The Impact of Physical Navigation on Spatial Organization for Sensemaking

Christopher Andrews and Chris North

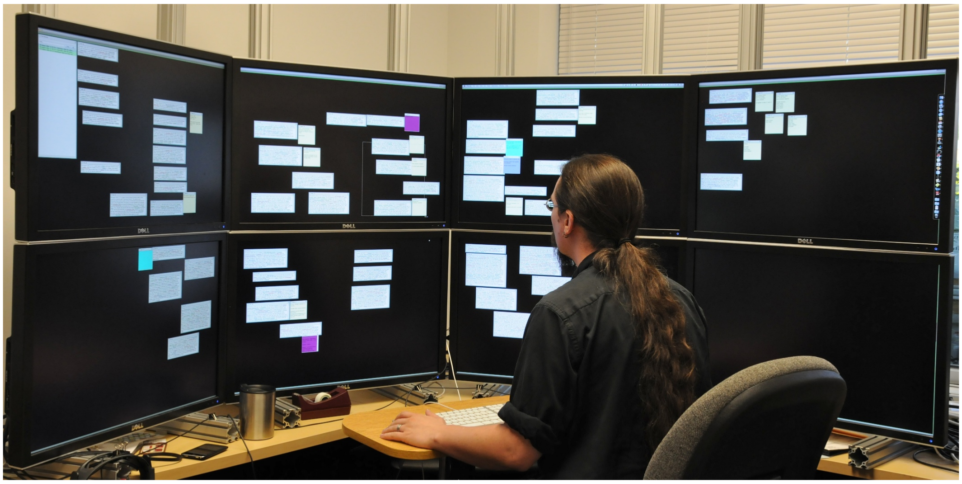


Fig. 1. Working on the 33 megapixel large, high-resolution display.

Abstract—Spatial organization has been proposed as a compelling approach to externalizing the sensemaking process. However, there are two ways in which space can be provided to the user: by creating a physical workspace that the user can interact with directly, such as can be provided by a large, high-resolution display, or through the use of a virtual workspace that the user navigates using *virtual navigation* techniques such as zoom and pan. In this study we explicitly examined the use of spatial sensemaking techniques within these two environments. The results demonstrate that these two approaches to providing sensemaking space are not equivalent, and that the greater embodiment afforded by the physical workspace changes how the space is perceived and used, leading to increased externalization of the sensemaking process.

Index Terms—Sensemaking, visual analytics, physical navigation, embodiment large, high-resolution displays

1 Introduction

The process of sensemaking is a cognitively demanding task that requires an analyst to piece together disparate items of information into a coherent whole. It involves not just gathering facts, but understanding them. In the domain of intelligence analysis, the task is further complicated by fragmentary, conflicting information, and the presence of agents who seek to deliberately confuse and obscure their activities [30].

Among the various approaches to sensemaking, we find spatial organization to be particularly compelling. The human perceptual system is well adapted to using spatial relationships to find patterns, categorize information, and otherwise simplify internal computation [17]. Externalizing information in the form of spatial relationships provides a way to reduce the load on memory, while at the same time adding information to the documents based on the context of their surroundings.

Another advantage of spatial organization is that it is very flexible, and can be used in a variety of ways. For example, it can be used to break objects up into discrete sets, or it can be used to indicate gradations of a relationship (e.g., an object that is placed between two other objects, but closer to one than the other can indicate that the new object is related to both of the original objects, but that there is a stronger connection to the closer one). The organizational structures can range from simple piles to hierarchical structures with internal categorization or ordering. A truly free-form space can even permit a variety of spatial metaphors to all coexist and interrelate [24].

Spatial structures can also be created without any formal rules governing their creation. This supports the process of incremental formalism, where the meaning of the structures (and the structures themselves) develop in step with the understanding of the objects beings structured [26].

In our previous work, we demonstrated how the spatial environment provided by large, high-resolution displays leveraged these advantages for sensemaking [1]. The available space allowed analysts to spatially organize ordinary text documents just as they might lay paper out on a table. While the large display creates this spatial environment implicitly, there are other ways to provide an identical amount of space, including many techniques for working with large information spaces on small displays [7]. In fact, there are

- *Christopher Andrews is with Middlebury College. E-mail: candrews@middlebury.edu.*
- *Chris North is with Virginia Tech. E-mail: north@cs.vt.edu.*

Manuscript received 31 March 2013; accepted 1 August 2013; posted online 13 October 2013; mailed on 4 October 2013.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

a number of "spatial hypermedia" tools that have been developed specifically to take advantage of the power of using a free-form organizational space for sensemaking [5, 25]. Occulus' Sandbox takes a similar approach, adding specialized marshaling tools to the organizational space specifically to support intelligence analysis [31]. What is notable about these systems is that they are explicitly designed to leverage human spatial abilities, but they are all designed to work on conventional displays, using small representations and a *virtual workspace* coupled with an overview+detail approach to provide the tool's space.

This leads us to ask if there is really a difference between the space provided by the large, high-resolution display and these virtual workspaces for sensemaking. We can distil the issue down to one of physicality. On a large, high-resolution display, the space can be physical. It is entirely available through the use of *physical navigation*, actual physical movement on the part of the user (e.g., walking, turning, glancing) [4]. On the other end of the spectrum, a conventionally sized display must use a virtual workspace, providing a small viewport that reveals only a portion of the available space, with access provided through *virtual navigation* (e.g., panning, zooming) [10].

There are a number of studies that have examined ways in which the physicality of large displays can affect the user behavior. Ball has looked explicitly at the difference between physical and virtual navigation for basic visualization tasks such as route tracing and search. He showed significant performance improvements when the display permitted the use of physical navigation, as well as significant drops in user frustration [4]. Shupp built upon these findings by showing that curving the display around the user, thus reducing the physical effort required to navigate the space, resulted in further improvements over virtual navigation [27]. Czerwinski et al. have also demonstrated that the wider field of view afforded by large displays aids navigation in virtual environments by providing more visible landmarks [9].

While many of these results could be attributed to pure mechanical efficiency (speed of turning to look at part of the display vs. an explicit interaction), there is also a cognitive component at play as well. For example, in Czerwinski's study, the performance improvements came from a greater spatial understanding fostered by the greater access to landmarks rather than pure mechanical efficiency. This raises an important question: does the way in which the space is presented to the analyst affect the way it is used for sensemaking? More specifically, how does it affect how externalization is manifested in spatial organization of the information? To answer this question, we conducted a study in which users were asked to analyze a collection of sample intelligence reports, using the provided workspace to help them manage the information.

2 Study Design

The goal of this study is to observe how the increased availability of information through physical navigation affects how a free-form spatial workspace is used during a sensemaking task. We will be looking at how users conscript the space to externalize information, specifically looking at the overall layout of the space and the structures constructed by the users. We also want to look for evidence that any such externalizations are actually used as part of the cognitive process, rather than just being created and abandoned. Finally, if the use of physical navigation does change the nature of the externalizations, we will look for how this impacts the analytic approach adopted by the users.

2.1 Experimental setup

To support physical navigation, we used a 10,240x3200 tiled LCD display constructed from eight 30" panels (Figure 1). The panels are arranged in a 4x2 grid, and curved around the user. As can be seen in the figure, the user is provided with a rolling chair and a tray table for keyboard and mouse. This system is an interesting platform for a couple of reasons. First, the entire display can be driven from a single computer, which allows us to run conventional applications on the display without any modifications. For our purposes, it also allowed us to develop a single environment that would run for both conditions. Second, the relatively small form factor of the display means that it can be deployed on a desktop, where it could be used as a personal workspace for daily work. As we are looking at a process that is a daily activity for intelligence analysts, this is an important point to consider.

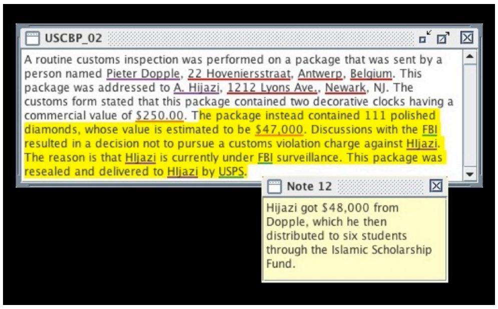


Fig. 2. An example document showing user highlighting and an attached note. Entities are underlined to aid recognition.

For the virtual navigation condition, we provided our participants with a 17" LCD panel with a resolution of 1280x1024. While 22" to 24" displays have started to replace 17" displays on the desktop, it is a relatively recent trend, and our desire was to emphasize the difference between the two environments, thus emphasizing the amount of virtual navigation required to interact with the space.

2.2 Sensemaking environment

The tool that we built for this study provides the user with a very basic workspace in which full text documents can be spatially arranged to convey relationships. The familiar selection rectangle tool was provided to support the rapid movement of documents in groups, and encourage experimentation and rearrangement.

In our previous work with analysts, we found that the ability to annotate text was seen as very important [1]. The environment supports two forms of annotations: highlights and notes (Figure 2). A highlight can be added to a document very simply by holding down a modifier key while selecting a block of text with the mouse. The notes follow the "sticky note" metaphor, and can be "stuck" to documents, following them wherever they are moved. They can also be stuck to the background as labels for regions of space, or just as freestanding text. The notes are yellow by default but the color can be easily changed to label documents or regions.

A custom file browser provides access to the documents in the dataset (Figure 3). The browser lists all of the available files, sorted by date, and color coded by the state of the document (unseen, open in the workspace, and seen, but no longer open). The browser also supports full text search, powered by the Lucene search engine [8].

Both conditions were provided with identically sized workspaces. On the large display, there was a direct one-to-one correspondence between pixels in the workspace and the physical pixels of the display, meaning the entire workspace was always available at the maximum level of detail. We also purposely fixed the font size (10pt) so that the participants in the two conditions would have equal information density within the workspace. For the large display users, this meant that they had to employ physical navigation to access the entire workspace.

For the small display condition, we tried to keep the basic environment as similar to the large display as possible, while adding virtual navigation techniques similar to those found in existing sensemaking tools (e.g., [5, 25, 31]). The small display condition makes use of the overview+detail approach to support access to the full space by providing a floating, resizable overview window (Figure 3). We chose this approach over other techniques for working with virtual workspaces because it has been shown to be the

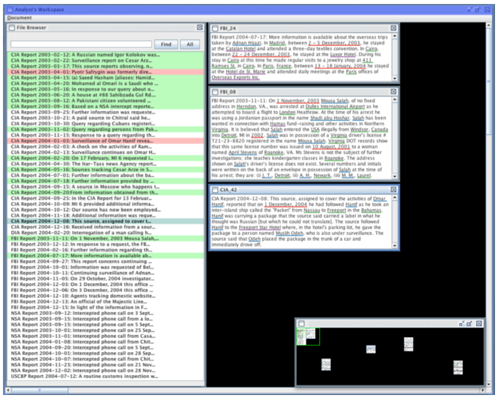


Fig. 3. The small condition environment. A file browser is on the left, and the overview is shown in the lower right.

preferred technique [7], and it is the technique used in the aforementioned tools. In our implementation, the overview is fully functional, allowing the user to select a document or documents and move them around the workspace directly from within the overview. In addition, a single click within the overview moves the viewport of the main view to the location in the workspace. This overview was available in both conditions, but open by default in the small display condition.

The viewport used in the small display also supports a number of panning techniques. Clicking and dragging the background moves the viewport around the space, similar to the move tool found in most image viewers. This implementation, however, adds a simple physics model so that if the mouse button is released in the middle of a drag, the viewport will drift to a stop, allowing the user to traverse the space faster. Dragging objects to the edge of the viewport also triggered an automatic panning, to mirror the process of dragging objects across the workspace on the large display as closely as possible. Finally, simple scrollbars provide the conventional viewport panning interaction.

2.3 The task

The task to be performed by our users was to solve a synthetic intelligence analysis problem using a basic analytic tool that relies on manual spatial organization as the primary evidence marshaling technique. This is the same basic task that we used in our previous explorations of large displays for sensemaking [1], and Robinson used in his studies [24]. The single independent variable in this study is the use of physical or virtual navigation to interact with the full extent of the workspace.

Our participants were asked to make sense of a collection of synthetic intelligence reports concerning the purchase and movement of bioweapons and their imminent deployment in the United States. The dataset consists of 58 documents, five of which are deliberately misleading, and another ten of which only provide background information. Each document is between five and twenty sentences long and contains reports of the activities of various “persons of interest”. There is no fixed starting place in this dataset, and we expected that the participants would need to read every document in order to correctly identify the threat with enough detail and supporting evidence to prevent the attack and arrest the instigators.

Each session began with a short introduction to the analytic environment. Participants were told that the underlying concept behind the environment was that the space could be used to express relationships between documents, but no particular organizational strategies were discussed. They were also shown samples of the kind of documents they would be working with and briefly introduced to the idea of making inferences and finding evidence to support or refute them. We then presented the participants with the document collection and asked them to identify the underlying plot. They were then allowed two hours to work with the data in the environment. After the two hours, each participant was asked to fill out a report listing the people, places, and events that they identified as significant, and to write a short description of what they thought was actually going on in the scenario.

Table 1. Summary of Quantitative Results

Quantitative metric	Large display	Small display	p value
score	20	20	-
# of structures	15.75	9	< 0.01
% complex structures	30%	4%	0.05
# of events	97	45	0.01
# of notes	16	9	0.03
avg. note length (chars.)	188	81	0.04

In addition to these reports, we collected a number of other types of data. Every interaction with the tool was logged, screenshots were captured every 20 seconds, and every session was video recorded. We also provided each participant with an eye tracker, which produced a video with the target of the participant’s gaze marked in it. Each session also concluded with an informal interview about how the participant approached the problem and how he or she used the workspace.

2.4 Demographics and recruitment

We recruited 16 participants (six male, 10 female), eight in each condition. We will label these participants L1-L8 (large display), and S1-S8 (small display). The participants had a broad collection of backgrounds, being comprised of five undergraduates, nine graduate students, and two staff members, and representing a variety of disciplines including international studies, sociology, math, biomechanics, and several engineering areas. Due to the time requirements, the study was conducted between subjects. Participants were paid for their time, with a bonus for the top score within the condition offered to create added inducement to solve the scenario.

One participant dropped out twenty minutes before the end of the study due to a headache brought on by the eye-tracking equipment. While we do not include her score, we do include evidence from her session in the rest of the results since she completed most of the task.

3 Results

Our approach to this study is primarily qualitative, based on our observations of the participants and analysis of the screenshots and videos we collected. However, in some instances, we have tried to supplement this with quantitative metrics in support of our observations. In all cases, since we only have two groups, we used a two-tailed, unpaired t-test to determine statistical significance.

3.1 Analytic scores

While this study was not really about performance, we did score the reports using the approach proposed by Plaisant et al. [22]. We assigned a point to every correctly identified person, location, or event, and subtracting a point for every incorrectly identified one (false positives). These numbers were then added together to form the final score. When tallied, we found no significant differences between the groups (both groups had an average score of 20, with a standard deviation of 5.8 for the small display group and 5.8 for the large display group), nor were there any correlations between the score and any other metric we collected (using bivariate analysis with the other metrics as well as two-way ANOVA with the metrics and display condition). Similarly, there were no effects from demographic differences (two-way ANOVA with display condition). It is worth noting that only two participants declared themselves finished with their analysis at the end of the study, and one of those reported exhaustion as the cause rather than satisfaction with her analysis.

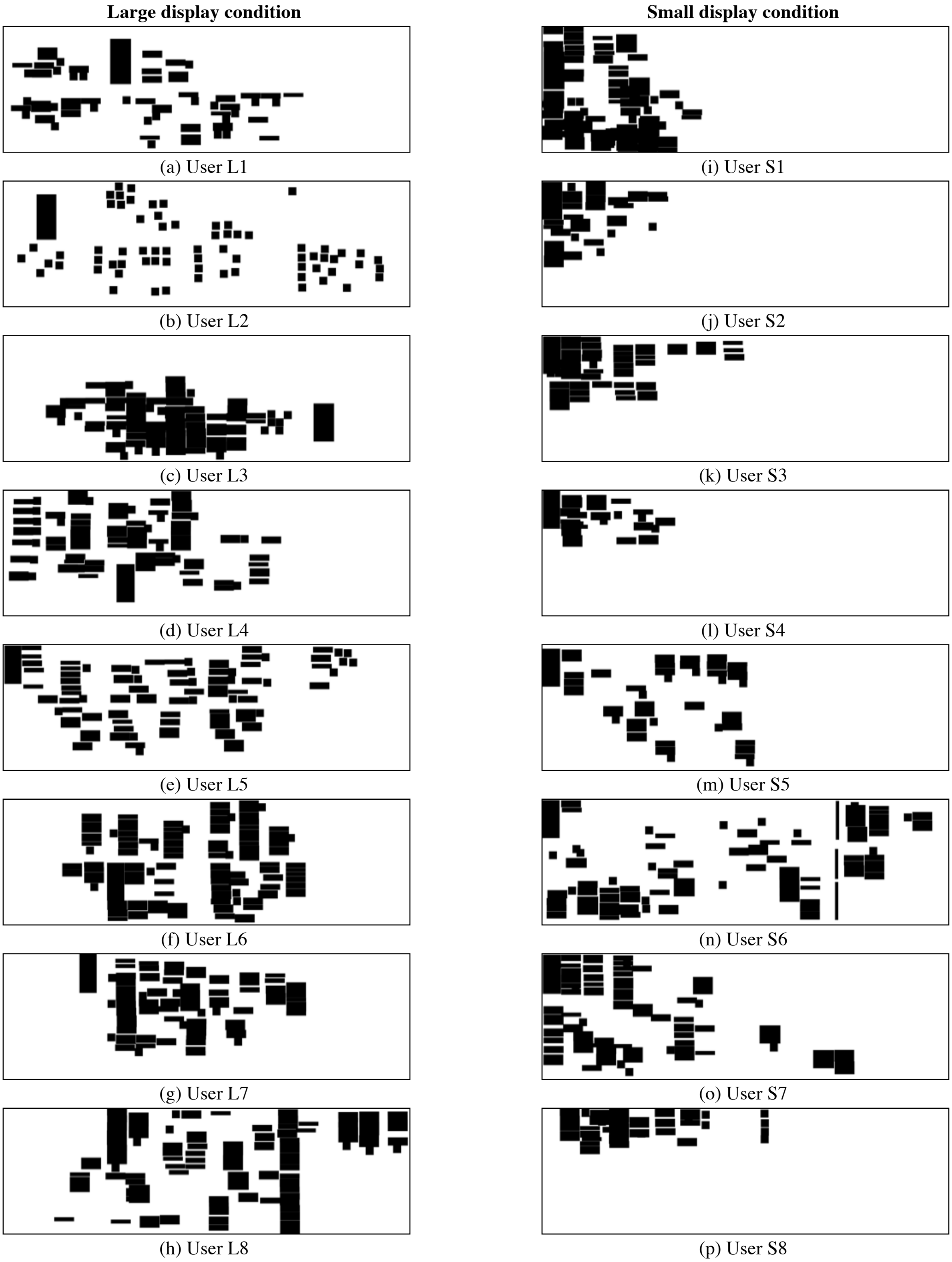


Fig.4. Simplified silhouette views of the final workspace state for all users. The large display condition is represented in figures a-h and the small display condition is represented by figures i-p.

3.2 Use of the space

As the main purpose of the freeform spatial workspace is to encode relationships through spatial proximity, we are interested in looking at the spatial structures created by these two groups. We used five indicators to identify a collection of documents as a structure. First, during the concluding interview, all of the participants described how they used the workspace for their investigation and how they interpreted each region of the space. Second, eleven of the participants left notes in their workspaces that labeled document collections, providing further insight into how they were grouping them. Third, we marked documents as belonging to a structure if they were brought together with intention during the investigation (i.e., the user read a document and then moved it to the proximity of another document in a different location in the workspace). Fourth, documents that shared a common fate (e.g., they were moved to a different location as a group) were marked as being a structure. Finally, we marked documents as being part of a structure if it was in close proximity to another document that was part of the structure, and there was some shared content between the document and the neighbor. Using these metrics, the screenshots were coded to generate a running tally of the number of structures at each point in time, their content and organization.

The two primary structures that we observed being constructed by both groups were timelines and clusters, which match previous observations of spatial sensemaking [1, 24]. The organizational bases for the clusters varied, but clusters based on geographic region or person were common. A number of structures were based on more than a single feature of the data. The most common combination was clusters that were internally organized in temporal order. Sub-clusters were also common. For example, several users created a cluster for activity taking place in the Bahamas and then within that structure made clusters for the two prime suspects who were operating there. The most complicated structure was a sprawling concept map developed by a large display user, in which documents were added to the structure based on some relationship to the already placed documents. We refer to any structure that either involves multiple spatial metaphors or an additional internal arrangement as being *complex*.

We begin our analysis by looking at the final state of the workspaces at the end of each session (Figure 4). That the use of space for sensemaking is a highly individualized process is clearly illustrated by these screenshots. While there are some common structural themes that run through these, they are all quite different. To further illustrate this, we will describe some of the more interesting examples.

In Figure 4(i), the S1 has created a loop of documents. This user made no attempt to use the space available, and instead just opened documents and let them lie where they were, moving the viewport whenever the view got too crowded. The loop was formed as the viewport arced down to the bottom of the display and then wrapped around to the left. In essence, this user quite literally worked himself into a corner. In contrast, S6 (Figure 4(n)) is one of the only small display subjects to attempt to make use of the whole available workspace. He used a structure based roughly on geographic relationships, with important documents corralled to the right. Most of the others made small clusters of related documents, but there were no other approaches that assigned any overall structure to the space.

On the large display, L3, our highest scoring user, produced aforementioned concept map that can be seen in Figure 4(c). Two of the users created timelines of important events and surrounded them with clusters based on important people, organizations, and locations (Figure 4 (d & g)). L6 (Figure 4(f)), on the other hand, split the space, with important people clustered above, and arranged by relationships, and events and places clustered below. The workspace shown in Figure 4(b) is a curious one in that the user (L2) made use of an overall geographic layout for the space, but rather than leaving the documents in place, she summarized each one into a note and then closed the original document. Unfortunately, she did not know what aspects of each document would be important and at the end only had a collection of names and places without any of the connecting events that tied them all together.

3.2.1 Spatial structures

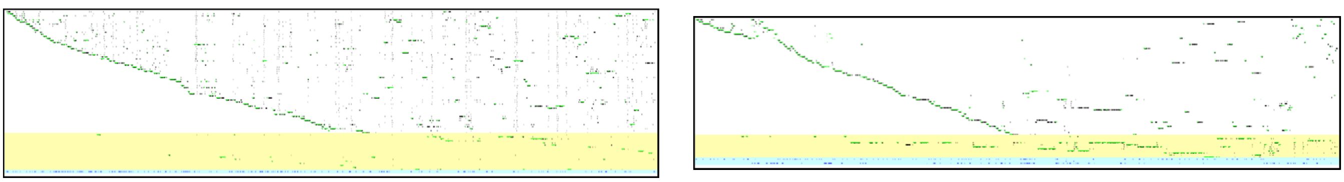
Visually, the difference between the two groups is fairly striking. S6 is an obvious exception (and to a lesser extent S5 and S7), but for the most part, visual inspection is sufficient to distinguish the two groups. However, we can also analyze the workspaces quantitatively.

In looking at the structures created by the subjects, the first thing we observe is that on average the large display group created 75% more structures over the course of the investigation (15.75 v. 9, $p < 0.01$). The structures produced by the large display users were also more complex, with 30% of the structures produced by the large display users combining spatial metaphors or integrating sub-clusters, while only 4% of the small display structures did the same ($p = 0.05$).

Unsurprisingly, the large display users also expended more effort organizing their documents. We looked at how the users moved documents and classified the movement based on how it affected the space. We consider an event to be structural if it moves an entire structure at once, or it changes the number of documents in a structure. We found that the large display users produced more than twice as many structural events on average than the small display users (97 v. 45, $p=0.01$).

3.3 Revisiting documents

In the previous section, we looked at how the space was used structurally. As the goal of externalizing information is to make it available perceptually so that it does not need to be retained in detail, we would like to know if the external representations are actually used. To assess this, we can look at what the user looked at during the session.



(a) Large display subject (L3) **(b) Small display subject (S5)**

Fig. 5. Visualization of eyetracking data for two subjects. The x-axis is time and each row represents the document (white background), note (yellow background), or tool (blue background) at which the subject is looking at that moment in time. The green traces indicate which document currently has input focus.

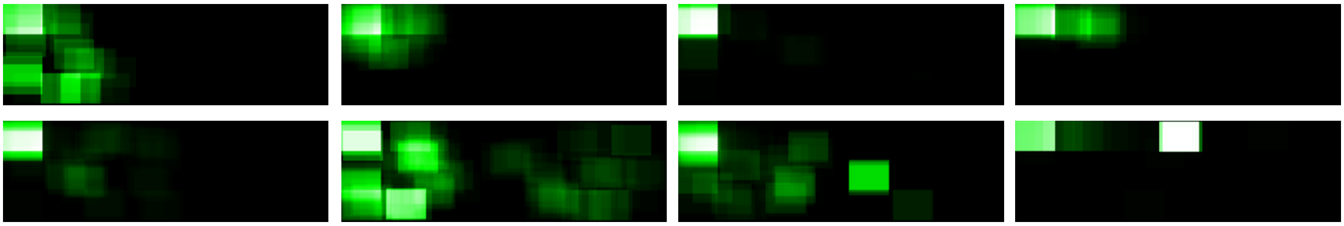


Fig. 6. Heatmaps showing the locations of the viewport for the eight small display subjects (top: S1-S4, bottom:S5-S8).

Our eye tracking system provides a video that approximates the user's view with the center of the user's gaze marked with a cross. We manually coded these videos using our screenshots to determine which document the user was currently focused on. Unfortunately, this is an extremely laborious process, and our tracker did not maintain calibration for some of our users. As a result, we were only able to code five of our users. For the rest, we can watch the videos, but do not have enough information to identify the documents being looked at. As a result, we can make some general observations, but we cannot quantify our observations.

The overall trend was that the large display users spent far more time revisiting documents and structures that were open in the space. We can illustrate this with the visualizations in Figure 5 based on the eye tracking data we did collect. In the visualization, each horizontal row represents a document or other object (note or tool) in the collection, sorted by type and then by title (the yellow region shows the notes, and the blue region is the tools). The black bars show where the user was looking, while the green bars show the document that currently has focus.

As can be seen, both subjects opened documents in approximately the same order, causing the initial diagonal. However, to the right of the diagonal, we can see that the large display user refers back to open documents far more often than the small display user. The thin, almost vertical lines that appear in the large display user's visualization illustrate moments when the subject was scanning the workspace, either looking for a particular document or just refreshing her sense of the workspace.

Also interesting is that for this small display user, window focus tends to follow gaze. This is because consulting already open documents usually required interaction, either selection to bring a buried document forward, or virtual navigation to access documents that were outside of the viewport.

To further illustrate this trend, we can look at data that we collected about the position of the small display user's viewport within the workspace. To examine this data, we created heatmaps that show how long portions of the workspace were visible during the session. The dwell time is encoded as a green ramp from black to white, with white indicating regions that were visible for longer periods. These regions can be compared to the layouts produced by the same users to demonstrate that in many cases, even when the workspace was used, the data was not often revisited. In three instances, there are very strong hotspots, indicating that most of the session was spent with the viewport fixed in one place. In the case of user S8 she reported after the session that she started trying to use layout to help manage information but she "gave up because it seemed like more effort than it was worth". After one third of the session had elapsed, she abandoned the documents she already had open, moved to an empty region of the workspace and never moved again.

3.4 Investigative approaches

In the previous sections, we have examined a number of ways in which the low-level behavior of the users was altered by the display environment. Unsurprisingly, this change was reflected in the overall approach to the task taken by the users.

To help illustrate the similarities and differences between the two groups, we will refer to another pair of visualizations that show the low-level window activation events (Figure 7). Again, each row corresponds to a single document, but this time, they are sorted by the order in which the user opened them. The horizontal bars show the period that the document is open. The bar is thicker when the document has focus, and the blue indicates that the document moved at some point between gaining and losing focus (movement is determined by looking at the bounds of the document at those two events). The small purple marks indicate that the document is a duplicate of the document above it.

Almost all of our users in both conditions generally followed a strategy Kang et al referred to as "overview, filter, and detail" [16], though none of our subjects adhered to it slavishly. We can see this in both of the visualizations. They are just marching through the documents in fairly rapid succession with the goal of reading all of the documents to get an overview. Half of our participants followed this approach to the end, while the others started pursuing leads and performing search after viewing enough documents to start identifying keywords.

Since the users were primed by the initial training, all of them, in both groups, began by using the space to store documents of interest. In most instances this meant clustering documents based on some surface feature that seemed distinctive, usually location, but sometimes person or activity.

An interesting difference between the two groups that manifested itself fairly early was that the entire large display group categorized all documents as soon as they were read. The small display group as a whole, however, frequently left documents where they opened unless the document seemed particularly interesting, was connected to other documents that were already categorized, or was related to another recently read document that had not yet been categorized. The result of this was that the small display users all formed piles of overlapping documents in the viewport.

Another difference was that the large display group made use of the whole workspace as they marshaled the documents. Unrelated documents would be quite far apart (multiple tiles of the display). The small display users, on the other hand, would just form separate piles within the viewport, moving the viewport when the current view was too crowded (with the exception of S6). As a result, S1-S4 and S8 all abandoned spatial organization before they had managed to read all of the documents.

3.4.1 Sensemaking

The more interesting division in behavior occurred as the participants started to move from just reading to trying to make sense of the data.

Both groups also had the same general approach at this stage — read documents in more detail, perform searches to help find related documents and create annotations (highlights and notes). However, with the exception of S6, the small display users mostly stopped interacting with their open documents. While S8 was the most extreme example, they all started to just perform searches, open the resulting documents, digest them or take notes, and then close them. S3 and S4 even started closing their already open documents.

The large display users, on the other hand, used the searches to identify documents that were already open, then closed the duplicate and read the original. The users clearly felt that the spatial location of the documents in the surrounding region added contextual information not available in the raw words that were already fully

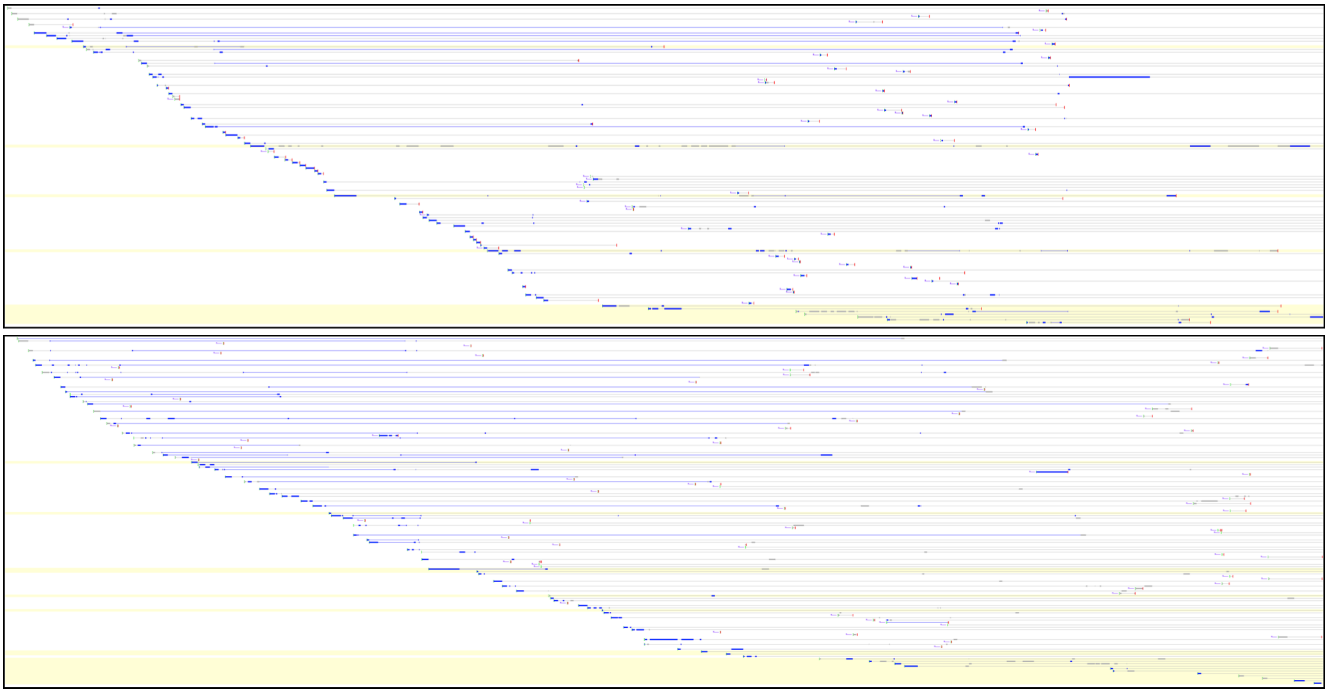


Fig. 7. Visualization of window activations and movement of documents (white) and notes (yellow). The horizontal bar shows that the document is currently open. The bars are thick when the window currently has focus, and blue if the window moved at some point during the marked interval. The green and red bars indicate open and close events, respectively, and the purple "L"s indicate that the document is a duplicate of the one above it. (top: S3, bottom: L3).

available in the second copy that they opened from the search results. Looking at the right sides of the two visualizations in Figure 7, we can see that both users opened a lot of duplicate documents, but the large display user closed them almost immediately once they had identified the document, while the small display user left them open longer and read them again. Many of the vertical lines in the eye tracking visualization are the result of this behavior.

The difference in the two groups also manifests itself in the note taking behavior. The large display group created approximately 75% more notes than the small display group on average (16 v. 9, $p = 0.03$) (L2 is an obvious outlier and was not included after confirming that she was a statistical outlier using Grubbs' test). The interesting bit, however, is that the use of these notes is quite different. The large display users created notes that labeled documents and space. These are short (a name or short summary of a document), and have no meaning if removed from the spatial context. L5 even used a number of notes with no text. She made them different colors and used them as a way to quickly show related documents that were on opposite sides of the workspace.

The small display users, on the other hand primarily produced narrative notes. They included long textual descriptions that summarized their understanding of what was going on. Rather than leaving these in place, the users moved them as they went, and they served the role of an external notebook. So, while the large display users generated more notes, the notes produced by the small display users were more than twice as long on average (188 v. 81 characters, $p=0.04$).

3.5 Reactions to the space

In our follow up interviews, one of the things we talked to the users about was the frustration that they had performing the task and the things that they would different. When we spoke to the other small display users, they reported having difficulty using the space well. S5 reported "at one point I tried to put an important document in the middle, but then I kind of forgot about it". S3 said that "when I wanted the specifics [the organization] didn't really help and I had to search". Finally, S4 said, "initially I thought that I would leave everything open and then I would remember where I put them… but it didn't work".

S6 was an exception to this, just as he was to most of our observations of the small display group. He made the greatest effort to use spatial organization to marshal the documents. He spent more time organizing his documents than anyone in the large display group, and from our interview with him at the end of the session it was obvious that he had formed a fairly strong high-level mapping of the entire workspace. His was also the lowest score among the small display users. This is explained when we compare his concept of where everything was in space (as described in his interview), and the documents we found. He made frequent mistakes in placing documents (i.e., they were misfiled according to his conception of the space). He also opened many duplicate documents, in some cases filing them right next to their doubles because "they seemed related". So, despite his efforts, the need to access the space virtually confounded him and impoverished his actual spatial sense.

When we talked to the large display users, we got fairly different answers. Most reported that they wished they had more time for organization. L8 wanted time to organize all of the documents very carefully to draw out all of the connections. L6 described a fairly complex approach that would start by opening documents faster and organizing them by person so he could start the process of putting things together faster. L4 and L7 both said that they would probably have paid more attention to dates and made more chronological structures. Only L3 didn't express desire for more time.

The striking thing here is that the physical environment appears to be biasing the large display group to continue to think in terms of physical layout, unlike the small display group. The large display users certainly asked for additional tools, such as visual search results that showed already open documents (L3), some way to draw connections across the space (L4 and L6) and some larger note taking facilities (L4, L6, and L7). However, with the exception of the note-taking tool, these are all additional spatial tools.

4 Discussion

The net outcome of these results is that we did indeed observe behavioral differences between the two groups that ranged from low-level behavior such as how many documents were open and how they were spread across the space to high-level differences in approach to the task.

4.1 External representations

The primary difference between how the two groups approached the problem is the form of the externalizations that they produced. Heuer has identified externalization as an important technique for helping to clarify the main elements of a problem, identify relationships and reduce the effect of cognitive bias [12], so any influence of the environment on how they are created would be important.

The large display users seemed to view the space as a more cohesive whole, arranging not just documents with relation to one another, but whole structures as well. Any textual notes were primarily used as labels on the space. The smaller display users created more narrative notes and while they also performed some document organization, it was on a smaller scale.

This difference is important because the form of the externalization affects how it is used, which can fundamentally change the nature of the task being performed. While the most obvious role of external representations is to act as a memory aid, they can also be conscripted into the cognitive process and used directly, serving to anchor or structure cognitive behaviour [32].

Kirsh makes the claim that interacting with external representations makes cognition more efficient (faster), and more effective (able to cope with harder problems). He identifies several attributes of external representations that he claims "may increase the efficiency, precision, complexity, and depth of cognition": persistence, reordering and reformulating [18]. Persistence refers to the stability and material properties of external representations, which can be relied upon to stay fixed unless explicitly acted upon, unlike internal representations. This supports reordering, where the externalizations can be rearranged for comparison and construction. This makes it easier to perceive relationships between representations. Reformulation, on the other hand, is the change of representations to illuminate properties of the represented object or idea by making them more explicit.

While Kirsh and Zhang were contrasting external representations to internal representations, we believe these same properties can be used to compare different kinds of external representations. The narrative notes favoured by the small display users are indeed external representations, but they lack a great deal of flexibility. The textual representation requires the user to be more explicit and limits the degrees of relationships that can be quickly expressed.

The structures formed by the spatial organization of the documents, however, are explicitly formed to permit easy reordering and reformation. Ordering the documents across the space is the task, with understanding emerging from the relationships that are expressed this way. While all of the users started by making spatial structures (since they were initially primed to), it was the large display users who stuck with it, with only two of the small display users continuing to make spatial structures throughout.

It is also important not to overlook the use of the external representations to expand memory. By using the documents themselves as components of the external representations, much more detail can be encoded into the representation. The user can regard the document as a simple label or entity for the purposes of structuring, but the full details are also immediately available. When the users created narrative notes, they were creating representations with no link back to the source material, thus increasing the cost of obtaining details that weren't recorded. L2, the large display user who converted all of her documents to summarized notes (Figure 4(b)), provides interesting evidence of this in action. Her organization was quite good, but the low level artefacts had no details in them, so she couldn't make any of the important connections (nor did she even seem to realize that there were connections not represented). As a result, she received the lowest score in the large display group, largely because she simply did not have the information she needed to solve the scenario at the end.

There is, of course, a role for the more narrative representation, however, we would argue that this more structured and explicit representation is more appropriate for later in the sensemaking process. This is a representation better suited for communication and could be seen as a reformulation of the more flexible spatial structures used during the exploratory stages of the investigation.

4.2 Role of navigation

We have shown that despite having the same spatial workspace, the two groups used the space differently and created different types of externalizations. It is useful to consider the source of this difference.

We attribute most of the differences to the physicality of the large display, and the fact that it allows the user to rely more heavily on embodied resources such as vision and proprioception. There is growing support for the idea that cognitive processes work to minimize effort [6, 11, 13, 18]. Gray introduced the idea of "soft constraints", which are essentially constraints on an interface imposed by low-level costs of accessing information [11]. He demonstrated that even small differences in the time required to access information would lead users to rely on imperfect or partial information if it was faster to access.

When we consider the behaviour of the two groups, it becomes obvious that the cost to access information in the workspace was quite a bit higher when virtual navigation was required. For example, looking back at Fig 7, we observed that the small display user was activating windows before almost every gaze event, implying an extra action that was required by the small display user to access the documents. We would also point again to the fact that with only two exceptions, all of the structures created by the small display users were constrained to fit within the bounds of the display (rather than the space). The large display users all created structures that spanned multiple tiles of their display, indicating that the viewport offered by the small display was the constraining factor, which would be the point where those users would be forced to switch from physical navigation to virtual navigation.

It is possible to attribute this difference entirely to the efficiency of physical navigation. However, in some instances, virtual navigation should be faster. Our environment allowed the user to navigation and move documents using the overview, so moving around the space or placing objects is potentially faster than physically moving, and the difficulties of working with a mouse over long distances is well-known [23]. As such, it would be fair to believe that virtual navigation would be more efficient in some cases.

However, we need to consider the problem of space constancy (the mechanism by which the world appears stable). While the precise mechanism is still subject to debate, it seems that some combination of the visual system and proprioception is used in the visual cortex to create the perception of a stable environment even as we move our eye, heads, necks, and bodies [14, 20, 29]. This low-level mechanism supports our spatial awareness and understanding. This potentially changes the cost structure of accessing information through virtual navigation. While the actual motor cost may be lower for virtual navigation, there may be a higher cognitive cost since the small display users cannot rely on this low-level mechanism for maintaining space constancy and must instead rely on the reduced size overview and their own internal model of the space.

A related factor is the physical presence of the display itself. Tan demonstrated that even if a display occupies the same portion of the visual field, larger displays still bias the user towards adopting an egocentric frame of reference with regard to the objects displayed on it, thus improving performance on spatial tasks [28]. As such, it is possible that the large display environment allows the user to relate the space to her own body sense, creating more cognitive affordances that can be conscripted as memory aids [21].

4.3 Implications for future research and design

As we previously noted, leveraging human spatial abilities as a sensemaking tool has been touted by many researchers and developers (e.g., [19, 25, 31]). However, this study suggests that to properly tap into those abilities, the environment should engage the user's embodied resources, which are already accustomed to perceiving and interacting with spatial information. From the perspective of embodied cognition, we might go further and say that human "spatial abilities" are rooted in the interaction of body and world, so that it not surprising that decreasing the fidelity of that interaction would have a corresponding effect on how readily a user might make use of space. We will note that this is not an argument for large displays exclusively; any related technology that engaged the user with a physical environment is likely to exhibit similar affordances.

Given our observation about structures being limited by the viewing region, it is possible that the size of the representations we used in the study (short documents) played a role in any difficulties the small display users encountered in organizing their data. Less space filling representations such as icons would have allowed the small display users to fit more representations on the display, potentially leading to larger, more complex structures. However, it is not clear how the reduced information content of icons would affect their utility as sensemaking tools. This clearly warrants more study, but again user L2 (who summarized all of the documents onto notes) provides a valuable caution. Her experience illustrates the importance of having detailed information available, as well as the need for evidence marshaling techniques to evolve with understanding of the problem.

On a more pragmatic side, the study indicates the need for new tools for working with the space, such as the visual search tool requested by L3. Some of these we have already addressed in our development of Analyst's Workspace [3].

5 Potential Issues and Future Work

There are, of course, some issues and open questions raised by this study. One potential confound is the reliance on the large display to provide the physical navigation condition. There are other technologies, such as HMDs, that allow for physical navigation, but they do not provide the resolution to do text analytics at this scale. As such, it is difficult to separate effects of the large display from the use of physical navigation. As mentioned earlier, Tan has demonstrated that just the size of display has an effect on perceptions [28]. This implies that there is more than pure physical navigation that creates the spatial environment, but it is not clear that there is a way to meaningfully separate the two, or that other solutions for providing a "human scale" environment [2] would not provide similar cognitive cues.

It could also be argued that the environment we developed favored the large display, as the full text documents can fill the small display rapidly, without much opportunity for the interesting construction of spatial structures within the confines of the viewport. However, part of the point was to juxtapose virtual and physical navigation, and any interaction that was required to access the contents of a document would have added virtual navigation to both conditions. More importantly, the very fact that the environment *could* favor the physical space condition only serves to strengthen the assertion that virtual navigation is not a direct replacement for a physical environment.

Another potential issue is the presence of the bezels on the large display. It could be argued that they provided navigational waypoints or organizational bins on the large display that were not present on the small display. We have certainly seen some evidence in the past that the bezels can be used as an organizational structure. However, in this study, every large display user created a structure that spanned across at least one bezel, suggesting that they were not a strong constraint on the space. The only real alternative would have been to place a grid on the spatial workspace for the small display group. However, we opted not to do this since we considered a grid built into the interface to be a stronger constraint than the obvious artifacts of technological limitations.

The inexperience of our subjects is another issue. The inexperience and different analytic abilities of our users certainly played a role in the analytic scores that we recorded. However, the goal of the study was not to see if a large display just made a better analytic tool. Our question was more fundamental. We wanted to explore how physical and virtual navigation affected the conscription of space as a sensemaking tool. To that end, the experience of our subjects is less important beyond their ability to read and understand what they read. The strategies for using space is of less interest than how fundamental the space became to their process.

A more serious concern is the cognitive effect of managing large quantities of information spatially. This is, of course, more of a general issue with the approach than a problem with the study. It seems clear that cognitive cycles must be used to maintain spatial understanding and to perform spatial organization. In about half of our large display subjects we observed some evidence that documents were being organized and reorganized without any real understanding being developed by the process. It is possible that the large display subjects actually hurt their scores by spending too much time reorganizing rather than actually reading and making new connections.

We continue to pursue this particular approach to sensemaking because we feel that the cognitive overhead of managing the space is outweighed by the cognitive overhead of managing large information collections internally, or with more linear or structured externalization techniques. However, this is clearly an area that needs more research.

The last question we should address is the issue of scalability. Virtual navigation allows the user to work with workspaces of any size, while strict adherence to physical navigation imposes fairly clear boundaries on the amount of information that can be displayed. Our interest is in demonstrating the fundamental difference between the environments rather than proposing a complete sensemaking environment. Actual applications may need to employ a hybrid approach that integrates a mixture of physical and virtual navigation or aggregation techniques. However, as indicated by Jakobsen and Hornbæk [15], it seems clear that future research will need to look into how the integration of these will affect the environment.

6 Conclusion

Physical navigation is fundamentally more embodied than virtual navigation, which requires technological mediation and internal mappings to maintain spatial understanding. In this study we have shown how physical navigation led to the development of more effective externalization. Users of the large display made use of more of the available space, treated the workspace as a more coherent whole, and created more complex structures. The environment also biased the users into adopting a spatial view, a perspective dropped by most of the small display users. Using evidence from how spatial structures were maintained and documented, we showed evidence that users were coupling with the documents in the physical space, conscripting them into their cognitive process in a way not evident in the virtual space condition.

It is true that we did not find a performance difference between the two environments. However, this was an exploratory study with the goal of looking for a difference between the two environments, which we did find. We think that this is a valuable contribution to the ongoing conversation about the nature of physical and virtual spaces.

The virtual workspace is still a useful technique for interacting with large information spaces, but this study has demonstrated that all space is not equal. Physical navigation cannot simply be replaced with virtual navigation techniques with the expectation that the user's spatial abilities and thus the facility to leverage the space to externalize information will carry through identically.

7 Acknowledgments

This research was partially supported by NSF grants CNS-1059398, CCF-0937133 and IIS-1218346.

References

[1] Andrews, C., Endert, A., and North, C.: "Space to think: large high-resolution displays for sensemaking". In *CHI*, Atlanta, GA, 2010, pp. 55-64.

[2] Andrews, C., Endert, A., Yost, B., and North, C.: Information Visualization on Large, High-Resolution Displays: issues, Challenges, and Opportunities, Information Visualization, 10, (4), pp. 341-355, 2011.

[3] Andrews, C., and North, C.: 'Analyst's Workspace: An embodied sensemaking environment for large, high-resolution displays', in *VAST*, Providence, RI, 2012, pp. 123-131.

[4] Ball, R., North, C., and Bowman, D.A.: "Move to improve: promoting physical navigation to increase user performance with large displays". In *CHI*, New York, NY, 2007, pp. 191-200.

[5] Bernstein, M.: "Collage, composites, construction". In *HYPERTEXT '03*, New York, NY, USA, 2003, pp. 122-123.

[6] Clark, A.: (2008)*Supersizing the Mind: Embodiment, Action, and Cognitive Extension*, Oxford University Press.

[7] Cockburn, A., Karlson, A., and Bederson, B.B.: A review of overview+detail, zooming, and focus+context interfaces ACM Comput. Surv., 41, (1), pp. 2:1-2:31, 2009.

[8] Czerwinski, M., Horvitz, E., and Cutrell, E.: "Subjective Duration Assessment: An Implicit Probe for Software Usability". In *Proceedings of IHM-HCI 2001 Conference*, Toulouse, France, 2001, pp. 167-170.

[9] Czerwinski, M., Tan, D.S., and Robertson, G.G.: "Women take a wider view". In *CHI '02*, New York, NY, USA, 2002, pp. 195-202.

[10] D. Austin Henderson, J., and Card, S.: Rooms: the use of multiple virtual workspaces to reduce space contention in a window-based graphical user interface, ACM Trans. Graph., 5, (3), pp. 211-243, 1986.

[11] Gray, W.D., and Fu, W.-T.: Soft constraints in interactive behavior: the case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head, Cognitive Science: A Multidisciplinary Journal, 28, (3), pp. 359--382, 2004.

[12] Heuer, R.J.: (1999)*Psychology of Intelligence Analysis*, Center for the Study of Intelligence.

[13] Hollan, J., Hutchins, E., and Kirsh, D.: Distributed cognition: toward a new foundation for human-computer interaction research, ACM Trans. Comput.-Hum. Interact., 7, (2), pp. 174--196, 2000.

[14] Jaekl, P.M., and Harris, L.R.: Space constancy vs. shape constancy, Seeing Perceiving, 23, (5-6), pp. 385-399, 2010.

[15] Jakobsen, M.R., and Hornbæk, K.: "Sizing up visualizations: effects of display size in focus+context, overview+detail, and zooming interfaces ". In *CHI '11*, 2011, pp. 1451-1460.

[16] Kang, Y., Gorg, C., and Stasko, J.: "Evaluating visual analytics systems for investigative analysis: Deriving design principles from a case study". In *VAST '09*, 2009, pp. 139 -146.

[17] Kirsh, D.: The intelligent use of space, Artif. Intell., 73, (1-2), pp. 31-68, 1995.

[18] Kirsh, D.: Thinking with external representations, AI Soc., 25, (4), pp. 441-454, 2010.

[19] Marshall, C.C., and Rogers, R.A.: "Two years before the mist: experiences with Aquanet". In *ECHT '92: Proceedings of the ACM conference on Hypertext*, New York, NY, USA, 1992, pp. 53--62.

[20] O'Regan, J.K.: Solving the "real" mysteries of visual perception: the world as an outside memory., Can J Psychol, 46, (3), pp. 461-488, 1992.

[21] Patten, J., and Ishii, H.: "A comparison of spatial organization strategies in graphical and tangible user interfaces". In *DARE '00*, New York, NY, USA, 2000, pp. 41-50.

[22] Plaisant, C., Grinstein, G., Scholtz, J., Whiting, M., O'Connell, T., Laskowski, S., Chien, L., Tat, A., Wright, W., Gorg, C., Liu, Z., Parekh, N., Singhal, K., and Stasko, J.: Evaluating Visual Analytics at the 2007 VAST Symposium Contest, Computer Graphics and Applications, IEEE, 28, (2), pp. 12-21, 2008.

[23] Robertson, G., Czerwinski, M., Baudisch, P., Meyers, B., Robbins, D., Smith, G., and Tan, D.: The Large-Display User Experience, IEEE Comput. Graph. Appl., 25, (4), pp. 44--51, 2005.

[24] Robinson, A.C.: "Collaborative synthesis of visual analytic results". In *VAST*, 2008, pp. 67-74.

[25] Shipman, I., Frank M., Hsieh, H., Maloor, P., and Moore, J., Michael: "The visual knowledge builder: a second generation spatial hypertext". In *HYPERTEXT*, New York, NY, 2001, pp. 113-122.

[26] Shipman, I., Frank M., and Marshall, C.C.: "Formality Considered Harmful: Experiences, Emerging Themes, and Directions on the Use of Formal Representations in Interactive Systems", CSCW '99, 8, (4), pp. 333-352, 1999.

[27] Shupp, L., Andrews, C., Dickey-Kurdziolek, M., Yost, B., and North, C.: "Shaping the Display of the Future: The Effects of Display Size and Curvature on User Performance and Insights", ACM Trans. Comput.-Hum. Interact., 24, (1&2), pp. 230 - 272, 2009.

[28] Tan, D.S., Gergle, D., Scupelli, P., and Pausch, R.: Physically large displays improve performance on spatial tasks, ACM Trans. Comput.-Hum. Interact., 13, (1), pp. 71-99, 2006.

[29] Tatler, B.W., and Land, M.F.: Vision and the representation of the surroundings in spatial memory, Phil. Trans. R. Soc. B, 366, (1564), pp. 596-610, 2011.

[30] Thomas, J.J., and Cook, K.A. (Eds.). (2005) *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, (National Visualization and Analytics Center).

[31] Wright, W., Schroh, D., Proulx, P., Skaburskis, A., and Cort, B.: "The Sandbox for analysis: concepts and methods". In *CHI*, New York, NY, 2006, pp. 801-810.

[32] Zhang, J.: The Nature of External Representations in Problem Solving, Cognitive Science, 21, pp. 179-217, 1997.