

Evaluating the Benefits of the Immersive Space to Think

Lee Lisle*
Center for Human-Computer Interaction
Department of Computer Science
Virginia Tech

Xiaoyu Chen†
Grado Department of Industrial and
Systems Engineering
Virginia Tech

Edward J.K. Gitre‡
Center for Human-Computer Interaction
Department of History
Virginia Tech

Chris North§
Center for Human-Computer Interaction
Department of Computer Science
Virginia Tech

Doug A. Bowman¶
Center for Human-Computer Interaction
Department of Computer Science
Virginia Tech

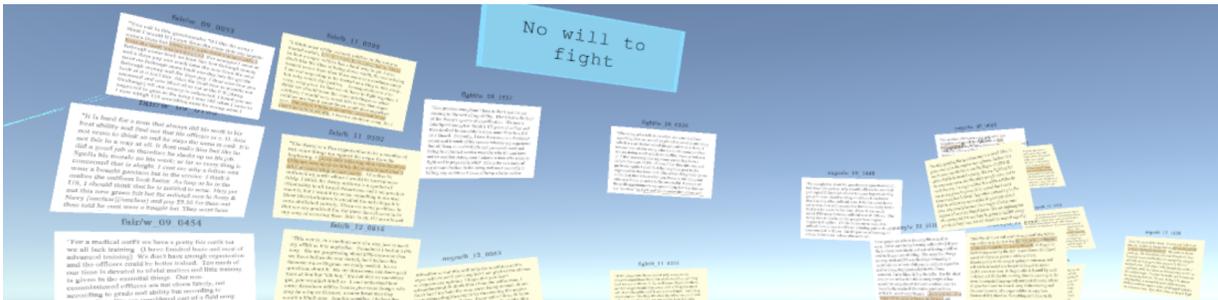


Figure 1: The Immersive Space to Think displaying text-based documents & user-defined labels. Users can create clusters of multimedia documents with labels, highlights, and notes to denote additional meaning.

ABSTRACT

Sensemaking with large multimedia dataset is a cognitively intensive task that requires analysts to understand the underlying stories that the dataset tells. Often, analysts use tools in order to offload cognition as well as convey their new understanding of the dataset; however, existing tools are limited by their underlying technologies. We have proposed a novel virtual reality tool to support sensemaking called the Immersive Space to Think (IST). IST can aid the process of analyzing multimedia data, but it remains unproven whether IST improves sensemaking performance over a traditional desktop setting. In a study performed over six weeks, one participant used both IST and traditional methods of sensemaking with a dataset of 100 text documents of transcribed survey responses from World War 2 soldiers to perform historical analysis. The participant was asked guided questions that produced three essays with their understanding of the data. After conducting a blind evaluation of the participant's interpretation of the data, a team of three experts in historical analysis concluded that the essays written with IST displayed a better understanding of the dataset. Furthermore, the participant gave positive feedback on IST, and also suggested possible improvements.

Index Terms: Human-centered computing—Human-computer Interaction—Interaction Paradigms—Virtual Reality; Human-centered computing—Visualization—Visualization techniques;

1 INTRODUCTION

Large multimedia datasets are problematic to analyze and understand thoughtfully. It could be that a history researcher wants to know

more about the motivation of a small town during the Civil War and has personal letters, farm yields, and pictures of its inhabitants. Perhaps an intelligence analyst has various sources of information on a possible terrorist cell and wants to figure out what their target is through cell phone records, video recordings of meetups and bank records. In both of these cases, it requires a significant effort on the analyst or researcher's part to understand the story behind the data. Sometimes the problem is that the analysts have too many data artifacts. In this case, powerful searching algorithms have been developed to identify documents with the highest relevance. Sun et al, for example, developed four different algorithms to search a dataset of 1000 text documents with only 34 relevant to the associated problem. Their best algorithm selected 5% of the dataset for further analysis that contained 95% of the actually relevant documents [17]. While this is impressive, the analysis of roughly 50 documents is still a difficult task for an analyst. When an analyst starts trying to make sense of a dataset, other issues arise such as how to organize the data artifacts into categories, corroborate evidence across artifacts, and synthesize a coherent account from the evidence [18]. While single data artifacts can be easily understood, they can only provide a partial understanding. Data analysts must interpret large, complex datasets to tell a complete, coherent story.

As Pirolli and Card establish, sensemaking is a cognitively intensive task where an analyst takes a dataset, forages for key evidence to support a claim, and revises their claims as they understand more of the dataset [13]. While sensemaking can cover tasks like writing an essay or coordinating a schedule, a common example of meaningful sensemaking is intelligence analysis. Often, sensemaking tools are evaluated using this kind of dataset [4, 9, 14]. Counter-terrorist datasets are often large, which requires an analyst to develop critical thinking skills in order to discard artifacts with no relevant data or combine several artifacts to create one coherent plot point.

Despite the importance of sensemaking with large multimedia datasets, technology-based tools to support this process have not advanced rapidly. Common practice uses traditional desktop computers, monitors, and input devices for document analysis. While these tools can take advantage of digital features such as search, copy/paste, and expressive markup, they are fundamentally limited

*e-mail: llisle@vt.edu

†e-mail: xiaoyuch@vt.edu

‡e-mail: egitre@vt.edu

§e-mail: north@cs.vt.edu

¶e-mail: dbowman@vt.edu

by available screen real estate, which only allows viewing of a few documents at a time, and which does not offer opportunities to spatially organize documents. In contrast, traditional approaches using paper documents that can be pinned up on a wall or spread out on a floor support rich spatial organization, but lack the powerful features of digital tools. Andrews et al. combined these two approaches by using large, high-resolution displays in their "Space to Think" work, which showed the potential of externalizing the analysis process into a large, interactive 2D space. We hypothesize that this approach might be even more powerful using immersive 3D technologies such as virtual reality (VR) and augmented reality (AR). VR and AR offer a rich three dimensional space that can place virtual artifacts anywhere around the user. Furthermore, recent advancements in VR displays have resulted in wireless interactions that remove the cables from the HMD to their desktop counterparts. These improvements have afforded better user experiences and should be explored for sensemaking opportunities.

Most existing immersive analytics research is confined to using quantitative datasets, rather than using multimedia data. Immersive analytics enables analysts to see data visualizations in new ways that can (sometimes literally) flip the data on its head to get a different perspective. However, multimedia datasets require a different kind of sensemaking. To explore AR and VR's potential for supporting multimedia sensemaking, we have developed the "Immersive Space to Think" (IST), which allows analysts to work in VR with text, image, video, and audio-based data artifacts. We hypothesize that IST could improve an analyst's sensemaking process and outcomes as compared to traditional methods and tools.

The study reported in this paper is a first step in comparing IST to traditional desktop sensemaking tools. We asked a single participant to perform historical interpretation tasks using a large dataset of primary source documents using both IST and common desktop tools for viewing, marking up, and note taking. The output of each task was a written essay; the essays were evaluated by three experts in historical interpretation. Results supported our hypothesis that IST would result in a better understanding of the source documents and a more coherent interpretation. While far from conclusive, these findings demonstrate the potential of the IST approach. We also discuss our observations of the sensemaking process in IST and the participant's feedback on our tool.

2 RELATED WORK

2.1 Sensemaking

As Stasko et al. establish, sensemaking is a difficult, cognitively intensive task where the analyst has to understand relationships between data artifacts in a large dataset [16]. Pirolli and Card cover the sensemaking process extensively in their work, where they detail the twelve phases and subloops performed in effective sensemaking [13]. The process is iterative, through analysts gathering data sources and meaning in the foraging loop and then better understanding how the sources fit together and mean more collectively in the sensemaking loop.

A key area of research for sensemaking is for education and self-learning. Butcher et al. researched how sensemaking is performed when the analyst (in this case, a student) has little domain knowledge [2]. They developed an application that automatically identified knowledge gaps to provide learning resources as a student wrote an essay. They found that their program supported effective sensemaking to help students learn new knowledge. Kloser et al. also worked on sensemaking in education through developing a framework for science teachers to facilitate sensemaking in secondary education [8]. They examined eight educators and each of their science methods courses through pre- and post-course interviews. They detail how sensemaking is difficult as is the facilitation of meaningful sensemaking, and what strategies educators can adopt to better convey knowledge.

2.2 Sensemaking of Multimedia Data

Understanding multimedia datasets has its own difficulties. Hicks et al. developed a set of guiding questions for students to aid in historical inquiry with source documents [7]. Hicks and Doolittle later developed an assistive tool called the SCIM Historical Inquiry Tutorial to guide students through their process and make sense of historical multimedia documents. They evaluated their tool and found that, with sustained teacher involvement, the use of the SCIM tool improved recall and strategic thinking in the test subjects [6]. Ntuen et al. developed a sensemaking and visualization application called AVALANCHE to utilize cognitive models to assist with sensemaking, collaboration, and situated cognition tasks [12]. They found in a military simulation study that their tool improved planning time, accuracy of interpretations, and reduced unnecessary information foraging. However, AVALANCHE deals with mostly quantitative datasets and the SCIM tool deals with a limited number of documents at the same time, while our approach can resolve these issues.

A key influence for our approach was Andrews et al.'s Space to Think sensemaking tool for multimedia documents [1]. Their tool leveraged a large, high-resolution display to create a better sensemaking environment for users and ran a study comparing their tool to working with a 17-inch display. Their study revealed that users would manage documents in completely different ways based on screen size. Small screen users would maximize each window and use separate files and handwritten physical notes to keep track of their sensemaking process while the large screen users took no physical notes and used relative spatial location as a form of external memory. We extend this work by exploring the use of three-dimensional space to think in VR.

2.3 Immersive Data Analytics

VR and AR systems have afforded the development of new ways to understand large amounts of data in a virtual environment. Chandler et al. define these visualizations and sensemaking opportunities as "immersive analytics" [3]. Marriott et al. detail how immersive analytics builds upon the fields of data visualization, visual analytics, virtual reality, and human-computer interaction [10]. Skarbez et al. make the argument that these VR/AR applications share a goal of supporting user sensemaking through abstract data visualizations [15]. They go on to define the general research areas of immersive analytics as ways of combining human and computer intelligence, leveraging immersion, designing immersive analytics systems, facilitating collaboration through immersion, and changing the process of analysis with immersion. Our design approach aims to address several of these research areas, while the study described in this paper aims to establish a preliminary level of effectiveness of the tool.

Immersive analytics has been helping researchers perform sensemaking on many different topics. For example, Nim et al. developed a way to better understand and visualize bird migrations [11]. Their design allows for users to see birds migrating alongside the geospatial data to provide more contextual data such as environmental concerns, thereby enhancing their understanding of the movement patterns. Cordeil et al. designed a system to explore multivariate data using modeless interaction through embodied data axes [5]. Users can place axes like physical objects, allowing users to change the visualizations in novel ways. However, while these studies assist with visualizing and sensemaking with quantitative data, our study aims at improving sensemaking using multimedia document types through interaction, not just visualization.

3 METHODS

3.1 IST Tool

Our tool, *The Immersive Space to Think* (IST), was designed to support data analysis for multimedia sources. Its design and capabilities

are listed below.

3.1.1 Technical Details & Design

We began with the goal of designing a version of “Space to Think” that allowed analysts to operate in a virtual environment rather than on a large two-dimensional screen. Documents and other multimedia artifacts would be represented as 2D windows within a 3D space, and users would be able to view/read, markup, and place the artifacts anywhere in a large 3D workspace to reflect their analysis. In our implementation, we use a VIVE Pro head-mounted display (HMD) with wireless attachment running on a desktop with an Intel i7-8700k processor and an NVIDIA 1070 graphics card. Additionally, our approach uses two VIVE Pro wireless controllers to interact with the documents. The user’s movement is tracked by a SteamVR 2.0 lighthouse covering a 4 meter by 8 meter space that was kept clear of obstacles.

The physical layout of the space also informed several design decisions within the virtual environment. The tracked area was visualized in IST through a virtual wooden floor. This informs users where they can safely travel without colliding with the walls or desks just beyond the tracked area. Furthermore, one of the long edges of the tracked area is displayed as a bulletin board where the documents are initially displayed with vertical bars between sets of documents to delineate separate categories. These choices were made to assist the users to understand the virtual environment and how it related to the real world.

3.1.2 Interaction Methods

IST’s main contribution is the ability to interact with a large set of multimedia documents in various ways in order for the user to both extract and convey additional meaning. These interactions are as follows.

Single-hand Movement of Documents - Users can select a document by pointing at it with the virtual ray emanating from one of the controllers and pressing the trigger button. At this point, the document is “grabbed” at the initial point of intersection of the ray and the document, and can be translated and rotated at the end of the ray at that fixed initial distance. The user can also press on the VIVE Pro controller’s trackpad to move the document closer or farther away along the ray.

Multi-hand Resizing and Moving of Documents - Similarly, users can “grab” documents using both controllers simultaneously. In addition to translation and rotation, this enables the user to resize documents based upon the two initial points the user grabbed the document. IST utilizes a “pinch-to-zoom” metaphor where the user moves the controllers and rays, and the entire document is translated/rotated/resized to keep the grab points fixed on the document.

Text scrolling - When documents are larger than the given space for text, IST allows for the user to “scroll” the documents much like a window on a traditional two-dimensional desktop. To scroll a document, the user needs to point at a document with the ray cast from the controller and drag their finger on the trackpad up or down to scroll in the opposite direction (similar to scrolling apps on a smartphone).

Text Highlighting - Often when reviewing text documents, users will want to highlight key phrases or sentences to easily retrieve passages that contain extra insight or represent a key theme or idea for a group of documents. IST allows users to do this through pointing at a document and pressing the “menu” button on the VIVE Pro controller. The document will then temporarily highlight the word that the controller’s ray intersects with. If the user holds down the menu button, it will highlight all words from the initial word to the endpoint of the ray. The user can then use a software button on a user interface (UI) panel that automatically appears when the user points the ray at a document, as seen in figure 2, to permanently

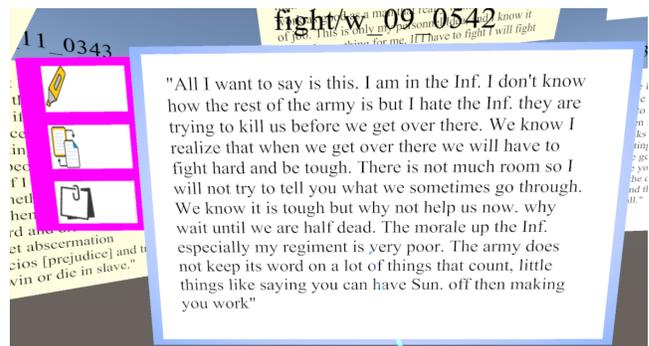


Figure 2: An example of a text document in the Immersive Space to Think. The UI panel on the left features document-specific interactions that can manipulate the document through highlighting words, copying the document, or taking notes on the document.



Figure 3: An example of a label in the Immersive Space to Think. The UI panel on the left features a search tool that highlights title bars of documents that contain the label's text.

highlight the selected text. The user can highlight different portions of the text while the permanent highlights are retained.

Document Copying - Some documents can support multiple arguments at the same time, so users may want to place the same document in multiple locations. In order to facilitate this, users can press a software button on the document’s UI panel that will create a copy of that document. The document is created as a “grabbed” object immediately, so the user can put the fresh copy in a new place.

Note Taking - To support annotation and cognitive offloading while users interpret documents, we included a “note-taking” feature that allowed users to write about each document in a separate text panel placed beneath the original document. To activate the notes, the user would press a software button on the document’s UI panel. Then the user would dictate the note verbally. In our case, we used a wizard-of-oz approach where an experimenter would add the desired note in real time to the document, though potential enhancements such as adding a Google Speech-To-Text feature are being explored.

Label Creation - Similar to the note-taking, users can also create labels with one to three words. This affords users the ability to create topic headers or key phrases that they can arrange supporting documents around. Label creation is executed through a wizard-of-oz verbal request, with the user stating they want a label and what it should say. The label will then appear one meter in front of the user.

Keyword Search - Once a label is created (see previous paragraph), users can use the labels to search for documents that contain their word or phrase. Each label has a unique UI panel with a single button depicting a magnifying glass (seen in figure 3). Clicking this button causes all documents with the label’s text to have their title bar highlighted. This allows the user to quickly see which documents have the desired word or phrase within their text.

PDF Viewing - The user can also view PDFs while using IST. These act like linked images where each image represents one PDF page and only one page is available at a time. Users can switch

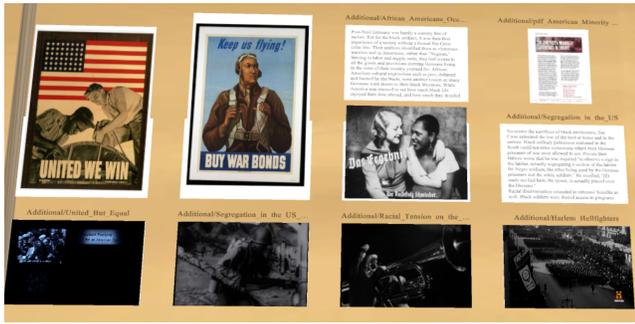


Figure 4: IST can handle many different files types, including images, text files, *pdfs* and videos (seen on the bottom row). Additionally, these are the resources the participant provided for phase three.

between pages through swiping left or right on the trackpad to go to the previous or next page respectively. An example of a PDF object can be seen in figure 4.

Video Playback - Videos are displayed with their starting frame visible, but playback is paused. Users can point the controller ray at a video and press the 'menu' button on the VIVE Pro controller to play or pause the video. Users can also swipe left or right to rewind or skip fifteen seconds in the playback.

3.2 Hypothesized Design Benefits

We designed IST to support sensemaking of complex multimedia datasets in an expressive 3D space, and we hypothesized that it would enable better sensemaking, both in terms of process and outcomes, as compared to a traditional desktop setup.

Modern desktops rely on 2D displays that are extremely limited in the amount of space provided. A single 61 cm (24 inch) diameter screen at the United States Occupational Safety and Health Administration (OSHA) minimum recommended distance of 50 cm provides a 61.7 degree field of view and field of regard. The HTC VIVE Pro, on the other hand, provides a 110 degree field of view with a field of regard of 360 degrees and as much depth as the tracked area allows. This increase in space allows for users to create much more complex document organization structures and leverage distributed cognition more easily. Although desktop monitors currently offer greater pixel density, VR users have the ability to move close to documents for readability and step back to see an overview of many documents.

Users can also leverage distributed cognition to quickly recall their thought process in IST. Since the window layout on a desktop is limited, users will often only have one to three documents open, implying that their relative layout will likely have no additional meaning. In IST, users can create structures of meaning, where the relative spatial organization of related files denotes more information about their relationship. For example, a group of documents placed together in IST can mean that they share a common theme or support an argument denoted by a nearby label. In addition, VR users can make use of depth and document orientation in their layout, resulting in more expressive structures.

Finally, searching for specific words or phrases can be difficult if the user's dataset is spread amongst a large set of files. With IST, however, we ameliorate this through an emphasis on visual feedback. The title bar can be seen on either side of each document and is highlighted during search with a contrasting color to quickly draw the user's attention. If the desktop has a variant of this multi-file search, it typically displays a list of files with the matching string. In IST, we leverage spatial memory to allow the user to assign meaning to locations in the virtual environment.

3.3 Experimental Design

The goal of the study was to gather preliminary data about the effectiveness of IST compared to traditional desktop tools for sensemaking of multimedia datasets. Our participant was an undergraduate male student working for honor's credit in history; he had no prior experience with AR or VR. To address our goal, the study was split into four phases: writing the first essay using a traditional desktop method, writing the second essay using the Immersive Space to Think, writing a self-guiding topic exploration essay, and a post-study questionnaire. Since we wanted to use a historical analysis task, we selected a set of 100 responses written by white and black World War II US soldiers on a race-relations survey conducted by the US Army (Survey #32). We then generated two prompts related to this dataset:

1. In many corners of the United States, race relations were tightly regulated by custom and law into and beyond the 1940s. What impact did military service have on black service members' views on race and race relations in the US?
2. As a whole, men who served in World War II were better educated than soldiers in previous US conflicts. How in World War II did having a higher education impact soldiers' views on the army in relation to their own personal experience in uniform?

These prompts were designed to elicit 500-700 word essay responses within a time limit of two hours. The time limit was intended to allow for time spent weeding out irrelevant survey responses to the prompts as well as finding particularly compelling responses with supporting evidence for their answer. Both the first and second phases used the same desktop computer as described in section 3.1.1.

For the first phase of the study, we created a directory of the 100 survey responses split into five sub-directories that served as categories. The responses were chosen by selected set of keywords relevant to the survey. Furthermore, these documents were in plain-text format (*.txt*), and the filenames had "w_" or "b_" preceding the ID number of the file to denote the response was from a white or black soldier, respectively. The participant worked at a desk with one 61 cm (24 inch) monitor with an aspect ratio of 16:9 and a display resolution of 1920x1080. The participant was instructed to use the Microsoft Notepad application to view the files and to use Microsoft Word to type their essay. They were told they could edit the files, create new directories, or move the files to different directories or sub-directories. All interactions were recorded through a Python tracking script that recorded window focus changes, files and directories being open or closed, files or directories being created, clipboard interactions, file or directory hierarchy changes, general key-logging, and window resizing. Furthermore, the screen was being video-recorded for later reference. No training was required for this phase.

After a two-week interlude, the second phase required the participant to answer the other generated prompt using IST. This first required a training period that lasted approximately 20 minutes to acclimate the participant to the features described in section 3.1.2; this was not considered part of the two-hour time limit. All interactions were logged in a separate text file by the IST software, as well as camera and controller positions every 100ms. Additionally, the Unity development environment was screen-recorded so we could reference actions taken by the experimenter as well as record what the participant was seeing, with audio recorded by the VIVE Pro's built-in microphone. The document set was the same as in phase one, separated into the same categories and denoted by a blue label above the grouping. Race of the respondent was denoted by the associated artifact being white for a white soldier and light yellow for a black soldier. During this phase, the participant was given the option to have a keyboard to write the essay while sitting at a desk

in the center of the tracked area. However, the participant asked if they could instead dictate their response to the experimenter, and their request was granted.

After phase two, the participant was then tasked to choose their own question that they wanted to investigate further. During the two-week break between phases, we asked the participant to find ten to twenty new sources in *pdf*, *mp4*, *jpg*, or *txt* formats as additional resources to answer their own question.

The third phase involved writing a third essay with the participant-designed prompt: “What were the lives of African American soldiers serving overseas like compared to those serving at home?” The participant was given two hours using IST to analyze sources and outline/write their essay using the original 100 sources plus their additional 12 sources. The sources they provided included four videos, three images, one PDF document, and four text files. All of these except two text files can be seen in figure 4.

Lastly, phase four involved asking the participant to give 1-2 paragraph answers to a series of six questions on their experiences during the study. These questions included:

1. What was your overall impression of the Immersive Space to Think and its ability to support document analysis?
2. If you had to write a long form essay or paper, would you use IST, and why or why not?
3. Was there any tool or feature missing from IST that could have helped you perform your tasks, and could you describe them?
4. Were there any difficulties that you encountered while using IST, and what were they?
5. Did you change your workflow or thought process while using IST? If so, how?
6. Please compare and contrast the two methods of research you performed.

After all the essays were written, they were given to three experts in historical analysis (faculty members in the Department of History at Virginia Tech) for evaluation. The experts were asked to state which of the first two essays was of higher quality and why in three to four sentences. They were then asked to give a one-paragraph evaluation of the third essay. None of the experts were informed what tool was used to write each essay. It should be noted that one of the authors participated as an expert evaluator, but like the others, he was blind to the source of the essays.

4 RESULTS & DISCUSSION

All three experts concluded that the second essay (written with IST) showed more understanding and analysis of the source materials compared to the first essay (written using the desktop setup). Expert one stated that the essay written on the desktop, contained “too much uninformed opinion,” and that the essay written with IST was “much more grounded in the actual source material, with a number of nuanced observations that are reflective of broader historical trends.” Expert two agreed, saying that essay two was “by far the superior essay, both in terms of prose and context.” Furthermore, expert three stated that the essay written with IST “makes a concerted effort to provide a coherent analysis of the collection set, observing rhetorical patterns that run across multiple documents.” They went on to state that the participant “not only notes rhetorical themes but the relation between these patterns and their social and geographical context.”

The experts all coming to the same conclusion on the outcomes of the analytic process provides some evidence that IST can afford greater understanding of the content as well as a tool for better organization of thoughts than the traditional desktop setup. Of course, results from a single participant cannot be said to be conclusive.

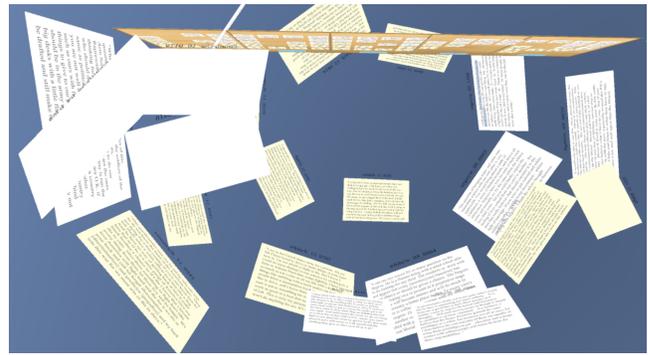


Figure 5: The final layout of documents arranged by the participant for writing essay two (top-down view). The layout was a rough dome shape, where the participant sat in a chair in the center of the documents where they could reference documents while writing their essay.

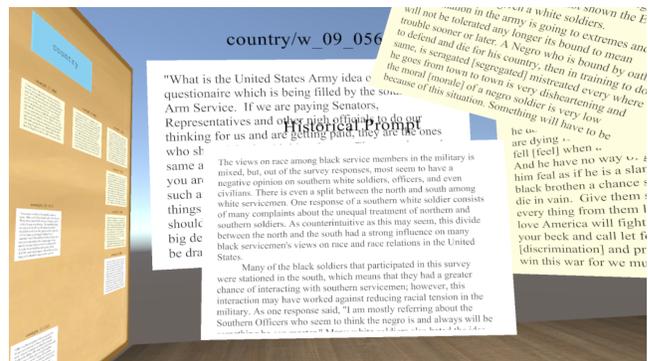


Figure 6: The final layout for the participant for writing essay two, as seen by the participant while they were writing the essay. From this angle, the bulletin board is seen on the left side of their view, while the right side contains documents they wanted to reference during writing. The central document in the foreground is the window they used to write the essay.

Looking at the details of the analytic process can also be instructive. In IST, the participant created a dome-like structure of documents when they prepared to write the document, as seen in figure 5. Their point of view can be seen in figure 6. All of the documents they wanted to reference formed the right side of the dome, while the left was formed by the bulletin board. If they needed to reference a document while writing, they would either turn to look at the document or bring it closer to them. Once they were finished referencing, they would return the document to the prior position. The participant kept documents that were on related topics close together in the dome, but with no clear separation of categories. The dome layout affords the ability to quickly look at any document they needed, whether to quote or reacquaint themselves with the context or meaning. Furthermore, it puts the essay writing in a prominent place at the center of all of the documents, while still keeping references visible. It should be noted that this is in stark comparison to their strategy for essay one, where they had, at most, one file open while writing and it was always partially obstructed by the essay window.

Through the log files we can observe a few notable tidbits of information. Essay one, on the desktop, took 89 minutes for the participant to write, while they spent 104 minutes in IST writing essay two. This increase in time may be due to having to dictate their answers while using IST. Interestingly, the participant only

had one file open at a time in the desktop scenario, much like the original Space to Think study [1]. However, the source document was always open next to their essay document in a way that they could read either easily. This small difference could be due to our screen being a 24-inch monitor versus the 17-inch monitor used in their study. Furthermore, the participant interacted with 60 of the source documents with an average time open and in focus of 49.3 seconds while working on the desktop, while interacting with 31 documents with an average time in the center of view of 41 seconds (during foraging phase) while working in IST. However, this can be accounted for by some familiarity with the dataset, as the participant had already completed essay one using the desktop setup.

During the foraging phase with IST, the participant would form rough temporary clusters of documents that had shared meaning. For example, they created a “trash” pile of documents that didn’t add relevant content or insight. While many of these clusters did not appear in the final dome-like structure, they did indicate meaningful relationships that formed during the sensemaking process. The formation of the final dome structure started at 28 minutes into their VR session and involved seventeen of the source documents. Lastly, the participant did not make use of highlighting text, taking notes, search, or copying (text or documents) in either setup. We do not believe this will be typical behavior.

The feedback on the essay from phase three proved to be a little more mixed. Expert one stated that he liked the prompt that the participant decided to use, and thought the participant did well in addressing the differences the prompt asked about, but also that he would have “liked to see some specific examples of this referenced in the essay.” Overall, however, expert one thought it was solid. Expert two continued this theme in stating that essay three “has a lot more context ... but the attempt makes the overall essay less organized.” While the response for essay three is less enthusiastic than essay two, it seems that the main problem with it, according to the experts, was that it lacked good examples. This may point to a weakness in dataset curation, as examples of black soldiers abroad could only come from the additional sources provided by the participant.

The post-study survey yielded positive feedback from the participant on our IST design. The participant started with “[IST] looks beautiful, especially the homepage. It is nice how you could organize the documents in a 360 environment.” However, they said that they wouldn’t use IST for an assignment unless they became more experienced with it. They also mentioned that the HMD became heavy after during the experiment which impacted their concentration towards the end of essays two and three. The participant also provided a few feature suggestions. “The control feels limited since you are unable to hide or minimize any documents like you could on a PC.” Instead, they “threw not useful files under the floor.” However, the latter quote proved to also be an annoyance, as they complained in the difficulties question that “if a window was moved behind a wall or under the floor, there doesn’t seem to be a way to get it back.” They also stated that the process of typing was hindered as they felt that they didn’t want to “switch between the keyboard and controllers constantly” and that, since they were dictating, it interrupted their normal flow of writing. Therefore, they came to the conclusion that “despite the innovativeness of IST [they] cannot be as efficient with it as [they are] with a computer.” The participant also stated that their process changed depending on the environment. With a laptop, they would “organize different sources into different folders and pick the useful ones,” though they admitted they didn’t do that in phase one. With IST, they “moved useful files to different locations” This does describe their process in IST, where they formed the dome of data artifacts.

5 LIMITATIONS

Obviously, a clear limitation of this study is that there was a single participant. This resulted from an opportunity to perform a long-

term project with this student in the context of other research on IST, so we designed this experiment specifically to take advantage of this opportunity. Our intention was to learn from their process to inform a more complete study in future research. Nevertheless, the insights gained from the study still indicate the potential of the IST approach.

Second, we cannot rule out the possibility that the historical analysis may have been better with IST simply because the participant used IST second, after using the desktop. While the ordering of the questions, as well as which was performed in IST, was kept from the experts, they may have deduced the ordering on their own. As one reviewer stated, “*historical analysis is iterative, honed by practice.*” Furthermore, the participant’s familiarity with the dataset may mean that they were adding to their understanding and sensemaking with the documents over time, though a two-week interval between essays may have reduced this somewhat. These issues can be accounted for in future studies with more than one participant through counterbalancing, as well as using separate datasets for each prompt.

One key part of this experiment was that the participant had to completely write their essays in the same two-hour session during which they analyzed the data set. Furthermore, while we offered the participant a keyboard to type their document while using IST, they declined and preferred to dictate their response to the experimenter as they wrote it. This was mentioned in their phase four survey as a major roadblock for their sensemaking process. We believe we can address this in later versions of this experiment both through giving the participants more time as well as using AR in either a pass-through or see-through method. A pass-through AR implementation would allow the user to be completely immersed until it was time to write the document, while a see-through AR implementation would allow the user to not be detached from the real world tools they have. Ultimately, we plan to implement both and explore the tradeoffs to address this issue.

6 CONCLUSIONS & FUTURE WORK

While we cannot provide definitive proof that IST affords better performance for sensemaking tasks, we can conclude that it is a promising tool demonstrating the potential of immersive 3D space to support spatial organization during analysis. We also can see that it may allow users to connect and understand concepts and themes in multimedia datasets in ways that outperform a simple desktop setup. IST allows users to view a large dataset easily as compared to a directory of files, enabling them to quickly evaluate documents and organize them in three-dimensional space. Spatial layout and annotation/markup tools provide users with an opportunity to offload cognition onto the environment. Even though the participant in our study did not use many of our additional interaction techniques, this would imply that the base features of moveable data artifacts, 3D space, and labeling may be effective on their own.

While our study showed promising preliminary results, we intend to perform further studies with similar designs to continue our research. With a higher participant count, we would have a more reliable pool of results to draw any conclusions. Furthermore, we plan on tweaking the experimental design such that it is easier for the participants to write the essay. It may not be necessary to get an entire written deliverable, but have them talk through their essay outline to see if they have drawn good insights into the dataset. Furthermore, we want to try to eliminate the wizard-of-oz features in the system such that the labeling and note taking features are more robust and can be done with minimal experimenter involvement.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Jessica Taylor and Dr. Bradley Nichols of Virginia Tech’s history department for their assistance as additional experts in historical analysis.

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] K. R. Butcher and T. Sumner. Self-directed learning and the sense-making paradox. *Human-Computer Interaction*, 26(1-2):123–159, 2011.
- [3] 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.
- [4] I. Colville, A. Pye, and M. Carter. Organizing to counter terrorism: Sensemaking amidst dynamic complexity. *human relations*, 66(9):1201–1223, 2013.
- [5] 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.
- [6] D. Hicks and P. E. Doolittle. Multimedia-based historical inquiry strategy instruction. do size and form really matter. *Research on technology in social studies education. Greenwich, CT: Information Age Publishing*, 2009.
- [7] D. Hicks, P. E. Doolittle, and E. T. Ewing. The scim-c strategy: Expert historians, historical inquiry, and multimedia. *Social Education*, 68(3):221–226, 2004.
- [8] M. Kloser, M. Wilsey, T. C. Madkins, and M. Windschitl. Connecting the dots: Secondary science teacher candidates’ uptake of the core practice of facilitating sensemaking discussions from teacher education experiences. *Teaching and teacher education*, 80:115, 2019.
- [9] 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.
- [10] 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.
- [11] H. T. Nim, B. Sommer, K. Klein, A. Flack, K. Safi, M. Nagy, W. Fiedler, M. Wikelski, and F. Schreiber. Design considerations for immersive analytics of bird movements obtained by miniaturised gps sensors. In *Proceedings of the Eurographics Workshop on Visual Computing for Biology and Medicine*, pp. 27–31. Eurographics Association, 2017.
- [12] C. A. Ntuen, E. H. Park, and K. Gwang-Myung. Designing an information visualization tool for sensemaking. *Intl. Journal of Human-Computer Interaction*, 26(2-3):189–205, 2010.
- [13] 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.
- [14] A. C. Robinson. Supporting synthesis in geovisualization. *International Journal of Geographical Information Science*, 25(2):211–227, 2011.
- [15] 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.
- [16] J. Stasko, C. Görg, and Z. Liu. Jigsaw: supporting investigative analysis through interactive visualization. *Information visualization*, 7(2):118–132, 2008.
- [17] Z. Sun, E.-P. Lim, K. Chang, T.-K. Ong, and R. K. Gunaratna. Event-driven document selection for terrorism information extraction. In *International Conference on Intelligence and Security Informatics*, pp. 37–48. Springer, 2005.
- [18] W. Zhu, P. Cui, Z. Wang, and G. Hua. Multimedia big data computing. *IEEE multimedia*, 22(3):96–c3, 2015.