supported *highlighting, searching, note taking, label making, and copying/pasting to/from clipboard.* For text entry, a tracked keyboard on top of a wheeled table was provided to our users.

Based on feedback from previous studies using IST, we added one additional feature was added to the IST prototype used in this study: **Quick Search**. This new feature enabled the ability to select a word directly in a document and search for it in the rest of the documents using a virtual button displayed on the side of the document. We still supported general search based on text entry, but the new Quick Search feature allowed our users to quickly see other instances of a word in the whole document set without needing to type the word manually.

4 USER STUDY

4.1 Goals and Research Questions

The goal of this study was to understand how the learning effect of becoming more familiar with an immersive analytics tool affects the spatial layouts and strategies used during sensemaking and how the refined strategies might lead to more effective sensemaking performance.

Our research questions were as follows:

- RQ1: Spatial Layouts
  - RQ1A: What spatial layouts do users form in IST when performing sensemaking tasks?
  - RQ1B: How do these change across multiple sensemaking sessions? (learning effect)
- RQ2: Strategies
  - RQ2A: What strategies do analysts use in IST during sensemaking?
  - RQ2B: How do these change across multiple sensemaking sessions? (learning effect)
- RQ3: How do the spatial layouts, strategies, and interaction patterns in IST correlate to the quality of the sensemaking task?

4.2 Apparatus

For this study, we used an HTC VIVE Pro [1] VR head-worn display with a wireless adapter to allow for a tether-free VR experience. The HTC VIVE Pro has 1440 x 1600 pixels per eye (2880 x 1600 pixels combined), a refresh rate of 90 Hz, and 110 degrees field of view. A Steam VR 2.0 Lighthouse tracking system was used. The tracked area for the study was a 3x3 meter area in which the participants could freely walk around while completing their analysis. The participants were trained on the size of the space during the onboarding phase of the study, and the floor in the virtual space also represented the boundary of the tracked area. In addition, the experimenter watched to ensure that participants did not get too close to the physical room boundaries. For interacting with the system, participants used a Valve Index controller [2] in their dominant hand, and for text entry, a wheeled table with a Vive tracker [3] and wireless keyboard with number pad included were provided for the participant to move freely around the space as needed for their analysis. To assist the participants in typing, we cut out some of the foam padding from the bottom of the headset, which allowed participants to peek out of the headset to see their fingers on the keyboard. The IST system was developed and rendered using Unity Game Engine [4].

4.3 Datasets

We selected three separate datasets for sensemaking tasks. These datasets were presented to the participants in a randomized order to prevent unwanted effects of dataset order on the results of our study. The details of each dataset can be seen outlined below:

Sign of the Crescent is a fictional intelligence analysis training dataset comprising 40 text-based documents for analysis. This dataset has been used in previous studies [13, 24, 45]. In this study, we selected 20 documents from the original text corpus, with one main storyline, two locations of interest, and a few distractor documents. The documents were about a paragraph in length on average.

Manpads is a fictional intelligence analysis training dataset comprising 23 text-based documents for analysis. In this study, we selected 20 documents from the original text corpus with one main storyline, four locations of interest, and two distractor documents. The documents were about a paragraph in length on average. This dataset has been used in previous work [45] and had a similar format to that of Sign of the Crescent.

Stegosaurus is a 2006 VAST challenge dataset [19], which consists of 240 documents which are a mixture of news articles, maps, spreadsheets, and some other supporting documents. However, only ten documents have relevance to the ground truth. This dataset has been used in visual analytics sensemaking tasks before [20, 48], so we believed it would be a good match for our participants to complete their analysis. In this study, we selected 20 documents from the original text corpus based on the VAST challenge solution. These documents had one main storyline for analysis with two main locations of interest. The documents were about a couple of paragraphs in length on average.

4.4 Experimental Tasks

Participants were given a new dataset and task to complete sensemaking on each session. The tasks were assigned in a randomized order and each day the task was loaded into the immersive space as seen in figure 1. The three task prompts were as follows:

Sign of the Crescent "From this set of reports, generate a hypothesis about any action(s) terrorists are planning in the near future. The near future means after April 27, 2003, which is the date of the final report you have received. Construct defensible arguments supporting why your hypothesis should be taken more seriously than others that are possible. You must be prepared to issue a warning notice to the new Office of Homeland Security and any other interested offices. At the end of today's session, you will be asked to generate an outline of your findings."

Manpads "From this set of reports, generate a hypothesis about any action(s) terrorists are planning in the near future. The near future means after November 12, 2002. Construct defensible arguments supporting why your hypothesis should be taken more seriously than others that are possible. Both the DOD and the Department of Homeland Security (DHS) are vitally interested in your reporting immediately on any terrorist threats you believe to be both serious and imminent. At the end of today's session, you will be asked to generate an outline of your findings."

Stegosaurus "Welcome to Alderwood, Washington, a fictitious city in central Washington State. Due to some recent revelations of corruption and conspiracy, the people of Alderwood are left wondering 'What next?' The head of the Alderwood Police Department has a hint towards the answer to that question. An incident from early 2004 (see article IntoxMan.txt) has been nagging him since it happened. Now, with the FBI's attention focused on Alderwood, he has decided to put together a special task force to look into the situation. When he handed you the report he said, 'This incident still doesn't sit right with me. Despite initial reports, no alcohol or drugs were found in the man's blood. That stuff he was spouting sounded like something you'd hear from a cult or something. If not for all the recent goings-on, I might think I was being paranoid. I want this looked into.' Your task is to identify evidence and generate a hypothesis about potentially suspicious activity in the town of Alderwood. At the end of today's session, you will be asked to generate an outline of your findings."

During the outlining stage, participants were encouraged to answer the questions **who, what, when, where, and how** about the threats or suspicious activities; they were also instructed to include any other details they felt were relevant to their hypothesis. Outlines were generated by the participants within an IST note artifact.

4.5 Participants

We recruited 14 professional intelligence analysts as participants. The analysts had an average of 12.2 years of professional analysis

[1] www.vive.com/eu/product/vive-pro/
[2] valvesoftware.com/en/index/controllers
[3] business.vive.com/eu/product/vive-tracker
[4] unity.com

experience with a standard deviation of 8.23 years (median of 9.5 Years). The age range of our participants was representative of the United States working-class population. We had seven male and seven female participants with the following VR/AR experience levels: Never (8), Once or Twice (4), 3-10 Times (2), and 10+ Times (0). All participants had normal or corrected vision (glasses or contacts). The participants were recruited using word of mouth and email listservs. This study was approved by the Institutional Review Board at the authors' university.

4.6 Measures

We administered a pre-study questionnaire that collected background information on the participants, such as analysis experience and VR experience. During the user study task, we collected both log files and save files. The log files captured all participant actions within our prototype, including participant and keyboard movement once a second and interaction events (new Note, Search, etc.). The save files captured where all IST artifacts (i.e., documents, notes, labels) were located (x, y, z, yaw, pitch, and roll orientation) once per minute. After each analysis session, we conducted a semi-structured interview with the participants about their analysis strategy and document organization for the session. After the final user study task, we completed a post-study interview to gather insight from the participants on strategies they perceived as most effective, layouts, and suggestions for future features. Finally, we collected a first-person screen recording of the participants' viewpoints while completing their analysis tasks.

4.7 Procedure

The user study was broken into three 75-minute sessions. Participants were instructed to schedule their sessions on three separate and preferably consecutive days. Session one began with the formal consent process and the pre-study questionnaire. Next, participants were taught how to use the IST prototype during a ten-minute training phase. Participants were provided a training dataset (distinct from the three datasets used in the main study sessions) and instructed on using all system features. The participants were given five minutes to practice using the features and manipulating the documents within the system.

After the training phase of session 1, and after arrival and quick system review for sessions 2 and 3, participants began their standardized 60-minute analysis task for the day. Participants began the analysis portion of the user study by reading the experimental task prompt (see Section 4.4) for the day, and they were informed that they would be expected to produce an outline of their findings using a note within the IST system by the end of the session. Time warnings were given at fifteen minutes, five minutes, and one minute remaining during the analysis portion.

Each session ended with a semi-structured interview while participants remained immersed in the IST system with their analysis documents surrounding them. At the end of session three, participants took part in a post-study interview conducted outside of VR to gain insight into strategies and structures used and collect feedback on the system. Participants were provided an opportunity upon completion of the study to ask any questions of the researchers and provided with contact information if questions or concerns arose.

5 Results and Discussion

To address RQ1, we looked at the overall spatial structures formed by participants, how these structures differed from session to session and studied the deeper organizational features used for analysis. To address RQ2, we examined the videos, interview files, and log-file interactions to understand the strategies used across the sessions. To address RQ3, we evaluated the outlines generated by the participants and made correlations between schemas, strategies, and system-level interactions to understand the sensemaking quality.

	Semi-Cylindrical	Cylindrical	Environmental
Session 1	7	1	6
Session 2	9	2	3
Session 3	5	5	4

Table 1: Final layout categorizations for the end of each sensemaking session

5.1 RQ1: Spatial Layouts and Deeper Organization

5.1.1 Categorizations of Spatial Layouts

Previous studies of the IST approach [25] categorized the types of layouts formed within immersive analytics sensemaking spaces into three types: *Semicirular*, Environmental, and Planar. Semicircular has been previously used for layouts in which documents are placed curved around the user. Environmental layouts use the physical boundaries of the space (i.e., floors/bulletin board) for organizing documents in a plane, and planar layouts use planes that exist throughout the space that are not curved around the user. We propose two small changes to the previous layout categorizations in this work. 1) To reflect its three-dimensional nature, we propose that semicircular be renamed semi-cylindrical. 2) We add a distinction between semi-cylindrical and cylindrical as supported by recent studies [27, 28, 30, 39], as well by comments from some of the participants. In our new definitions, we define semi-cylindrical layouts as those in which documents are curved around the user but not on all sides, while cylindrical layouts position documents in the full 360-degree space surrounding the user. Our definitions for environmental and planar remain the same. Figure 2 shows an example of each type of layout.

At the end of each session, a high-level structure of the space was evaluated using a top-down view such as that seen in Figure 1. We found 21 semi-cylindrical layouts, with 50% in session 1, 64.3% in session 2, and 35.7% in session 3. We had eight cylindrical layouts, with 7.1%, 4.3%, and 35.7% in the three sessions, respectively. We found 13 environmental layouts, with 42.9%, 21.4%, and 28.6% across the three sessions. In total, 50% of the layouts were semi-cylindrical, 19.05% were cylindrical, and 30.95% were environmental, as seen in Table 1. Using a chi-squared test of independence, we did not find a statistically significant difference in spatial layouts used across sessions or datasets.

5.1.2 Spatial Layout Change Over Time

Of our 14 participants, we had eight (57.1%) use at least one different high-level structure across their sensemaking sessions (i.e., P11: $\rightarrow$ $\rightarrow$). This suggests that the participants tried different layouts during their sensemaking as they learned more from using the system for previous analysis tasks. The other six (42.9%) participants used the same high-level structure across all three sensemaking sessions (i.e., P14: $\rightarrow$ $\rightarrow$). Although these participants did not use the space differently across the sessions based on this high-level classification, we did find differences based on deeper organizational features of these spatial layouts, as discussed in Section 5.1.3.

Self-Reported Change Across Sensemaking Sessions

Going beyond the high-level classification of the spatial layouts, during the post-session interviews, participants were asked to comment on their overall spatial structure and if the structures they used during sessions two and three differed from those they used in the previous sessions. The goal was to understand how the participants felt their use of space changed across the sessions. We had ten (71.4%) participants who reported a change in spatial use from session 1 $\rightarrow$ 2 and five (35.7%) participants who reported a change in

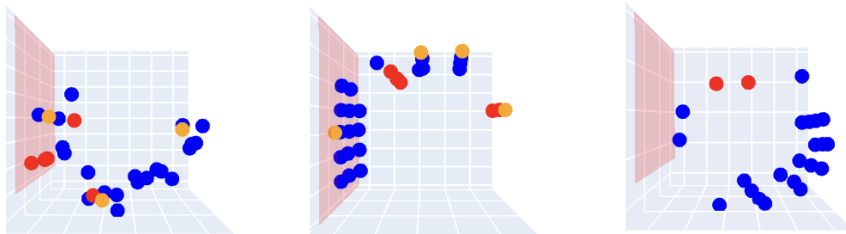

Figure 2: From left to right, Semi-Cylindrical, Environmental, and Cylindrical layouts created by P13, P7, and P3, respectively. The red mesh represents the virtual bulletin board, red dots represent notes, yellow dots represent labels, and blue dots represent the documents used during analysis.

spatial use from session 2 → 3. The changes in perceived space usage can be illustrated by participant quotes:

> *"Much more organized ... Yesterday [session 1] I didn't really have an organization. However, I think it helped me see everything in a way that made sense"* -P4 Session 2 interview

> *"I think on last time I was sort of throwing things wherever to get them out of the way, and then today I took advantage of having the space and using it to organize this time"* -P6 Session 2 interview

> *"My use of the spatial layout continues to get better as I get more familiar with ... using the tools here."* - P14 Session 3 Interview

Angular Space Usage Change over Time

In our high-level classification of the spatial layouts, we saw a shift in more participants using cylindrical layouts across the sessions from 1 → 2 → 3. Additionally, in looking at all layouts, we saw a shift from primarily 180-degree document placement in session 1 to what appears to be more angular space used for document placement in sessions 2 and 3. To take a deeper look at this shift, we analyzed the total angular space used by the participants across the sessions.

To examine the total angular space used for document placement, we: 1) determined the location each participant spent the most time at during a session, 2) translated all document locations to be relative to that location, and 3) calculated the spherical coordinates of each document relative to the center. This allowed us to generate BarPolar charts as seen in Figure 3, which represent the total count of documents located in each 15-degree angular bin across the three sensemaking sessions for all participants. The figure shows that in session 1, participants had a strong tendency to place documents in the front 180 degrees of the space. Then as we look at sessions 2 and 3, we see a shift towards using more of the full 360 degrees of the space for document placement. In addition to looking at the angular use of space across sessions, we also looked at the angular use across the three different datasets, but we did not find any distinct differences between datasets for this measure.

This increase in the use of angular space could have occurred for multiple reasons. First, as the participants become more familiar with immersive analysis, they may have been more comfortable in trying to use "space to think" during their analysis. Second, based on the interviews, participants came in with new ideas for how they wanted to use the space after their first analysis session, suggesting that they learned from their first session. Another possibility is that participants may have become more comfortable physically navigating and maneuvering within the VR system after an initial session with the system.

5.1.3 Deeper Organization

Clustering Change Over Time

In order to evaluate how participants grouped related documents in space over time, we looked at multiple clustering techniques such as K-Means [31], Hierarchical clustering [33], DBSCAN [16], and OPTICS [3]. Based on the clustering results of running these different techniques, we selected OPTICS (Ordering Points to Identify the Clustering Structure) [3] for use in our clustering analysis since it visually fits the ground truth of the data best. Additionally, OPTICS uses density-based clustering and detects noise (documents that are located too far away from a cluster to be categorized within that cluster). Figure 4 shows an example of OPTICS clustering. The goal of this analysis was to identify an overall clustering of each of the final layouts from each session so that clustering evaluation methods such as the Calinski Harabasz (CH) [9] score of the layouts could be used to evaluate the overall organization of the space. The CH score evaluates the goodness of clustering, which is the ratio of the sum of between-cluster variance and within-cluster variance, also known as the variance ratio criterion.

We hypothesized that the CH scores based on the OPTICS clustering would increase over time as the participants became more familiar with analysis in IST. This would indicate well-defined clusters with a large between-cluster variance and a small within-cluster variance, meaning that documents within the same cluster are close together and far from other document clusters.

We ran OPTICS on each final layout using the following parameters (min_samples=2, metric=l1, and all other default parameters as seen in [35]). We ran a full-factorial analysis of variance (ANOVA) to evaluate the effects of the session, dataset, and the session*dataset interaction on the CH-score from our OPTICS clustering. We found a main effect of the session on the CH score ($F(2,33)$, $p = 0.0114$). Tukey HSD tests for pairwise comparisons found that session 3 had a significantly higher CH score than session 1 ($p = 0.0087$). There was no significant difference between sessions 2 and 3 ($p = 0.4038$) or between sessions 1 and 2 ($p = 0.1399$). The average CH scores were 6.866, 14.882, and 20.186 for sessions 1, 2, and 3, respectively. We found no significant effects of dataset or session*dataset on the CH-Score.

In addition to changes in clustering, there were changes in the use of other organizational strategies over the sensemaking sessions. In previous studies using IST, many participants formed vertical *columns* of documents to encode timelines and form clusters of information. An example of a column can be seen on the left side of the session 3 layout in Figure 1. Based on visual inspection in this study, 9/14 participants used columns in their spatial layouts during session one, while all 14 used columns during session 3.

Some participants also organized their notes into a *scratch space*, an area used only for notes that were separate from the space used to organize documents. Based on visual inspection, we found scratch

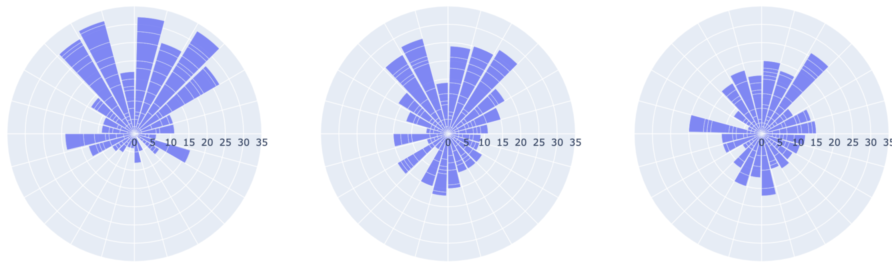

Figure 3: From left to right, session 1, session 2, and session 3 angular placement of documents. The direction towards the bulletin board is at the top of the circle. The space is broken into 15-degree bins, and each bar represents the total count of documents across all participants in that 15-degree slice of the space.

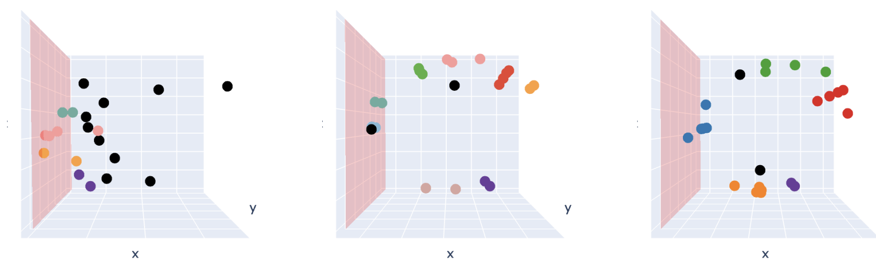

Figure 4: From left to right, session 1, session 2, and session 3 clustering using OPTICS for P11's final spatial layout. The red plane represents the bulletin board, the black points indicate documents not assigned to a cluster ("noise"), and the colored points indicate the clusters to which the documents were assigned.

space usage by one participant in session 1, two in session 2, and two in session 3. While there is not a trend with scratch space usage over time, we still highlight it, since this approach to spatial organization has not been reported in previous research related to IST.

The following interview quote highlights how participants were able to use deeper organization within their spatial layouts.

"I am actually very proud of this spatial layout. I have it all labeled on top, and I have them all in chronological order but in a way that is easily visible at a quick glance. And I have it broken out by different reporting entity as well." - P7 Session 2 Interview

Overall, we found that participants utilized the space provided within IST differently as they became more familiar with the tool over multiple sessions. We saw more cylindrical layouts in later sessions, which could be due to comfort levels within the VR system or a willingness to explore different spatial layouts. Most participants also reported that their use of space changed in the later sessions. We also saw more participants using the full 360 degrees around them for document layout as we looked across the different sensemaking sessions and found that the clustering score increased over time. Finally, deeper organizational features were used more frequently during the later analysis sessions. Together, these changes indicate that participants were learning and adapting to the use of IST as they progressed through the sensemaking sessions, with later sessions tending to exhibit more significant space usage and more organized layouts. We saw the greatest change in the use of space between sessions 1 and 2, with mostly smaller changes between sessions 2 and 3.

5.2 RQ2: Strategies for Analysis

As approaches to sensemaking are very personal, we expected to see different strategies for each participant in the study. Additionally, we expected to see individuals make changes to their strategy as a result of learning from previous analysis sessions what did not work well and what did. In this analysis, we wanted to look at the participants' strategies from a high-level sensemaking process point-of-view to categorize the overarching strategies during the analysis sessions. Using the interview files and notes from observations of the analysis sessions, we identified five high-level strategies used in our study.

Strategy 1: Read all documents first, then sort the documents into themes. After initial sorting, continue to refine the clusters based on connections between the documents. (Read, Sort, Refine)

Strategy 2: Skim the document set for themes, then begin to sort the documents into groups. After initial sorting, read documents in detail while performing a second sorting. Make refinements to the groupings as needed. "Skim" in this analysis refers to quickly reading over a single document to get a general idea of its contents. (Skim, Sort, Read, Refine)

Strategy 3: Skim documents for themes, then begin searching on recurring themes. Read the documents from the search and, using these results, cluster documents. Use the information from the search to guide the next search. Refine groupings as needed. (Skim, Search, Read, Sort, Refine)

Strategy 4: Sort documents into groups based on a quick scan, then read the documents in detail. While reading, sort the documents into smaller groupings as needed. Refine the space. "Scan" in this analysis refers to a very high-level view of all documents to understand their structure without skimming though them for general details like one does during skimming. (Sort, Read, Sort, Refine)

Strategy 5: Read the documents first, then search for recurring themes. Sort the results from the search and then refine the groupings as needed. (Read, Search, Sort, Refine)

Some of the main differences among these strategies relate to when the "Read & Extract," "Schematize," and "Search for Evi-

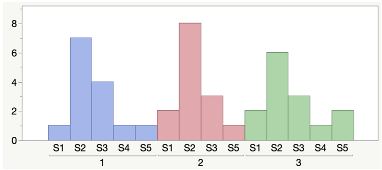

Figure 5: Total count of strategy usage across the sessions. Blue is session 1, red is session 2, and green is session 3.

dence" stages of the process occur. For instance, in strategy 1, users read the documents first, sorted them in space, made minor refinements to their organization, and then presented their findings in the outline. This strategy is a very linear sensemaking process compared to strategy 2, where participants started by skimming the documents, then sorting the documents into rough groupings. Then the documents were read in detail and sorted again through a refinement stage while the participants prepared their hypotheses. This approach involved more iterations up and down the sensemaking loop. Strategy 3 and Strategy 4 were also more iterative high-level strategies, while Strategy 5 was more linear. Our study's strategies are similar to those defined by Kang et al. [22], with our S1 mapping well to "Build from Detail", S2 mapping to "Overview, Filter, and Details", S3 mapping to "Hit the Keywords", S4 mapping to "Overview, Filter, and Details" with an initial triage stage at the beginning, and lastly, S5 mapping to "Find the clues, follow the trail."

Figure 5 shows the counts of strategy uses per session in our study, with Strategy 2 being the most popular strategy in all sessions. Using a Chi-Squared test of independence, we found no statistically significant difference in using different strategies across the sessions or the different datasets. In order to evaluate interaction-level differences among the strategies, we ran Student's T-Tests on system interaction data collected in the log files. We found that participants using S2 walked farther than those in S1 ($p = 0.0259$), with an average of 151.28 meters compared to 101.559 meters of total movement. Participants using S4, with an average of 218.35 meters, also walked more than those in S1 ($p = 0.0046$) and S3 ($p = 0.0374$), which had an average of 124.57 meters of total movement. We found that S3 ($p = 0.0278$) and S5 ($p = 0.0311$) users executed more searches than those who used S1. Lastly, we found that S1 users moved documents more often than S2 ($p = .0029$), S3 ($p = 0.103$), and S5 ($p = 0.0463$). One interpretation of these findings is that the participants who used S1 spent more time reading and moving documents around than using features to help them make connections, such as the search feature or physically navigating the space to help them understand the overall dataset. Another interpretation of these findings is that since S3 and S5 executed more searches than S1, they better understood where the documents connected on a thematic level and therefore needed less time moving documents to create meaningful organization. Additionally, as S2 and S4 navigated more during their analysis than S1, this could suggest that S2 and S4 used physical navigation via walking around the space to aid in referencing documents during their analysis instead of using the controller to bring documents to them for examination.

During the post-session interviews, participants were asked to comment on their overall strategy and if their strategies during sessions 2 and 3 differed from the ones they had used in the previous sessions. Ten participants reported a change in strategy from session $1 \rightarrow 2$, and six reported a change in strategy from session $2 \rightarrow 3$. We found that the majority of participants self-reported that their strategies changed after session 1. These self-reported changes capture high-level perceived strategy differences (e.g., reading all documents first $\rightarrow$ sorting documents first) and system-level interaction changes in strategy (e.g., notes, labels, and search usage). Overall, participants utilized different high-level strategies for analysis as they became more familiar with using IST, as we expected. While these changes in strategy may also show a learning effect, we believe that these changes provide insights into how IA systems, such as IST, would be used in a real-world setting by providing the participants time to learn how the system could truly benefit them during their analysis process. A few quotes highlighting strategy changes from the participant interviews can be found below.

"I don't think I really had a strategy the first time ... I kept jumping around on which piece of information I wanted to group by, and this time it seemed to make more sense to make more of a timeline strategy." - P6 Session 2 interview

"Yesterday, I read every report before I did the buckets, and I kind of saw that that was a waste. So today I skimmed them more to kind of get the ideas together ... I felt more comfortable with the controls and I don't think I reread as much today." - P8 Session 2 Interview

5.3 RQ3: Correlations with Sensemaking Quality

To this point, we have presented spatial layouts and high-level strategies used during analysis. However, we also aimed to understand how layouts, strategies, or system-level interactions correlate with the quality of the sensemaking task.

In order to evaluate sensemaking quality, we scored the outlines generated by participants based on the ground truth of the dataset. The rubrics awarded points for every correct piece of information presented and specifically focused on the **Who, What, When, Where, and How** questions the participants were guided to include in their outlines. Where a point was awarded for every correct piece of information that was reported on. The primary experimenter scored the outlines in a randomized order across all three datasets. Each rubric had a variable number of points that could be received (Crescent - 25, Manpads - 26, Stegosaurus - 31), so we report normalized scores where 1.0 represents a perfect score.

We ran a full-factorial ANOVA to evaluate the effects of Session, Dataset, and the Session*Dataset interaction on the outline scores. We found no significant effect of Session or Session*Dataset; however, we did find a main effect of the dataset on the score ($F(2,33)$, $p = 0.0004$). Using an LSMeans Differences Tukey's HSD Test, we found that Sign of the Crescent ($p = 0.0020$) and Manpads ($p = 0.0007$) scored higher than Stegosaurus. Due to the apparent difficulty of the Stegosaurus dataset, we removed it from the following correlation analyses.

We also ran ANOVAs to evaluate the effects of Session and Dataset on all other variables listed below (spatial layouts, strategies, and interactions). We did not find any effects of Session, Dataset, or Session*Dataset on any of these variables, and therefore suggest that the correlations presented below are not influenced by dataset or familiarity. Thus, we can have some confidence in a general relationship between sensemaking quality and specific layouts, strategies, and interaction patterns.

5.3.1 Spatial Layouts and Quality

The mean score for cylindrical was 0.5535 (sd = 0.2162), environmental was 0.5689 (sd = 0.1320), and semi-cylindrical was 0.5394 (sd = 0.2646). We used a Student's T-Test (alpha = 0.05) and found no significant effect of the layouts on scores.

5.3.2 Strategies and Quality

The mean score and standard deviation for each strategy can be seen in table 2. We ran a Student's T-Test on strategies and scores, and we found that S2 scored higher than S1 ($p = 0.0084$) and S5 ($p = 0.0247$), and S4 scored higher than S1 ($p = 0.0372$). In combination with our analysis of the different strategies in RQ2, we believe that since S1 had less movement, fewer searches, and more document movements, perhaps the participants who used S1 spent too much time focused

	Average	Standard Deviation
S1	0.2250	0.0212
S2	0.6392	0.1537
S3	0.5261	0.2445
S4	0.7400	0.0000
S5	0.3483	0.3483

Table 2: Average score and standard deviation for each strategy

on reading the documents in detail and then sorting them one by one. Leading to the participants not iterating through their sensemaking process to explore connections between the document set.

5.3.3 Interactions and Quality

We looked at all log-file interactions (searches, grabs, labels, notes, highlights, text entered, and copy/paste) as well as calculated features (keyboard and participant movement, time spent interacting with documents, and distances between documents) in an attempt to understand how these correlate to the quality of the sensemaking.

We found two interesting correlations that we want to highlight here. We found a positive correlation (0.4596) between the score and the number of **new notes** created during analysis (p=0.0139). We wanted to understand the **"spread"** of documents across the provided space (i.e., the amount of space used to lay out the documents). To calculate the spread, we looked at the total distance between all documents in the space. Then we normalized by the total number of artifacts (docs, labels, and notes) within the participant's layout. We found a positive correlation (0.4703) between this factor (Sum Distances / Number of Artifacts) and the score (p=0.0116), showing that a layout with greater spread correlates with higher-quality sensemaking.

5.4 IST Analysis Best Practices

Based on the analysis conducted in this study and the findings presented above, we propose a set of best practices for effective analysis within IST-like systems. Our study suggests that these guidelines can help users perform effective sensemaking in IA systems.

Use a spatial organization schema that works best for the task and analyst preferences There are many ways to encode meaning in a system like IST. Having the availability to organize in 3D allows for new opportunities to organize sensemaking schemas. We saw no differences between spatial layouts in terms of the quality of sensemaking. Thus, we suggest that the overall spatial layout should match the document set and the analyst's preferred style.

Use a more iterative approach to sensemaking We found that participants who used S2 and S4 scored higher than participants who used S1. Additionally, S2 also scored higher than S5. Our analysis found that S2 and S4 were more iterative approaches to the sensemaking task, where the participants moved up and down the sensemaking process as needed to support their strategy. Based on this and user feedback in the final interview session, we believe frequent iterating through the sensemaking stages leads to better analysis both in and out of IST.

Use more space for document layout As participants became more familiar with the tool, they became more organized over time. Additionally, the participants who used more of the provided space (better spread between documents) scored better on these sensemaking tasks. We suggest that using more space to create meaningful organization benefits the overall sensemaking quality. In future iterations of IA systems, additional organization features such as semi-automated clustering [43] or semi-automatic column detection could provide scaffolding for users to create meaningful structures.

6 Limitations

One limitation of this work is the text-entry system. For this study, we used a tracked keyboard and table that participants could wheel around the space with them. While we removed some of the foam in the headset to assist the user in seeing their fingers on the keyboard, some participants still struggled with text entry, which could have led to less detailed outlines or fewer notes. In future work, a system like the portal proposed by Giovannelli et al. could be used [18]. Moreover, as a result of using the VIVE Pro headset, the text-readability when documents were placed across the 3x3 meter tracked space decreased due to some blurriness. While we don't believe this influenced our results, we suggest that future work use higher resolution displays to enhance text-readability from across the tracked space.

Another limitation of this work is that the Stegosaurus dataset differed from both Manpads and Crescent in document length and style, which could explain the difference in scores we observed, thus limiting our ability to evaluate the effects of session and other factors on scores.

Finally, we were limited by a small participant pool, which led to less power in our statistical analyses. We present all statistically significant findings with caution and provide directions for future research that should be verified in the future.

7 Conclusions and Future Work

This work explored how learning and familiarity influence sensemaking in immersive analytics. We explored three main research questions that focused on spatial structures, strategies, and correlations for performance. We found high and low-level spatial structure and strategy changes across the different sensemaking sessions. Together, these indicated that as the participants became more familiar with the tool, the participants adapted and learned how to analyze within our prototype IST. Lastly, our correlation analysis found that iterative sensemaking, note usage, and document spread were key indicators for effective sensemaking. Using what we learned in our exploratory analysis, we suggest guidelines for sensemaking in IA systems.

We also covered many types of analysis within this paper to explore users' sensemaking in IA systems. The analyses included were not limited to spatial structure categorization, deeper organization features, clustering analysis, degrees used, high-level, and strategy categorization. We believe these analyses provided us with many insights into our participant's sensemaking and highlight these novel analyses to be used in future work for IA sensemaking. One of the future directions of this research will be implementing new organization features intending to understand how these features affect the sensemaking task. Specifically looking at new features that would promote some of the best practices described above. To promote spatial organization schemas, we believe some automated layout features can be developed to assist the users in the overall organization. Additionally, additional organizational features can be developed to encourage more space usage to help users offload more sensemaking into the environment. Some of these features include the clustering of documents which provides the ability to move sets of documents together as a group. Another idea would be to provide assistive linking features to help analysts tie information within and across clusters. Allowing for the ability to have cross-space references when needed. We also plan to explore how the guidelines presented in the paper affect sensemaking in a controlled between-subjects users study investigating the effects of guidance vs. no guidance in the quality of sensemaking tasks.

Acknowledgments

This work was supported in part by NSF I/UCRC CNS-1822080 via the NSF Center for Space, High-performance, and Resilient Computing (SHREC).

References

[1] C. Andrews, A. Endert, and C. North. Space to think: large high-resolution displays for sensemaking. In *Proceedings of the SIGCHI*

conference on human factors in computing systems, pp. 55–64. ACM, 2010.

[2] C. Andrews, A. Endert, B. Yost, and C. North. Information visualization on large, high-resolution displays: Issues, challenges, and opportunities. *Information Visualization*, 10(4):341–355, 2011. doi: 10.1177/1473871611415997

[3] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander. Optics: Ordering points to identify the clustering structure. In *Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data*, SIGMOD '99, p. 49–60. Association for Computing Machinery, New York, NY, USA, 1999. doi: 10.1145/304182.304187

[4] P. Bandyopadhyay, L. Lisle, C. North, and D. A. Bowman. Immersive space to think: The role of 3d space for sensemaking. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, 2020.

[5] A. Batch, A. Cunningham, M. Cordeil, N. Elmqvist, T. Dwyer, B. H. Thomas, and K. Marriott. There is no spoon: Evaluating performance, space use, and presence with expert domain users in immersive analytics. *IEEE transactions on visualization and computer graphics*, 26(1):536–546, 2019.

[6] D. A. Bowman and R. P. McMahan. Virtual reality: How much immersion is enough? *Computer*, 40(7):36–43, 2007. doi: 10.1109/MC.2007.257

[7] L. Bradel, A. Endert, K. Koch, C. Andrews, and C. North. Large high resolution displays for co-located collaborative sensemaking: Display usage and territoriality. *International Journal of Human-Computer Studies*, 71(11):1078–1088, 2013.

[8] L. Bradel, C. North, L. House, and S. Leman. Multi-model semantic interaction for text analytics. In *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 163–172, 2014. doi: 10.1109/VAST.2014.7042492

[9] T. Caliński and J. Harabasz. A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods*, 3(1):1–27, 1974.

[10] T. Chandler, M. Cordeil, T. Czauderna, T. Dwyer, J. Glowacki, C. Goncu, M. Klapperstueck, K. Klein, K. Marriott, F. Schreiber, et al. Immersive analytics. In *2015 Big Data Visual Analytics (BDVA)*, pp. 1–8. IEEE, 2015.

[11] S. Cohen, J. T. Hamilton, and F. Turner. Computational journalism. *Communications of the ACM*, 54(10):66–71, 2011.

[12] M. Cordeil, A. Cunningham, T. Dwyer, B. H. Thomas, and K. Marriott. Imaxes: Immersive axes as embodied affordances for interactive multivariate data visualisation. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, pp. 71–83. ACM, 2017.

[13] K. Davidson, L. Lisle, K. Whitley, D. A. Bowman, and C. North. Exploring the evolution of sensemaking strategies in immersive space to think. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–15, 2022. doi: 10.1109/TVCG.2022.3207357

[14] T. Dwyer, K. Marriott, T. Isenberg, K. Klein, N. Riche, F. Schreiber, W. Stuerzlinger, and B. H. Thomas. Immersive analytics: An introduction. *Immersive analytics*, pp. 1–23, 2018.

[15] A. Endert, P. Fiaux, and C. North. Semantic interaction for sensemaking: Inferring analytical reasoning for model steering. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2879–2888, 2012. doi: 10.1109/TVCG.2012.260

[16] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, vol. 96, pp. 226–231, 1996.

[17] A. Galati, R. Schoppa, and A. Lu. Exploring the sensemaking process through interactions and fnirs in immersive visualization. *IEEE Transactions on Visualization and Computer Graphics*, 27(5):2714–2724, 2021. doi: 10.1109/TVCG.2021.3067693

[18] A. Giovannelli, L. Lisle, and D. A. Bowman. Exploring the impact of visual information on intermittent typing in virtual reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 8–17, 2022. doi: 10.1109/ISMAR55827.2022.00014

[19] G. Grinstein, T. O'Connell, S. Laskowski, C. Plaisant, J. Scholtz, and M. Whiting. Vast 2006 contest - a tale of alderwood. In *2006 IEEE Symposium On Visual Analytics Science And Technology*, pp. 215–216, 2006. doi: 10.1109/VAST.2006.261420

[20] P. Isenberg, D. Fisher, S. A. Paul, M. R. Morris, K. Inkpen, and M. Czerwinski. Co-located collaborative visual analytics around a tabletop display. *IEEE Transactions on visualization and Computer Graphics*, 18(5):689–702, 2011.

[21] D. Kahneman, S. P. Slovic, P. Slovic, and A. Tversky. *Judgment under uncertainty: Heuristics and biases*. Cambridge university press, 1982.

[22] Y.-a. Kang, C. Gorg, and J. Stasko. Evaluating visual analytics systems for investigative analysis: Deriving design principles from a case study. In *2009 IEEE Symposium on Visual Analytics Science and Technology*, pp. 139–146. IEEE, 2009.

[23] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon. Visual analytics: Definition, process, and challenges. In *Information visualization*, pp. 154–175. Springer, 2008.

[24] T. Li, K. Luther, and C. North. Crowdia: Solving mysteries with crowdsourced sensemaking. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–29, 2018.

[25] L. Lisle, X. Chen, J. Edward Gitre, C. North, and D. A. Bowman. Evaluating the benefits of the immersive space to think. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 331–337, 2020. doi: 10.1109/VRW50115.2020.00073

[26] L. Lisle, K. Davidson, E. J. Gitre, C. North, and D. A. Bowman. Sensemaking strategies with immersive space to think. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 529–537, 2021. doi: 10.1109/VR50410.2021.00077

[27] J. Liu, A. Prouzeau, B. Ens, and T. Dwyer. Design and evaluation of interactive small multiples data visualisation in immersive spaces. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 588–597. IEEE, 2020.

[28] J. Liu, A. Prouzeau, B. Ens, and T. Dwyer. Effects of display layout on spatial memory for immersive environments. *Proceedings of the ACM on Human-Computer Interaction*, 6(ISS):468–488, 2022.

[29] W. Luo, A. Lehmann, H. Widengren, and R. Dachselt. Where should we put it? layout and placement strategies of documents in augmented reality for collaborative sensemaking. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3491102.3501946

[30] W. Luo, A. Lehmann, Y. Yang, and R. Dachselt. Investigating document layout and placement strategies for collaborative sensemaking in augmented reality. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI EA '21. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3411763.3451588

[31] J. MacQueen. Classification and analysis of multivariate observations. In *5th Berkeley Symp. Math. Statist. Probability*, pp. 281–297. University of California Los Angeles LA USA, 1967.

[32] K. Marriott, F. Schreiber, T. Dwyer, K. Klein, N. H. Riche, T. Itoh, W. Stuerzlinger, and B. H. Thomas. *Immersive Analytics*, vol. 11190. Springer, 2018.

[33] D. Müllner. Modern hierarchical, agglomerative clustering algorithms. *arXiv preprint arXiv:1109.2378*, 2011.

[34] D. A. Norman. Natural user interfaces are not natural. *interactions*, 17(3):6–10, 2010.

[35] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[36] P. Pirolli and S. Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, vol. 5, pp. 2–4. McLean, VA, USA, 2005.

[37] D. Sacha, A. Stoffel, F. Stoffel, B. C. Kwon, G. Ellis, and D. A. Keim. Knowledge generation model for visual analytics. *IEEE transactions on visualization and computer graphics*, 20(12):1604–1613, 2014.

[38] K. A. Satriadi, B. Ens, M. Cordeil, T. Czauderna, and B. Jenny. Maps around me: 3d multiview layouts in immersive spaces. *Proceedings of the ACM on Human-Computer Interaction*, 4(ISS):1–20, 2020.

[39] K. A. Satriadi, B. Ens, M. Cordeil, T. Czauderna, and B. Jenny. Maps

around me: 3d multiview layouts in immersive spaces. *Proc. ACM Hum.-Comput. Interact.*, 4(ISS), nov 2020. doi: 10.1145/3427329

[40] R. Skarbez, N. F. Polys, J. T. Ogle, C. North, and D. A. Bowman. Immersive analytics: Theory and research agenda. *Frontiers in Robotics and AI*, 6:82, 2019.

[41] J. Stasko, C. Görg, and Z. Liu. Jigsaw: supporting investigative analysis through interactive visualization. *Information visualization*, 7(2):118–132, 2008.

[42] I. A. Tahmid, L. Lisle, K. Davidson, C. North, and D. A. Bowman. Evaluating the benefits of explicit and semi-automated clusters for immersive sensemaking. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 479–488. IEEE, 2022.

[43] I. A. Tahmid, L. Lisle, K. Davidson, C. North, and D. A. Bowman. Evaluating the benefits of explicit and semi-automated clusters for immersive sensemaking. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 479–488, 2022. doi: 10.1109/ISMAR55827.2022.00064

[44] J. Thomas and K. Cook. A visual analytics agenda. *IEEE Computer Graphics and Applications*, 26(1):10–13, 2006. doi: 10.1109/MCG.2006.5

[45] H. Wu, M. Mampaey, N. Tatti, J. Vreeken, M. S. Hossain, and N. Ramakrishnan. Where do i start? algorithmic strategies to guide intelligence analysts. In *Proceedings of the ACM SIGKDD Workshop on Intelligence and Security Informatics*, pp. 1–8, 2012.

[46] Y. Yang, T. Dwyer, K. Marriott, B. Jenny, and S. Goodwin. Tilt map: Interactive transitions between choropleth map, prism map and bar chart in immersive environments. *IEEE Transactions on Visualization and Computer Graphics*, 27(12):4507–4519, 2021. doi: 10.1109/TVCG.2020.3004137

[47] R. B. Zajonc. Attitudinal effects of mere exposure. *Journal of personality and social psychology*, 9(2p2):1, 1968.

[48] J. Zhao, M. Glueck, P. Isenberg, F. Chevalier, and A. Khan. Supporting handoff in asynchronous collaborative sensemaking using knowledge-transfer graphs. *IEEE transactions on visualization and computer graphics*, 24(1):340–350, 2017.