Visual Analytics on Large Displays: Exploring User Spatialization and How Size and Resolution Affect Task Performance

Gökhan Çetin, Wolfgang Stuerzlinger, John Dill
School of Interactive Arts + Technology (SIAT)
Simon Fraser University
Vancouver, Canada
{gcetin,w.s,dill}@sfu.ca

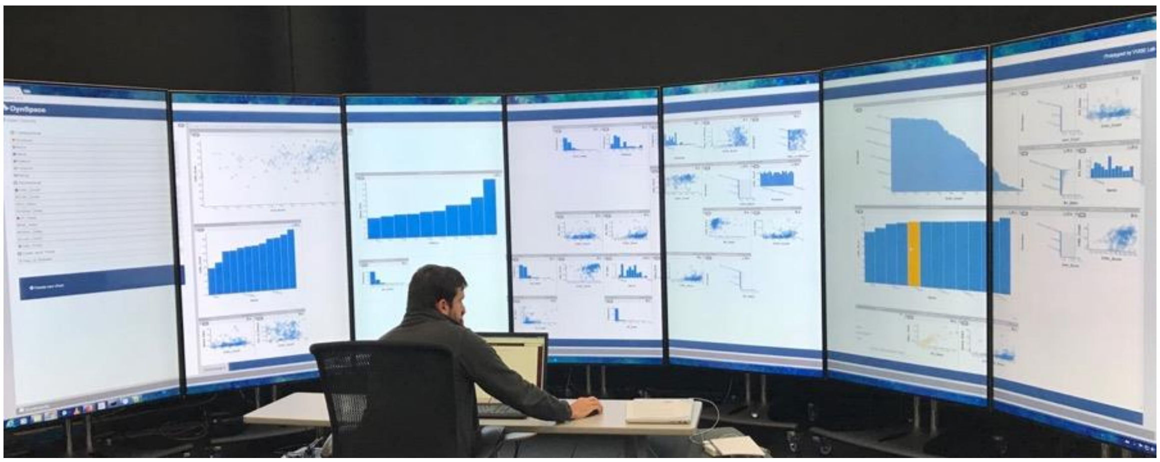

Fig. 1. A participant doing a VA Task on V4-SPACE, our large high-resolution display system.

Abstract—Large, high-resolution displays (LHRDs) have been shown to enable increased productivity over conventional monitors. Previous work has identified the benefits of LHRDs for Visual Analytics tasks, where the user is analyzing complex data sets. However, LHRDs are fundamentally different from desktop and mobile computing environments, presenting some unique usability challenges and opportunities, and need to be better understood. There is thus a need for additional studies to analyze the impact of LHRD size and display resolution on content spatialization strategies and Visual Analytics task performance. We present the results of two studies of the effects of physical display size and resolution on analytical task successes and also analyze how participants spatially cluster visual content in different display conditions. Overall, we found that navigation technique preferences differ significantly among users, that the wide range of observed spatialization types suggest several different analysis techniques are adopted, and that display size affects clustering task performance whereas display resolution does not.

Index Terms—Visual analytics, large high-resolution displays, spatialization, clustering, visualization, space in analytics

1 Introduction

When dealing with real world problems and due to the strong trend towards automatizing data collection processes, data volume and complexity are ever increasing, and inherently complex datasets are more frequently encountered. While *Big Data* can yield powerful insights, analysis is required to generate such insights. We categorize Big Data analysis approaches into two main groups: *algorithmic analysis* and *human-in-the-loop*. While powerful, algorithmic approaches still cannot deal with problems needing human-in-the-loop analysis. Visual Analytics, an area studied in this work, applies interactive visual analysis to such problems.

Visual analytics (VA) is a science-based activity supporting sense-making of large, complex datasets through interactive visual data exploration [39]. In VA, users reason and make sense of data through interaction with visualizations of the data. Related tools include those from Tableau, Microsoft, Qlik and others.

In VA applications using normal desktop monitors, interacting with visualizations of large volumes of info is difficult, making spatialization and semantic interaction awkward and difficult [14, 16, 17]. Radically increasing the display size changes this dynamic substantially, allowing users directly access more information and to directly arrange documents spatially to convey data relationships [16, 26]. Large, high-resolution displays (LHRDs) have been shown to improve productivity [1,5,12,49] over traditional desktop monitors and we can expect this to hold for VA applications as well.

We think of spatialization in VA as a 2D view of high-dimensional data where the pieces of information can be organized in the 2D view, spatially, to indicate various relationships such as similarity, and would expect LHRDs to support such activity better than typical desktop monitors. Large displays have usability issues though, such as cursor location [6], access to distal parts of the display and icons [9,23], and window management [8,31]. There is a need then for more research on LHRDs on spatialization issues, analytical task performance and usability issues, which together form the main motivation for this work. Correspondingly, our research goals are: (a) to observe users' spatialization strategies and semantic interaction with information in an analytical workspace on a very large display system, and (b) to study the effect of size and resolution of the display system on visual analytics task completion success. Our research questions (RQs) are:

RQ1: How do people use the (large) space of an LHRD to cluster items while trying to solve analytical problems that require re-organization of the content? More specifically, we ask which users consider when clustering; topical relations, visual similarity, or common dimensions of charts?

RQ2: Do larger displays help users to do better in VA tasks? How does task success change as screen size is increased?

RQ3: How does task success change as resolution is increased? Does higher resolution improve task success in VA tasks?

Our first study addresses RQ1. In this study, spatialization on LHRDs is not compared to desktop monitors. Since LHRDs are characteristically different from small monitors, we explore how spatialization takes place on LHRDs. To observe clustering behavior on large displays, we gave users a VA task involving 2D data charts.

The second study addresses RQ2 with a within-subjects design. RQ3, on the effect of display resolution on VA task success, used a between-subject design where subjects were randomly assigned to three groups, where each group was given one of 720p, 1080p or 2160p (4K) resolution on the same underlying physical displays to investigate the effect of higher resolution on task success.

The main contributions of this paper are:

- A study exploring navigation methods as well as spatialization and manual clustering behaviours of users during VA tasks on LHRDs. We observed large individual differences in both preferred navigational methods (physical vs. virtual) and spatialization methods.
- A study evaluating and analysing VA task successes under different display size and resolution conditions. We found that while large displays help, there are again large individual differences. Display resolution appears not to have a significant affect.

The rest of the paper is organized as follows: Sections 2 and 3 describe related work and the apparatus. The two studies conducted for this work are described in sections 4 and 5. Section 6 contains a discussion, summary and conclusions.

2 Related Work

A Virginia Tech study investigated the effects of information layout, screen size, and field of view (FOV) on user performance in an Information-Rich Virtual Environment (IRVE) [29]. IRVEs consist not only of 3D graphics and other spatial data, but also include more abstract or symbolic information related to the space [10]. The authors designed and evaluated two information layout techniques to support search and comparison tasks. They compared a single monitor and a tiled nine-panel large display in a between-subjects design. For the evaluation, users were timed, tracked for correctness, and asked to rate both difficulty and satisfaction on each task. The authors posed a set of research questions, including one in which we were particularly interested: “Do the advantages of a single layout space hold if the screen size is increased?” The best performing layout technique on a small display was “Viewport Space”; however, it was outperformed by another called “Object Space” on the nine-screen display [29]. As the study suggested, large displays are different environments than conventional desktop monitors, with their own set of challenges and advantages, which need to be identified by LHRD studies.

2.1 Content Spatialization on LHRDs

Previous research indicated that clustering information reduces the amount of visual search needed to find the elements required for problem-solving [34]. Another study shows that a spatial contiguity effect applies to how deeply a user learns [39], i.e., that students learn more deeply when extraneous material is excluded rather than included, and when related content is placed near the item being considered rather than far from it. This learning effect is relevant, since training is required for participants using a system for the first time [36], especially for analysis tasks that require tool understanding and analytical thinking patterns [25].

Endert et al. presented the results of a study where users performed a spatial sense-making task on a LHRD (a 5x2 grid of 17” LCDs) in the LightSPIRE system [16]. Our current work shares much common ground and methodology with this study. We are primarily interested in the analysis of spatial layout of visual information and about what structure exists within the user-generated clusters. Participants used three distinct patterns of spatial organization: Topical, temporal and hybrid clustering [16]. Most used topical clusters; while a third used temporal information when organizing their workspace. One participant used hybrid clustering to balance temporal awareness and understanding of topical relations. Other interesting clustering characteristics not used by Endert et al. include intra-cluster co-occurrences and transitivity – how connected different clusters are. Also of interest would be a measure of cluster independence and whether positioning or inter-cluster distance symbolizes any relation.

Spatial organization does not necessarily take place the same way on large displays [2]. One argument for this is that, compared to conventional monitors, the required type of navigation within the content is different. Previous work has shown that users spatially organized content quite differently when there was a physical workspace (which requires more physical navigation) such as a LHRD, as compared to a virtual workspace, where the users must navigate using virtual navigation techniques [2].

Andrews and North [2] explicitly examined the use of spatial sense-making techniques within two environments – a 33-megapixel LHRD and a small desktop monitor. Their study task used a basic analytic tool relying on manual spatial organization as the primary evidence marshalling technique [1,32]. The results demonstrate that the two approaches for providing a sense-making space, physical and virtual workspaces, are not equally effective, and that the greater embodiment afforded by the physical workspace changes how the space is perceived and used. Dourish explains the concept as follows: “Embodiment denotes a form of participative status. (It) is about the fact that things are embedded in the world, and the ways in which their reality depends on being embedded. So, it applies to spoken conversations just as much as to apples and bookshelves; but it’s also dividing a line between an apple and the idea of an apple [13]”.

Spatial representations of data on large displays are expected to aid understanding. However, the question of *what properties of a spatial representation significantly support cognitive processing* is still open. In this context, Ragan et al. explored how spatial layout and view control impact learning, by investigating the role of persistent visibility when working with large displays [30].

Their first experiment investigated how memory (detail recall) and learning (comprehension) performance are affected by spatial distribution of information in a visual presentation, while viewing was either automatic or interactively controlled. They hypothesized that a spatially distributed layout would support superior learning in comparison to a non-spatial one, i.e., a slideshow layout in which new information replaces with the previous, and that interactive, user-controlled viewing would improve task performance over non-interactive, automated viewing. Surprisingly the results did not support their hypotheses. Learning scores were significantly lower in the distributed layout than in the slideshow-style presentations [30]. Also, there were no significant differences due to the viewing mode (automatic vs. user-controlled).

A second experiment studied the effects of persistent visibility, where in contrast the images did not disappear after having been shown for a while. The results showed that learning scores with the persistent-visibility distributed layout were superior to the automatic and interactive distributed presentations, as hypothesized [30].

Thus, interestingly, spatial distribution of the information did not help learning on its own, but it did help with persistent visibility.

2.2 Impact of Display Size and Resolution

Others have also considered aspects related to our RQ1. Reda et al. present the results of a small-scale study to understand how display size and resolution affect insight. Although their results verify the generally accepted benefits of large displays, they also provide mixed results for extant work and propose explanations by considering the cognitive and interaction costs associated with visual exploration [48].

Arguably the closest work to our second study is Ni et al.'s paper suggesting that increased display size and resolution improve task performance in IRVEs [27]. They showed that among a high-resolution small monitor, rear-projected screen, and a tiled high-resolution display system, users were most successful at search and comparison tasks on LHRDs. In addition, users working with large displays became less reliant on wayfinding aids in acquiring spatial knowledge [27]. While Ni et al. investigated 3D spatial performance, which is different from the VA activities that are the focus of this work, the research questions are very similar. Thus, we follow a similar approach in terms of exploring the effects of physical screen size, display resolution and spatialization effects. In their work, the authors isolated display size and resolution as independent variables by using three different display technologies: A high-resolution small desktop display; a single projector provides a large, low-resolution display; and an array of projectors produces a LHRD.

Ni et al.'s experimental findings demonstrated the advantages of increased size and resolution [27]. As a general guideline, the LHRD was the preferred choice for IRVE applications, since it facilitates both spatial navigation and information gathering. Interestingly, a large low-resolution rear-projected screen outperformed a regular-sized monitor with a higher resolution. Our study tests this phenomenon in a different way but with a similar study design. The authors also ask in their future work section if there is an upper bound beyond which users will be overwhelmed by the displayed information. We investigate this aspect in terms of "information density".

3 Apparatus

The apparatus for our studies consists of physical displays (V4-SPACE), a software tool (DynSpace+), and datasets to be analyzed.

3.1 Display System: V4-SPACE

Our research uses a very large display system called V4-SPACE to support our research on how LHRDs can support VA tasks for a single user. Our goal was to observe and analyze how physical size, resolution, and content spatialization on the workspace affect user performance on LHRD in VA tasks.

V4-SPACE consists of a 1x7 array of large, tiled monitors, each a vertically oriented 85" 4K Samsung Smart TV. An additional 21" monitor sits in front of the user (visually below the view of the monitors) and is used for *auxiliary tasks*. The main display has a total of 15120 x 3840 pixels, which makes V4-SPACE a 58-megapixel display system with 52 PPI pixel density. While the aspect ratio of a single display is 9:16, the system ratio is 63:16, about 4:1. The main display is 7.41 m by 1.88 m. With the 1x7 grid, there are no horizontal bezels and only 6 vertical ones.

V4-SPACE is driven by a single computer with an Intel i7-6700K processor at 4GHz with four PCI Express Gen 3 slots. The displays are driven by two nVidia Quadro M5000 cards, which provide four 4K outputs each, hardware synchronized through a Quadro Sync card. The auxiliary desktop monitor was also present but used only to support answering questionnaires.

The system is designed for a single user who has a fixed position in front of the display system, about 3.3 m from each monitor. At this distance V4-SPACE is a "super-retina" display. The monitors are arranged in a circular arc (~131° horizontal field-of-view, FOV) such that each monitor is equidistant to the user. This avoids information legibility issues due to non-uniform distances. As a limited form of physical navigation, the user may rotate a swivel chair to look at different parts of V4-SPACE.

Interaction is through keyboard and mouse. To support the high resolution of V4-SPACE, we use a Razer DeathAdder Chroma 10000 PPI optical gaming mouse. We used this mouse with a high gain factor on mouse acceleration. Such a combination is the best available input option: it enables the user to access every corner of the display with a (large) wrist movement, while still affording pixel accurate pointing.

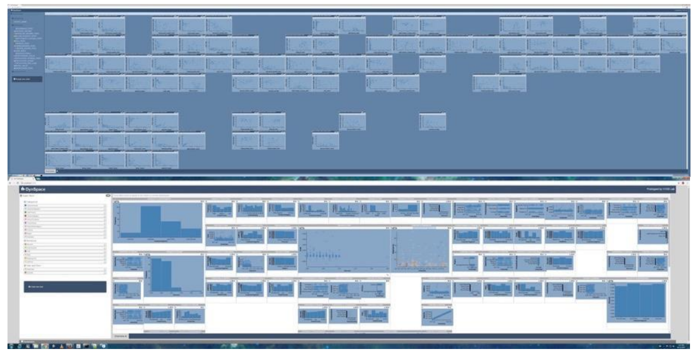

Fig. 2. Upper image: Simplified version of DynSpace+ used in Study I to observe content spatialization and how users cluster charts. Lower image: version of DynSpace+ used in Study II, with enriched analysis features.

3.2 Visual Analytics Tool: DynSpace+

The tool used for this work was DynSpace+, with slightly different versions used for studies I and II (Fig. 2). The first study on exploring users' spatialization and information clustering approaches on LHRDs used a simplified version; all inapplicable functionality was removed to avoid distractions during the classification tasks. All analytical functionalities were available for the sense-making tasks that they were the focus of our second study (Fig. 2).

DynSpace+ is a browser-based VA tool written in JavaScript. There are two panels: a Data Panel on the left side showing data dimensions, and a main visualization panel on the right. Data dimensions are classified as categorical, numerical, or date & time. Users can select data dimensions, drag and drop them into charts in the visualization panel for analysis.

The visualization panel contains data charts to explore relations between selected data dimensions. A chart in DynSpace+ is a 2D data plot (currently a scatterplot, histogram, bar chart, or row chart). Charts are in rectangular sub-panels that a user can move, resize, add, or remove. Users can filter charts in multiple ways. Through filters, users can focus on details, separate parts of data, and/or remove outliers. Filters can be applied either globally to all existing charts or locally to selected charts. All charts are coordinated through brushing and linking.

DynSpace+ was designed to aid analysis of complex datasets – to gain insights, or answer questions about the data. Sample task questions include: "Do critics and ordinary users usually have similar opinions of a video game?" (videogame dataset), "How high are the shipping costs for a delivery truck as compared to air shipping?" (superstore dataset), "Among the states in which females are <50% of the population, which state has the highest percentage of Hispanics?" (US Census data).

DynSpace+ initially displays a number of charts, each generated automatically by picking random data dimension pairs. These initial charts use about half of the V4-SPACE workspace. The initial display of charts is a simple array with no clustering. Enough free space is left for the user to arrange the spatial layout of the content as they wish, e.g., by moving charts and creating clusters. As the initially displayed charts are only a subset of all charts that can be generated, users were shown how to create additional charts, as needed.

DynSpace+ uses a layout manager that enforces complete visibility of charts at all times by not allowing charts to partially or completely overlay one another. The available space is divided into invisible rows and columns so that the clusters always appear as an array as in Fig. 3. A chart originally takes up a 1x1 "unit" space but it can be it can be increased to any multiple. The reason behind the design decision was to retain persistent chart visibility and ensure easy organization of over a hundred charts available on the screen.

We could have clustered the initial charts. This idea is not new, and many clustering algorithms have been used in various systems, including those focusing on VA tasks [19,20,23,33,37]. However, as Schreck et al. summarize [37], the unsupervised nature of the clustering algorithm can be disadvantageous, resulting in clusters and layouts that do not comply with user preferences, expectations or the application context [37]. Or worse, they may inappropriately bias the user's exploration. We thus intentionally provided no initial clustering, keeping the process entirely manual and up to the user.

3.3 Datasets for the VA Evaluation

The quality and quantity of the used data need to be considered carefully. Enrico Bertini discusses problems related to too much data, such as cluttered displays, performance drop, information loss and limited cognition [7]. On the other hand, if the dataset is too small, analysis becomes trivial.

We used the following datasets in this work: the 2016 US Census Data [43], an aircraft-bird strike dataset [41], a videogame sales dataset [42], and a superstore dataset [44]. All datasets we too large to be displayed in their entirety on V4-SPACE, but parts of the datasets are still "small" enough to be comprehensible within a time frame appropriate for a user study.

Consistent with Keim's VA mantra [21] (Analyze First, Show the Important, Zoom, Filter and Analyze Further, Details-on-Demand), a variant of Shneiderman's visual information-seeking mantra [38], users are first given an overview consisting of an initial random set of charts, and are able to "zoom" to request more details. Details for each data point on every chart are provided on demand, through mouse hover functionality, which shows all attributes related to the data point. This avoids distracting users with factors that may be irrelevant for the current context and thus enables them to focus on higher-level objectives, while being still able to get "detail on demand".

4 Study I: A Classification Task

This first study addresses RQ1: How do people use the (large) space of an LHRD to cluster data while trying to solve analytical problems that require re-organization of the content? Elaborating RQ1, we are led to ask: How much of the space do they use when clustering objects? How do these clusters look? How far are they from each other? What does the cluster distance symbolize? Are objects strictly separated or loosely grouped? Are there patterns for the space usage? Also, do different clustering approaches lead to significant differences in accuracy and/or task completion times?

The purpose of the study was to observe user behaviour while solving a classification task. The spatialization task used the simplified version of DynSpace+ to display many 2D plots visualizing relations among a subset of the 2016 US Census Dataset. This study used all seven monitors of the V4-SPACE display.

Users were given a workspace populated with charts. We observed in pilots that novice users' classification behaviour was sometimes purely based on the type of charts, which has no relation to the visualized data. To ensure classification decisions were not affected by chart type, Study I used only one type of data (numerical) and one type of chart (scatterplots).

On the given workspace, users could create new charts, delete existing charts, resize charts as needed and move them around freely. Selecting a data dimension on the data panel highlighted charts using that dimension in the visualization panel, allowing users see if their grouping criteria had a "common dimension", i.e., if charts having the same dimension in one of their axes were in the same group.

Grouping the information and generating clusters of objects in analytical tasks is a part of sense-making – the process of searching for a way to encode data in a representation to answer task-specific questions [34]. It takes place in the early stages of VA processes, and can impact task efficiency in the next stages, especially on large displays. For this reason, our objective for this study was to examine clustering on LHRDs.

A core inspiration for this study was Endert et al.'s work [15], which presents the concept of semantic interaction that seeks to enable analysts to spatially interact within their analytical workspace. However, we did not impose a common analytical goal for the users and looked for different approaches of their own. In other words, they were told to cluster the charts, but they were not given the specific clustering criteria that would be measured and compared directly. We believe it is important to observe how users interactively organize their information, unbiased by any initial or assistive clustering algorithm. This is another reason why we decided not to provide any algorithmic clustering in our apparatus.

4.1 Participants

There were nine participants, P0 to P8, four of whom were male. Ages ranged from 18 to over 40. Some were undergraduate students participating for course credit, others were volunteers with at least a bachelor's degree. We ensured that all users could read the text on the displays from the default chair position with no vision problems. In a pre-study survey, participants were given seven questions about VA terms and concepts in order to assess background knowledge. On average, they were familiar with 4.4 of 7 such terms.

Since the study did not require specific domain knowledge, mostly novice users were used for this study. Five participants out of 9 did not have any VA experience. One reported less than a year of experience and three reported 1 to 3 years of experience. Seven had never used any visualization or VA tools, while one had used d3.js and another used R in a statistics course.

In the pre-study survey we asked users if they could interpret data from scatter plots. They ranked their ability of interpreting data correctly from a scatterplot on a scale from 1 to 5. The results were 3 "sometimes", 4 "mostly" and 2 "yes, always" answers; which respectively stood for 3, 4 and 5 on the used scale.

4.2 Experimental Design

Participants were told that they could rotate/swivel the chair to see the full display, but that their chair's location had to remain in place until the end of the experiment, to retain the same distance to all displays in the system. Leaning back and forward was allowed, as needed. We trained participants in basic, yet frequent operations: keeping track of the cursor, how to find it when they lost track of it, and switching it between the LHRD and the auxiliary monitor. Participants were then introduced to the DynSpace+ VA tool and received training on its use. Participants could ask questions during the experiment and were encouraged to ask them as needed. We noted any direct or indirect feedback from the users during the tasks.

4.2.1 Tasks

Participants were asked to complete two tasks, consisting of two subtasks each. In the first task, participants were given charts randomly generated by DynSpace+ based on the dimensions of the dataset used, were asked to arrange the charts into groups based on *similarity*, and told that they could consider any similarity criteria. Participants were required to group all charts. We enforced a minimum of five groups, but they could create more. Following completion of the grouping, they were asked to explain their reasons for their chosen grouping. Since any clustering strategy was allowed and the task was open-ended, we believe that the lack of a list of pre-determined initial charts identical for all users did not affect the users' task results.

In the second task, participants were asked to select each dimension in turn (using the data dimension panel) to see which charts in the visualization panel used the given dimension, so they could observe any patterns that might become apparent. Participants recorded their observations by typing their response in the study questionnaire displayed on the auxiliary monitor.

4.2.2 Data Collected

User clustering activity was tracked both by screen-capturing and by recording the final grouping positions along with information about each chart in the groupings. Questionnaire answers were recorded using free-form text boxes and were stored in a survey system.

At each step of the procedure, we watched and noted observations, to be able to better understand the raw screen recordings, e.g., through manually recordings incidents of frustration at any point in time, which further informed the analysis of the results.

At the end of the study, we did an audio-recorded post-study interview with each participant, which typically took less than 5 minutes. This interview was semi-structured, using a list of questions regarding completed tasks, the software tool, the system, what they liked, what was challenging and/or confusing, and whether they had any further feedback.

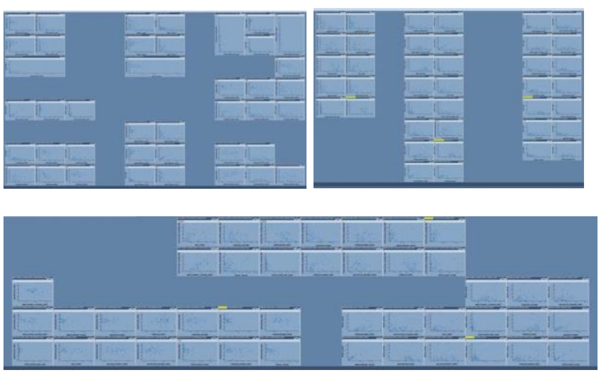

Fig. 3. Variation in aspect ratios used by participants. P6 used a roughly square orientation (aspect ratio ~1, top left) while P4 used a vertically-oriented layout (aspect ratio <<1, top right) and P5 used a horizontal orientation (aspect ratio >> 1, bottom).

4.3 Results

4.3.1 Clustering

Participants' group of charts appeared as clusters of visual information sets. Three participants created only the minimum allowed five groups and four created 6 clusters. The other two participants created 8 and 16 groups respectively. The participants used one of the following three equally prevalent strategies when clustering: Similar visual appearance (P1, P2, P7), commonality in labels (P0, P4, P5), or in topics (P3, P6, P8). When building clusters, charts were arranged horizontally or vertically (P0, P1, P4, P5), around a center (P6, P8), or in a mixed arrangement (P2, P3, P7).

For some users, there was no perceptible relation between different clusters in the workspace. (P0, P1, P4, P5). For P3, cluster shapes were determined by cluster types. For the rest (P2, P6, P7, P8), as the similarity between clusters increased, the distance between them got smaller. For those participants, relations between clusters were reflected by cluster separation. For some participants, there was either no (P0, P1) or only a minimal (P3, P5, P8) distance between clusters. For some (P4, P6), the bezels strongly influenced the cluster separation. The distance between clusters varied for some users (P2, P7) depending on the inter-cluster relations.

After the second task of the first study, participants recorded their observations in writing. Some of the users reported that their clustering mostly aligned with having common dimensions; however, others did not and those explained what approach they followed and why.

4.3.2 Space Usage

If a user runs out of free space, DynSpace+ permits vertical scrolling (but not horizontal), via the mouse wheel. We enabled vertical scrolling, as it is very easy to access with the mouse wheel. Horizontal scrolling is relatively less prevalent and typically requires more interaction work to access. Thus, we did not add support for it in DynSpace+. Seven of the of nine participants used the entire width provided by the system. The remaining two used only about a third of the available space, relying heavily on vertical scrolling.

4.3.3 Navigation Techniques

Six participants used physical navigation frequently, i.e., they rotated their head back and forth. The others kept their gaze mostly focused on a subset of the displays.

4.4 Discussion of Study I

4.4.1 Clustering

Each user's clustering behaviour seemed to be unique. Thus, we did not observe very strong patterns or commonalities among users (recall participants were required to create at least 5 clusters):

- 5 or 6 clusters: P0, P1, P2, P3, P4, P5, P7
- 8 clusters: P8
- 16 clusters: P9

Participants used various criteria for clustering the scatter plots. We identified their criteria by considering the visual appearance of the final layout, feedback from the interview and our own post experiment analyses.

- Visual appearance: Some participants interpreted scatter plots as "pictures" and clustered those that "looked similar".
- Topical relations: Some participants grouped charts with similar and/or related data dimensions. E.g., they created clusters with gender- or sales-related dimensions.
- Common axis labels: Some participants picked a specific data dimension and clustered all charts using that dimension. E.g., they created a cluster of LAND_AREA versus every other data dimension, whether that comparison made sense or not.

Each approach corresponds to a primary clustering criterion that determines which chart is put into which group. Some participants used only a single criterion for grouping while others applied one criterion for constructing clusters and another for in-cluster positioning and for depicting between-cluster relations. For example, P6 grouped charts using a common x-axis label, as the primary clustering criterion. P6 however, deliberately chose chart positions within each cluster based on "how they visually appear".

The participants' rearrangement efforts resulted mostly in simple shapes, either highly rectangular (aspect ratio substantially different from 1) or round/square (aspect ratio around 1). Classifying cluster shapes by aspect ratio (large ratio: horizontal; small: vertical) yields the following grouping (with examples shown in Fig. 3):

- Horizontal: P5.
- Vertical: P0, P1, P4.
- Square: P2, P3, P6, P7, P8.

For several participants, no semantic pattern of cluster arrangement could be identified. P0 and P1 did not have any space between clusters, which were arranged in a random order. For P4 and P5, clusters were separated with some distance between. However, the spacing seemed uniform and clusters appeared similar.

In contrast, other participants used cluster position to convey meaning. P3 used cluster shape to indicate what the clusters represented. While P3's categorized groups looked somewhat similar, P3's group of uncategorized plots looked different from other clusters in terms of shape and the distance from other clusters.

Both P2 and P7 clustered plots in terms of visual appearance, so that clusters of strongly correlated plots were placed far away from other clusters, possibly to further emphasize their difference. Other clusters, whose charts were visually more similar, were positioned closer to each other. These participants evidently used distance to symbolize the similarity of the relations they used. In other words, if members of different groups were unrelated, groups were far apart.

Finally, for P6 and P8, the same effect was seen with topical relations. Most groups were located depending on the information they revealed.

As in previous studies [28], it was clear that bezels played a role in their clustering. We observed participants attempting to utilize bezels as an aid to separate clusters. One participant (P4) in particular seemed to adapt their entire grouping strategy around bezels since they created vertical clusters separated by bezels.

Bezels also caused difficulties: on a few occasions participants missed details on a chart when the chart crossed a bezel. Those users then missed a trend or could not categorize the label name correctly, since the text ran over to the next screen.

4.4.2 Space Usage

An earlier study [1] suggested people cluster information to reflect their mental model and schema, after having looked at a lot of data. Our first study looks at this first stage of the analytics process where people form their mental models and schema, through experimental clustering/classification.

Interestingly, two participants used space in a radically different way, essentially orthogonal to the display system's characteristics and features. While our horizontally-wide display system supports viewing information just through physical movements, these two users' approach (P0 and P1) depended completely *on vertical scrolling* with the mouse wheel. They gathered all charts together and did not leave any distance between them to spatially encode different clusters. P0 reported that the large space was distracting and that they wanted to limit the content to a smaller, physical FOV. Due to this trade-off, P0 and P1 needed much more vertical space than available, 210% resp. 170%. As they used very limited horizontal space, their navigation through the charts required frequent vertical scrolling.

One potential explanation might be that some users prefer to avoid physical navigation, in contrast to previous reports [2,3,4]. In Study I (and in Study II, below) some users stated they wanted to keep everything within their central, i.e., foveal, vision without having to rotate their head. For such users, sitting still and using the mouse to control the system seems to be preferable to performing body and head rotations to access the information on the displays. One potential reason might be that mouse movements could be (or at least perceived to be) faster than head or even body movements. Also, some users do not seem to be comfortable with relying on their peripheral vision and want to gather all information in a compact space (i.e., within their foveal vision) that they can monitor without missing any content.

An alternate explanation is that current advances in technology are pushing users towards certain approaches. As smartphones increasingly pervade everyday life, we become increasingly used to scrolling (vertically) through content. This form of accessing content has potentially formed navigation habits that shape how users arrange content, even on wide screens.

As indicated both by layout results and participant interviews, it is certain that the observed behaviour was based on individual's differences: some participants spread content across the wide horizontal space while others purposefully scrolled up and down.

From a layout orientation point of view, P0 and P1's contents were aligned to the left borders of the space. The clusters of P5 and P7 were also slightly biased towards the left side. One potential explanation is that the data dimensions panel was on the left. In contrast however, P8's content was concentrated towards the right, as also confirmed in that participant's self-report. Other participant's clusters were distributed across the entire display space.

4.4.3 Navigation Techniques

Earlier studies reported a preference for physical over virtual navigation [2,3,4], as noted above. However, in our study, we observed mixed results around navigation. Users exhibited widely different navigational habits and preferences.

P7 leant forward as a different type of physical navigation, to be better able to visually focus on the displays. P3 reported enjoying physical navigation and preferred that over virtual navigation: "*even if I had needed to navigate virtually, I would have rather used tabs than vertical scrolling*".

However, P5 stated that they did not *like* navigating physically, despite organizing their content using that strategy. P8 used both virtual navigation through vertical scrolling and physical navigation by body/head/eye rotation to cover the complete 131° display. P0 and P1 however preferred virtual over physical navigation, using only 37° and 53° of the horizontal FOV respectively, leaving the remaining space blank. During clustering, both ran out of vertical space, and used the mouse wheel to virtually navigate, i.e., scroll up and down. They seemed to be content with this strategy and did not seem to want to navigate physically by placing charts in a wider space and rotating their heads to access them.

5 STUDY II: LHRD IMPACT ON VA

Our second study addresses RQ2 and RQ3:

- RQ2: Do larger displays help user to do better in VA tasks? How does task success change as screen size is increased?
- RQ3: How does task success change as resolution is increased? Does higher resolution improve task success in VA tasks?

The study investigates if display size and resolution affect VA task success, analysis approaches, or user preferences.

5.1 Participants

Of the 18 participants (Table 1), 13 were 23 to 28 years old while the rest were 29 or older. Participants were offered a cash stipend for their participation. As in the first study, participants had to be able to read text on the displays, which was in this study the size of 10 pt text on a paper at normal reading distance.

Participants were required to have a basic understanding of data analysis, visualization knowledge and some VA experience, either through academic (course completion, conducting research) or professional experience.

5.2 Experimental Design

We used a mixed 3 (display size) x 3 (display resolution) design. Display size was a within-subjects factor, while resolution was between-subjects. Each participant was randomly assigned one of three screen resolutions. To avoid ordering effects, we applied a two-dimensional Latin square for the three display conditions and the three datasets across participants. Each task used a different data set to reduce any potential learning effects. Datasets were equivalent in size and complexity and the associated tasks were carefully matched to ensure comparable levels of task difficulty. See Table 1.

To create the different screen size conditions, we temporarily disabled the outer two or four displays, which generated the 3-, 5- and 7-screen conditions of the experiment (S1, S2, S3). Horizontal viewing angles from the chair were ~130, ~90 and ~50 degrees for S3, S2 and S1, respectively.

Participants were randomly assigned to one of 3 resolution configurations, where all displays were set to one of:

- R3: 3840 x 2160 (2160p/4K)
- R2: 1920 x 1080 (1080p or half of original resolution)
- R1: 1280 x 720 (720p or a third of original resolution)

At the distance determined by the chair's position, the R2 condition is equivalent to a Retina Display, while R1 was a lower resolution and R3 had a higher resolution. Our purpose was to determine if resolution made a difference in text and data legibility (investigated in 3D space previously [11]) and to observe the effect of resolution on analysis with dense data. Fig. 4 shows a user working on V4-SPACE with the (S1, R2) condition.

In this study, each participant was given a 10 to 15 minute introduction to the VA tool by demonstrating the main features in a predefined sequence. Participants were then asked to complete a training task on their own, without time restrictions. They were permitted to ask questions if they forgot about functionality, or if they got confused. They were required to finish all steps of the training task until they could complete the tasks on their own, i.e., were reasonably comfortable with the analytic capabilities of the tool, after which the main experimental tasks started.

Each main task consisted of two parts: Exploratory analysis (part a) and answering specific questions about the data (part b).

In the exploratory analysis part, users were asked to find and write down up to ten insights about the dataset, where an insight was specified as an individual observation about the data, i.e., a unit of discovery [35]. Prior to the study they were given instructions on what an "acceptable level of complexity" of their findings was.

In the second part, users were asked a specific set of questions about the dataset; where the answers could be found using the VA tool to analyze the dataset. For each task and the associated dataset users were required to decide what data dimensions to plot and compare against, interpret intermediate level findings correctly, and use obtained information correctly to reach ultimate analytics goals.

Users were given eleven minutes for each of the two parts (part a and part b) of the three tasks (T1, T2 and T3), including the time for reading descriptions, which would give them approximately 10 minutes to spend on the actual analysis for each part. When the time was up, the online questionnaire proceeded to the next page, i.e., from Task 1 (a) to Task 1 (b), regardless of whether the user finished the task or not. Typing the insights on the auxiliary monitor was included in the given time. We did not think that this activity would interrupt analytical process significantly, as the users needed to type only after every individual, separate finding of information.

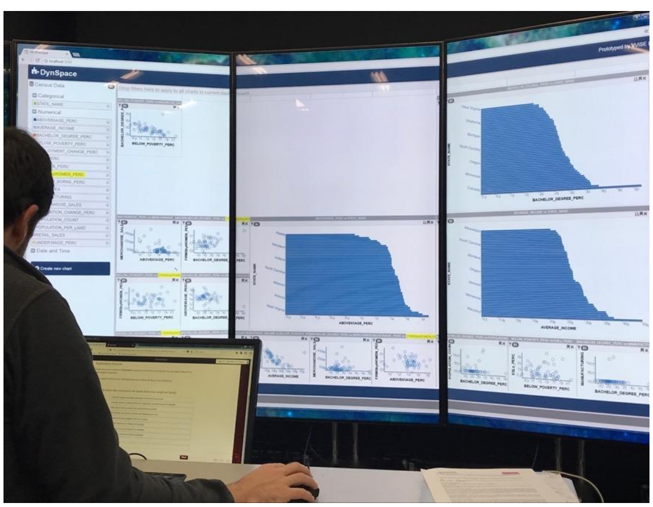

Fig. 4. Participant working with the 3-screen configuration (S1) at 1080p resolution (R2) conditions of V4-SPACE, using the US Census data on DynSpace+. Tasks and the questionnaire are displayed on the auxiliary display, visible in the foreground. The analysis itself was done on the LHRD visible in the background.

5.3 Results

5.3.1 Quantitative Results

The primary goal of this study was to quantify VA task performance. Each participant worked on three tasks with different datasets and performed two types of analysis for each, one exploratory and one structured. The measures of task performances were task accuracy and completion time.

Quantitative evaluation of results is crucial [45,46]. We scored results based on the number of insights found, correct answers to questions, time bonuses and incorrect answer penalties. In part a), each qualifying answer was one point, and time (<10 mins per task) added one point. To maintain a minimum standard in insights, prior to the study they were given instructions on what an "acceptable level of complexity" of their findings was.

In part b), each correct answer was worth 3 points and incorrect answers received -1 point, since we wanted to discourage guessing. Participants could thus score up to 23 points on each task. The results of the scoring are shown in Table 1. In the dataset columns (Task1, Task2 and Task3 on the right), C is the census dataset, V is videogames and S is the superstore dataset. For example, P00 first did the Census data task on size S1 with R1 (score 15), then the Videogame task on size S2 with R1 (score 19) and finally the Superstore data task on size S3 with R1 (score 16).

After the tasks, each user was asked to complete a questionnaire to rate the display conditions in terms of:

- Which display condition was their overall favorite?
- How well did that display condition enhance their effectiveness and efficiency for VA tasks?
- Ease of use of that display condition for VA tasks.

User responses to different display size conditions varied a lot. In general, the 5- and 7-display conditions were better in terms of efficiency and effectiveness. However, the 3-display condition—the smallest—was preferred for ease of use. We did not find statistically significant differences in the subjective user ratings.

Table 1. Study II Task parameters and score. S1/2/3 refers to display size: 3, 5 or 7 monitors; R1/2/3 refers to resolution: low, medium or full; C/V/S refers to the dataset/problem used: Census, Videogame or Superstore. Maximum possible score: 23.

Partici-pant	Size S1		Size S2		Size S3		Reso-lution
	Task type	Score	Task type	Score	Task type	Score	
P00	C	15	V	19	S	16	R1
P10	C	15	V	23	S	13	
P01	V	14	S	17	C	17	
P11	V	23	S	20	C	20	
P02	S	20	C	23	V	20	
P12	S	9	C	12	V	15	
P03	V	15	S	15	V	19	R2
P13	V	16	S	16	V	12	
P04	S	20	C	23	V	15	
P14	S	23	C	20	V	11	
P05	C	20	V	10	S	19	
P15	C	20	V	15	S	23	
P06	S	15	C	18	V	11	R3
P16	S	11	C	19	V	10	
P07	C	20	V	23	S	15	
P17	C	18	V	20	S	17	
P08	V	9	S	19	C	15	
P18	V	19	S	19	C	19	

5.3.2 Qualitative Results

Based on qualitative observations, different display sizes seem to have encouraged/forced participants to make on the fly changes to their strategies during analysis. After switching from the S3 condition to S1, P00 commented: "*I was using (all) screens. Now I feel trapped.*" To overcome space restrictions, they limited the data to a smaller subset of dimensions to reduce the amount of the data.

Characteristically, P01 always worked on a single chart and made it extremely large. "*Can we clear the space?*" was P01's first question when they started the analysis, as they did not seem to find it useful to have many simultaneously visible charts.

P02's analysis seemed to have been affected negatively by two factors: vertical display bezels and the small, fast mouse pointer on large displays. The interaction with the mouse affected time and bezels impacted task success.

P03's favorite was the middle case, S2, explaining "*[S1] was too small and in [S3] there was too much data at the periphery.*"

P07 had a mixed approach, making use of most of the VA tool's capabilities. This participant had a productive session and identified most of the answers. P08 thus discovered it was much faster to look at correlations by dragging a dimension onto a chart than to create a new chart.

At the end of the analysis sessions, participants were asked to answer questions regarding their experience. The purpose was to learn more about preferences on display sizes, navigation, space and content management strategies, and overall analysis approach.

Display Size: P17 summarized their experience by saying: "I was happy with the medium size configuration that uses 5 screens. I could quickly choose items from the menu on the left. The dragging was easier than the large [7-screen] configuration. I had more space to arrange visualization compared to the small [3-screen] configuration."

Space Management: Some participants managed to use the horizontal space efficiently. P14 said: "On a few occasions, I put together multiple charts to see insights." Others did not. P11 said: "I usually expanded the plots in the center of the screen, while there was extra white space [on the] far-right end."

Content Management: DynSpace+ provides an initial set of charts for the data exploration. Most participants liked this feature. P12 stated: "I feel the system was effective at getting a bird's eye view of the trends and correlations." P13: "Being able to see multiple graphs at the same time was a big bonus of this system. I liked to be able to eliminate and filter information and [still] only look the graphs I needed." P17: "Initial figures that are generated automatically made the process and exploration faster. The filters and sorting were quite useful to find specific detail about the data."

Navigational Preferences: Some participants explained they preferred virtual navigation. P6 responded with "*a mix of both*". However, the majority still preferred (the limited form of) physical navigation through chair rotation.

5.4 Discussion of Study II

5.4.1 Discussion of Quantitative Results

To eliminate any effect of experience level between the participants, we normalized the raw scores by setting the personal best of each participant to 10 and scaling their scores from the other two tasks accordingly. Moreover, we also removed 3 of the total of 54 scores as they were unusually low (less than 5) and were likely caused by fatigue. These 3 scores occurred once for each display condition.

To analyze the quantitative task results, we applied a two-way mixed ANOVA after verifying that it meets preconditions, which indicated a significant effect from the size condition: $F(2,9) = 3.32$, $p= < .05$, Fig. 5). According to the post-hoc test, S2 significantly outperformed S1 and S3. Resolution did not have a statistically significant impact. We also verified there was no significant interaction between the two factors.

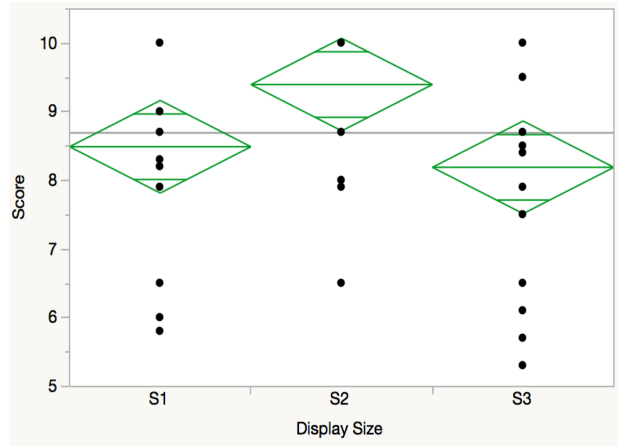

Fig. 5. Results of normalized scores by display size condition.

We also investigated effectiveness and efficiency, ease of use and overall popularity of each display size condition through subjective measures. Variance was very high and none of the conditions were statistically significant. Therefore, we can only state that participants seem to have different preferences for display size. Opinions were spread across the spectrum. This differs from some previous work [2], which demonstrated a clear domination of "physical navigation over virtual", and implied that largest displays will be preferred by most users.

5.4.2 Discussion of Qualitative Results

As discussed in section 5.4.1, display resolution did not directly affect task success. Still, it had qualitative impacts. In the R1 condition, participants enlarged their charts more frequently compared to R2 and R3. The extra resizing effort impacted task completion times as well as space usage and content organization strategies, but not in a significant manner. A potential explanation is that the tasks did not require participants to discern small details.

Concerning the size of the displays, participants expressed substantially different opinions. Although larger displays seemed to be favored slightly (~45% preferred S3, ~33% preferred S2), each option had some popularity with some participants. Looking deeper into the data, we identified 4 groups of VA experience levels with respect to display size preference:

- Level 0 (no VA experience): Highly prefer S3.
- Level 1: Mixed.
- Level 2: Mixed.
- Level 3 (very experienced with VA): Prefer S2.

Overall, more experienced participants seemed to prefer S2, while those least experienced (level 0) preferred S3. Also, for exploratory tasks participants preferred S3, but preferred S1 or S2 for structured tasks.

Physical navigation did not dominate virtual navigation. While some participants liked to turn with the chair and rotated their body and head to view the information, others preferred vertical scrolling through the mouse wheel.

Finally, we observed some task performance inhibitors during Study II, namely, the bezels and limits to mouse interaction. P02 was adversely affected by both factors. P08 seemed to have problems with using a mouse on such a large display, repeatedly losing track of the cursor, wrong clicks, and motions that were simply too fast. That caused P08 to lose time, which resulted in time pressure and a reduction in time task performance.

6 Summary and Conclusion

We conducted two VA studies on LHRDs using a VA tool in this work. The first study involved a spatial organization task, whereas the second study investigated effects of changes in display size and resolution on VA task success.

The first exploratory study was intended to identify challenges and yielded qualitative observations. All participants were given the same task and there was only one condition. We obtained new empirical information about participants' approaches when asked to prepare large volumes of data for analytical tasks on a VA space laid on large displays. The results of study I suggest that participants follow substantially different classification and spatial organization strategies, even though all started with the same state. We observed clear distinctions in terms of clustering strategies, space usage, and preferred navigation techniques.

In the second study, we investigated the effect of display resolution and size. Display size was the within-subjects factor while resolution was between-subjects. We tracked both quantitative and qualitative measures to analyze task performance and participant preferences. In this second study, resolution did not have a statistically significant impact on quantitative task success, but display size did, as measured through completion time or with a scoring system depending on task accuracy. Studying the effect of resolution is new in VA research. With an eye towards optimizing analytics performance, we looked asked how large a large display system should be and whether there is an optimum size, or better an optimal view angle (FOV) that the display system should span.

Participants in both studies complained about bezels going through a chart. In fact, in some cases bezels intruded upon the analysis and negatively affected results, suggesting that adaptation to bezels is important in VA for those LHRDs with bezels.

We kept the analysis space (LHRD) and evaluation space (auxiliary display) separate. While participants worked on LHRDs, administrative study information was displayed on the small monitor in front of them, and they answered task questions on this monitor, as well. Displaying questions on the main screen during analysis would either reduce the available screen space or would require frequent tab switches that would reduce the persistent visibility of the content. An alternative to an auxiliary monitor is paper-based instructions and forms. However, since we found that task switches on LHRDs are very costly, we did not require participants to switch modes from a mouse to pen-and-paper during the task.

Although still the most popular choice, physical navigation was not uniformly preferred by all users of large displays working on VA tasks. This differs from some previous work [2,3,4]. A possible explanation is that while some degrees of physical navigation (eye rotation, slight head movement) are easy for all participants, larger head turns and body rotations were less preferable to some. For those participants virtual navigation, such as vertical mouse scrolling, was preferred. Some participants reported that the larger display condition was irritating due to the need for physical movement for navigation.

As in the literature, large displays are preferred over conventional monitors [3,4]. However, bigger is not always better. S2 appeared to be best condition in our studies; we speculate this may relate to S2 providing ~90 degrees of (horizontal) FOV. Some works support this statement, though. A study that investigates relations between display size, information space, and scale ratio in separate experiments with interfaces that implement classic information visualization techniques (focus+context, overview+detail, zooming) for multi-scale navigation in maps state that a larger display does not improve performance in multi-scale navigation tasks where targets are visible at all zoom levels and that the completion times decrease with larger display size as users find the interfaces harder to use when displays are too large [24]. Another study reports hindering effects of navigational costs on the gains from reducing virtual navigation by using larger displays [47].

Higher resolutions, on the other hand, did not improve task success quantitatively, though it appeared to have had a qualitative impact.

In overall, we believe our findings and observations improve our understanding of the role of LHRDs in VA and see that there is a need for additional studies to compare obtained results.

7 Future Work

In the first study, two participants unexpectedly used only a very narrow portion of the LHRD, performing their tasks through heavy use of vertical scrolling, i.e., in a way they were accustomed to by everyday devices such as smartphones. It would be interesting to study this kind of behaviour further. Moreover, based on the observed differences between individuals, we speculate that one could identify groups of users with common behavioural patterns. Then, one could optimize the design of the system user interfaces for these behavioural patterns. Yet, this is clearly a subject of future work.

References

[1] Andrews, Christopher, Alex Endert, and Chris 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] Andrews, Christopher, and Chris North. "The impact of physical navigation on spatial organization for sensemaking." IEEE Transactions on Visualization and Computer Graphics 19, no. 12 (2013): 2207-2216.

[3] Ball, Robert, Chris North, and Doug A. Bowman. "Move to improve: promoting physical navigation to increase user performance with large displays." In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 191-200. ACM, 2007.

[4] Ball, Robert, and Chris North. "The effects of peripheral vision and physical navigation on large scale visualization." In *Proceedings of graphics interface 2008*, pp. 9-16. Canadian Information Processing Society, 2008.

[5] Baudisch, Patrick, Nathaniel Good, Victoria Bellotti, and Pamela Schraedley. "Keeping things in context: a comparative evaluation of focus plus context screens, overviews, and zooming." In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 259-266. ACM, 2002.

[6] Baudisch, Patrick, Edward Cutrell, and George Robertson. "High-density cursor: a visualization technique that helps users keep track of fast-moving mouse cursors." In *Proc. Interact'03*, pp. 236-243. 2003.

[7] Bertini, Enrico. "How do you visualize too much data?" Fell in Love With Data. http://fellinlovewithdata.com/guides/how-do-you-visualize-too-much-data (accessed October 19, 2017).

[8] Bezerianos, Anastasia, and Ravin Balakrishnan. "View and space management on large displays." *IEEE Computer Graphics and Applications* 25, no. 4 (2005): 34-43.

[9] Bezerianos, Anastasia, and Ravin Balakrishnan. "The vacuum: facilitating the manipulation of distant objects." In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 361-370. ACM, 2005.

[10] Bowman, Doug A., Larry F. Hodges, Don Allison, and Jean Wineman. "The educational value of an information-rich virtual environment." Presence: teleoperators and virtual environments 8, no. 3 (1999): 317-331.

[11] Chen, Jian, Haipeng Cai, Alexander P. Auchus, and David H. Laidlaw. "Effects of stereo and screen size on the legibility of three-dimensional streamtube visualization." IEEE transactions on Visualization and Computer Graphics 18, no. 12 (2012): 2130-2139.

[12] Czerwinski, Mary, Greg Smith, Tim Regan, Brian Meyers, George G. Robertson, and Gary K. Starkweather. "Toward characterizing the productivity benefits of very large displays." In *Interact*, vol. 3, pp. 9-16. 2003.

[13] Dourish, Paul. Where the action is: the foundations of embodied interaction. MIT press, 2004.

[14] Endert, Alex, Patrick Fiaux, and Chris North. "Unifying the sensemaking loop with semantic interaction." In *IEEE Workshop on Interactive Visual Text Analytics for Decision Making at VisWeek 2011*. 2011.

[15] Endert, Alex, Patrick Fiaux, and Chris North. "Semantic interaction for visual text analytics." In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 473-482. ACM, 2012.

[16] Endert, Alex, Seth Fox, Dipayan Maiti, and Chris North. "The semantics of clustering: analysis of user-generated spatializations of text documents." In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pp. 555-562. ACM, 2012.

[17] Endert, Alex, Patrick Fiaux, and Chris North. "Semantic interaction for sensemaking: inferring analytical reasoning for model steering." IEEE Transactions on Visualization and Computer Graphics 18, no. 12 (2012): 2879-2888.

[18] Few, Stephen, and Perceptual Edge. "Dashboard confusion revisited." *Perceptual Edge* (2007): 1-6.

[19] Han, Jiawei. "Data Mining: Concepts and Techniques" Lecture, Department of Computer Science, University of Illinois at Urbana-Champaign, Champaign, IL, October 22, 2007.

[20] Jain, Anil K., M. Narasimha Murty, and Patrick J. Flynn. "Data clustering: a review." *ACM computing surveys (CSUR)* 31, no. 3 (1999): 264-323.

[21] Keim, Daniel, Gennady Andrienko, Jean-Daniel Fekete, Carsten Gorg, Jorn Kohlhammer, and Guy Melançon. "Visual analytics: Definition, process, and challenges." *Lecture notes in computer science* 4950 (2008): 154-176.

[22] Kemp, Karen, ed. Encyclopedia of geographic information science. SAGE publications, 2007.

[23] Khan, Azam, George Fitzmaurice, Don Almeida, Nicolas Burtnyk, and Gordon Kurtenbach. "A remote control interface for large displays." In *Proceedings of the 17th annual ACM symposium on User interface software and technology*, pp. 127-136. ACM, 2004.

[24] Jakobsen, Mikkel R., and Kasper Hornbæk. "Interactive visualizations on large and small displays: The interrelation of display size, information space, and scale." IEEE Transactions on Visualization and Computer Graphics 19, no. 12 (2013): 2336-2345.

[25] Mandler, Jean M., Dale Seegmiller, and Jeanne Day. "On the coding of spatial information." *Memory & Cognition* 5, no. 1 (1977): 10-16.

[26] Marshall, Catherine C., and Russell A. Rogers. "Two years before the mist: experiences with Aquanet." In *Proceedings of the ACM conference on Hypertext*, pp. 53-62. ACM, 1992.

[27] Ni, Tao, Doug A. Bowman, and Jian Chen. "Increased display size and resolution improve task performance in information-rich virtual environments." In *Proceedings of Graphics Interface 2006*, pp. 139-146. Canadian Information Processing Society, 2006.

[28] Ni, Tao, Greg S. Schmidt, Oliver G. Staadt, Mark A. Livingston, Robert Ball, and Richard May. "A survey of large high-resolution display technologies, techniques, and applications." In *Virtual Reality Conference, 2006*, pp. 223-236. IEEE, 2006.

[29] Polys, Nicholas F., Seonho Kim, and Doug A. Bowman. "Effects of information layout, screen size, and field of view on user performance in information-rich virtual environments." *Computer Animation and Virtual Worlds* 18, no. 1 (2007): 19-38.

[30] Ragan, Eric D., Alex Endert, Doug A. Bowman, and Francis Quek. "How spatial layout, interactivity, and persistent visibility affect learning with large displays." In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pp. 91-98. ACM, 2012.

[31] Robertson, George, Eric Horvitz, Mary Czerwinski, Patrick Baudisch, Dugald Ralph Hutchings, Brian Meyers, Daniel Robbins, and Greg Smith. "Scalable Fabric: flexible task management." In *Proceedings of the working conference on Advanced visual interfaces*, pp. 85-89. ACM, 2004.

[32] Robinson, Anthony C. "Collaborative synthesis of visual analytic results." In Visual Analytics Science and Technology, 2008. VAST'08. IEEE Symposium on, pp. 67-74. IEEE, 2008.

[33] Rousseeuw, Peter J., and L. Kaufman. *Finding Groups in Data*. Wiley Online Library, 1990.

[34] Russell, Daniel M., Mark J. Stefik, Peter Pirolli, and Stuart K. Card. "The cost structure of sensemaking." In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, pp. 269-276. ACM, 1993.

[35] Saraiya, Purvi, Chris North, and Karen Duca. "An evaluation of microarray visualization tools for biological insight." In Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on, pp. 1-8. IEEE, 2004.

[36] Saraiya, Purvi, Chris North, and Karen Duca. "An insight-based methodology for evaluating bioinformatics visualizations." *IEEE transactions on visualization and computer graphics* 11, no. 4 (2005): 443-456.

[37] Schreck, Tobias, Jürgen Bernard, Tatiana Von Landesberger, and Jörn Kohlhammer. "Visual cluster analysis of trajectory data with interactive kohonen maps." *Information Visualization* 8, no. 1 (2009): 14-29.

[38] Shneiderman, Ben. "The eyes have it: A task by data type taxonomy for information visualizations." In *Visual Languages, 1996. Proceedings., IEEE Symposium on*, pp. 336-343. IEEE, 1996.

[39] Shrinivasan, Yedendra Babu, and Jarke J. van Wijk. "Supporting the analytical reasoning process in information visualization." In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 1237-1246. ACM, 2008.

[40] Skupin, André. "A cartographic approach to visualizing conference abstracts." IEEE Computer Graphics and Applications 22, no. 1 (2002): 50-58.

[41] "Bird Strikes 1996 – 2000" http://bit.ly/2zSfzTU. (accessed November 20, 2017)

[42] "Superstore Sales" http://bit.ly/2hNVA2g. (accessed November 20, 2017)

[43] "US CENSUS 2016 DATASET" http://bit.ly/2zkupn7. (accessed November 20, 2017)

[44] "Video Game Sales" http://bit.ly/2mIIOno. (accessed November 20, 2017.

[45] Shupp, Lauren, Christopher Andrews, Margaret Dickey-Kurdziolek, Beth Yost, and Chris North. "Shaping the display of the future: The effects of display size and curvature on user performance and insights." Human–Computer Interaction 24, no. 1-2 (2009): 230-272.

[46] Yost, Beth, Yonca Haciahmetoglu, and Chris North. "Beyond visual acuity: the perceptual scalability of information visualizations for large displays." In Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 101-110. ACM, 2007.

[47] Rønne Jakobsen, Mikkel, and Kasper Hornbæk. "Sizing up visualizations: effects of display size in focus+ context, overview+ detail, and zooming interfaces." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1451-1460. ACM, 2011.

[48] Reda, Khairi, Andrew E. Johnson, Michael E. Papka, and Jason Leigh. "Effects of display size and resolution on user behavior and insight acquisition in visual exploration." In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 2759-2768. ACM, 2015.

[49] Bi, Xiaojun, and Ravin Balakrishnan. "Comparing usage of a large high-resolution display to single or dual desktop displays for daily work." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1005-1014. ACM, 2009.