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

by

Gökhan Çetin

B.Sc., Bilkent University, 2015

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
School of Interactive Arts and Technology
Faculty of Communication, Art and Technology

SIMON FRASER UNIVERSITY
Spring 2018

Approval

Name:	**Gökhan Çetin**
Degree:	**Master of Science**
Title:	**Visual Analytics on Large Displays: Exploring User Spatialization and How Size and Resolution Affect Task Performance**
Examining Committee:	**Chair: Carman Neustaedter** Associate Professor
	Wolfgang Stuerzlinger Senior Supervisor Professor
	John Dill Supervisor Professor Emeritus
	Brian Fisher External Examiner Professor School of Interactive Arts + Technology Simon Fraser University
Date Defended/Approved:	March 28th, 2018

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University

or has conducted the research

c. as a co-investigator, collaborator, or research assistant in a research project approved in advance.

A copy of the approval letter has been filed with the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Abstract

Large, high-resolution displays (LHRDs) 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 environments, presenting both 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. I 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 experimental conditions.

Keywords: Visual Analytics; Large High-Resolution Displays; sense-making; spatialization; clustering; visualization

Dedication

To my parents, my sister and my entire family, who have always been on my side and have provided lifelong inspiration, endless motivation, unconditional love and all kinds of support regardless of the distance between us.

To Lindsay for her patience, understanding, kindness and positive energy; without whose love and encouragement I could not have completed this thesis.

Acknowledgements

The creation of this thesis was tough and challenging; however, it was also a joyful, adventurous experience thanks to many people that I mention here. The following are who helped this effort to become a complete thesis, all of them will forever be in my thesis-related memories.

I should start with my senior supervisor Professor Wolfgang Stuerzlinger. Many thanks to him for the incredible, unique infrastructure he provided to me, valuable ideas and wise directions throughout the process of writing, and the huge effort of reviewing and creation of the thesis. I feel privileged to have worked with him.

I would like to give special thanks to dear Prof. Emeritus John Dill, also my supervisor. Observing him has been a great life experience for me. Having the chance of working with Professor Dill was a fantastic opportunity as he has not only contributed to the research leading to this thesis from early on, but also inspired me from many different perspectives.

I also would like to thank all my lab mates, for making me feel to be a part of a great team. There has always been a positive atmosphere in the VVISE Lab, motivating us to do good work. I would like to highlight Maha, as I greatly appreciate her assistance, and collaboration, which made things easier for me.

Many nights' have been dedicated to this thesis. Thanks to my fellows, especially Kivanc, Mert and Mali for pulling all-nighters on numerous occasions.

I benefitted from the help of many study participants during my research. Thanks to Nghia, Crystal, Cailey and those undergraduate and graduate students of SIAT that I have not mentioned the name of, for their time and effort during my experiments.

Finally, I would like to thank from heart to all my friends, colleagues and those who have been a part of this effort but not have been explicitly named above.

Table of Contents

List of Tables

List of Figures

List of Acronyms

CCFL	Cold-Cathode Fluorescent Lamp
CRT	Cathode Ray Tube
DPI	Dots Per Inch
FOV	Field of View
GBD	Gap Between Displays
HCI	Human-Computer Interaction
HD	High Definition
InfoViz	Information Visualization
IRVE	Information-Rich Virtual Environment
LCD	Liquid Crystal Display
LED	Light Emitting Diode
LHRDs	Large High-resolution Displays
OLED	Organic Light Emitting Diode
PFOV	Physical Field of View
PPD	Pixels Per Degree
PPI	Pixels Per Inch
SciViz	Scientific Visualization
SFOV	Software Field of View
SFU	Simon Fraser University
SIAT	School of Interactive Arts + Technology
TAM	Technology Acceptance Model
UHD	Ultra-High Definition
VA	Visual Analytics
VE	Virtual Environment
VR	Virtual Reality

Glossary

4K	Sometimes called 4K resolution, refers to a horizontal screen display resolution in the order of 4,000 pixels. There are several different 4K resolution configurations. Displays used in this work use the 3840x2160 pixel configuration defined by the 4K resolution standard.
Analytical reasoning	Refers to the ability to look at information and discern patterns within the information. Analytical reasoning involves deductive reasoning with no specialised knowledge, such as: comprehending the basic structure of a set of relationships, recognizing logically equivalent statements and inferring what could be true or must be true from given facts and rules.
Bezel	The frame that surround conventional monitors.
Big Data	Data sets that are so voluminous and complex that traditional data processing application software are inadequate to deal with them. Big data challenges include capturing data, data storage, data analysis, search, sharing, transfer, visualization, querying, updating and information privacy.
CAVE	A cave automatic virtual environment (better known by the recursive acronym CAVE) is an immersive virtual reality environment where projectors are directed to between three and six of the walls of a room-sized cube. The name is also a reference to the allegory of the Cave in Plato's Republic in which a philosopher contemplates perception, reality and illusion.
Chart	A chart is a graphical representation of data. A chart in DynSpace+ is a 2D representation of selected dimensions of a given data.
Clustering	Be or come into a cluster or close group; congregate.
Dashboard	A “faceted analytical display”: a set of interactive charts that simultaneously reside on a single screen, each of which presents a somewhat different view of a common dataset and is used to analyze that information.

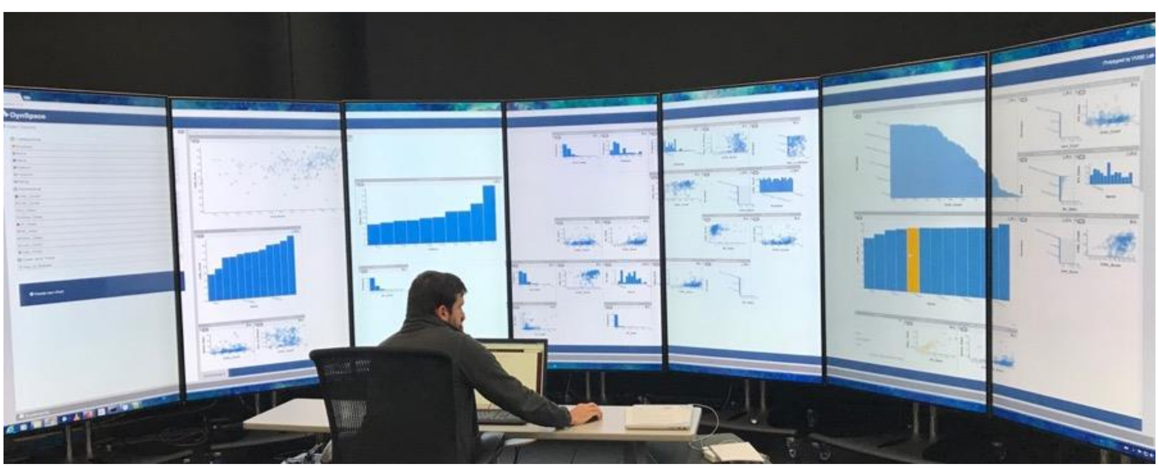
Display	The aggregate visual output intended to be treated as a single contiguous space.
Embodiment	Embodiment denotes a form of participative status. It is about the fact that things are embedded in the world, and the ways in which their reality depends on being embedded.
Information visualization	It is the study of (interactive) visual representations of abstract data to reinforce human cognition. The abstract data include both numerical and non-numerical data, such as text and geographic information.
Insight (visual analytics)	Gained through interacting with the system, exploring possible connections, investigating hypotheses during interactive visual data exploration with large, complex datasets.
Latin square	An arrangement of letters or symbols that each occur n times, in a square array of n^2 compartments so that no letter appears twice in the same row or column.
Powerwall	A powerwall is a large, ultra-high-resolution display that is constructed of a matrix of other displays, which may be either monitors or projectors.
Retina Display	A marketing term developed by Apple to refer to devices and monitors that have a high enough PPI so that a person is unable to discern individual pixels at a normal viewing distance
Semantic interaction	Interaction that seeks to enable analysts to spatially interact with their visualizations directly within the visual metaphor, using interactions that derive from their analytic process.
Sense-making	Sense-making is the process of searching for a representation and encoding of data in that representation to answer task-specific questions.

Spatialization	(in a VA context) Two-dimensional view of high-dimensional data such that similarity between information is represented by relative distances between data points.
Visual analytics	Visual analytics (VA) is the science of analytical reasoning facilitated by interactive visual interfaces.

Introductory Images

The studies in this thesis have been conducted on a system called V4-SPACE. Above is a picture of this 58-megapixel Very Large Ultra-High-Resolution Display System.

Below is a picture taken during a study, where a participant is working on a visual analytics (VA) task on V4-Space using DynSpace+, a VA tool built by our research lab.



Chapter 1.

Introduction

Ever-increasing volume and complexity of data is becoming the new standard. People are likely to encounter complex datasets more frequently when dealing with real world problems, especially since there is a strong trend towards automatizing the data collection processes. Millions of entries with dozens to hundreds of data dimensions can easily be found in many instances of Big Data sets. Such datasets can yield extremely powerful insights; yet need to be analyzed for understanding what the collected data means. We can categorize the analysis approaches for aforesaid data into two main groups; algorithmic analysis and human-in-the-loop. Current algorithms are not fully capable nor reliable for automatically analyzing data, also because they currently cannot understand the semantics of the data. The more practical alternative, human-in-the-loop processes, involves *Visual Analytics*; which is a basis for this thesis.

Visual analytics (VA) is the science of analytical reasoning facilitated by interactive visual interfaces [30]. Other work defines visual analytics as a science-based activity supporting sense-making of large, complex datasets through interactive visual data exploration [97]. All definitions share that the user reasons and makes sense of the data through interaction with visualizations of the data. This gave rise to an industry around VA applications including Tableau and INSPIRE.

In most VA applications, users interact with data through widgets, such as sliders and menus, that control the visualization through modifying the underlying model parameters. However, this is indirect interaction with the data. Endert et al. presented *semantic interaction*, which seeks to enable analysts to spatially interact with their visualizations directly within the visual metaphor, using interactions that derive from their analytic process [42]. Endert further showed that semantic interaction supports *sense-making* better [45]. In a spatial workspace, users can directly arrange documents spatially into clusters to convey similarity or relationships in the data [44]. The familiarity and flexibility afforded by a spatial workspace with semantic interaction allows users to easily establish implicit relationships among data elements in a large dataset [67].

When space is relatively limited, as it is on normal desktop monitors, existing visualizations must often be removed (at least temporarily) to make room for new ones, requiring additional cognitive effort from the user to remember such information. Increasing the display size changes this dynamic significantly, allowing the user to visually access more information at once. Comparisons can be made visually and directly rather than relying on memory and imperfect mental models, which is analogous to the traditional recognize vs. recall issue when comparing command line interfaces to Windows-Icons-Menus-Pointer (WIMP) interfaces. On a large display, a flick of the eye or turn of the head is all that is required to consult a different data source [2]. Moreover, compared to traditional desktop monitors, previous work [2,13,32] has found that large, high-resolution displays (LHRDs) improve productivity and we can expect this to hold for VA applications as well. With small monitors, we often face a trade-off between the level of detail and the number of different objects that can be displayed.

From a data point of view, large display systems not only contribute positively to analytic processes, but also become a necessity with increasing volumes of data. Considering how datasets have changed over time in terms of volume and complexity, large displays will become increasingly necessary if a contiguous space is desired and users want to see a sufficient level of detail for analyzing their data. Huber’s taxonomy of large datasets [53] reflects the fact that we will not be able to analyze large datasets on traditional desktop monitors with context and details being available at the same time.

When high-dimensional data is visualized in a 2D plane, users may wish to manipulate the layout of the data points to better reflect their domain knowledge or to explore alternative structures [43]. *Spatialization* in the VA context is defined as two-dimensional view of high-dimensional data such that similarity between information is represented by relative distances between data points, where, e.g., a cluster represents a collection of similar information [101]. Other sources define spatialization as the transformation of high-dimensional data into lower-dimensional, geometric representations basing on computational methods and spatial metaphors [57]. Considering the formal definition in a VA context, I assumed LHRDs to support content spatialization better as the available space is much larger, so that the *relative distances between data points* are more distinct and *clusters* are represented more visibly. I wanted to verify that assumption in my thesis work.

Furthermore, the increasing volume and complexity of data suggest great potential to improve analytical task performance via LHRDs. However, the switch from conventional desktop monitors is much more than just an arbitrary change in size and resolution. As much larger screens, such as wall-sized displays, become available we face a new set of challenges and usability issues, despite their potential advantages over conventional displays. Users encounter a variety of usability problems with large displays, such as keeping track of the cursor [14], access to distal[1] parts of the display and icons [17,58], and window management [16,84]. Those and probably more factors need to be considered when using LHRDs. Thus, more research is needed on all aspects of LHRDs to explore the spatialization of information and effects of semantic interaction on analytical task performance, as well as the impact of the increase in display size and resolution, which together form the main motivations for this work.

My 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 identify the effect of size and resolution of the display system on visual analytics task completion success. In my study, I pose three general research questions (RQs). The RQs investigated in this thesis are as follows:

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? What are the factors that users consider when clustering; topical relations, visual similarity, or common dimensions of charts? 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?

RQ2: Do larger displays help the user to do better in visual analytics tasks? How does task success change as screen size is increased? Task success is here defined as shorter task completion time and better accuracy in visual analytics tasks, such as answering fact-based questions or finding insights in relatively complex datasets among categorical, numerical or time data using scatterplots, bar charts and histograms.

[1] distal: situated away from the center of the body or from the point of attachment.

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

To address RQ1, I conducted the first experiment. RQ2 and RQ3 were explored in a second experiment.

In the first study, spatialization on LHRDs is not compared to conventional small monitors. Since LHRDs are characteristically different from small monitors, I wanted to explore how spatialization takes place on LHRDs. (Thus, there were no multiple conditions.) To observe clustering behavior on large displays, I gave users a VA task and 2D data charts to work with.

In the second study, RQ2 was the prioritized research question and it was investigated within-subjects. RQ3, which explores how display resolution affects VA task success, used a between-subject design where subjects were randomly assigned to 3 groups, where each group was given one of 720p, 1080p or 2160p (4K) resolution on the same underlying physical displays to investigate if higher resolution improves task success.

In this thesis, several concepts will be made use of frequently. I define and explain them here to avoid any confusion and ambiguity in the rest of the document. Frequently used definitions, concepts and terms are given in Appendix A.

This thesis presents:

- V4-SPACE, a very large display space with high-resolution to perform VA tasks,
- DynSpace+, a visual analytics tool built in the VVISE lab that runs on V4-SPACE, and
- two studies that analyze VA task success and explore users' content spatialization approaches on LHRDs, using DynSpace+ running on V4-SPACE.

According to current standards in the field, my system qualifies as a "Very-Large Ultra-High-Resolution Display (VLUHRD) System". The broad term for such systems is, however, Large High-Resolution Display (LHRD). Although VLUHRD suits my system better, especially when comparing it to other similar large displays, I will also be using a common LHRD abbreviation in the following chapters of the thesis.

The main contributions of this thesis are:

1. Exploration of spatialization and manual clustering behaviours of users during VA tasks on large displays.

2. Evaluation and analysis of VA task successes under different display size and resolution conditions.

The rest of the thesis is organized as follows: Chapter 2 discusses related work with an emphasis on more closely related work in the end of the chapter. Chapter 3 outlines the system design and specifications and discusses various design decisions made when extending the existing systems. The two studies conducted for this work are described in Chapters 4 and 5, along with their methodology, design, results and separate discussion subsections. Chapter 6 contains a general discussion. Finally, limitations and future work are presented in Chapter 7.

Chapter 2.

Literature Survey

This work explores *visual analytics* (VA) on LHRD systems. In this chapter I am discussing the research that took place in the literature in 5 sections. I start by defining the underlying concepts and explain the used terms. Next, I will discuss the need and the motivation for using large, high-resolution display systems in visual analytics research. I will introduce major large display systems constructed so far, followed by a separate discussion of work closely related to this thesis. Finally, I will talk about the methodology of similar research, to determine the vision to be adopted in this work.

2.1. Defining Visual Analytics

Visual analytics (VA), the science of analytical reasoning facilitated by interactive visual interfaces, combines techniques from information visualization, computational data analysis and other areas to support the analytical reasoning process [30]. It should not be confused with *information visualization*, which targets the conveyance of an understanding of data through interactive abstract visual representations [22]. Thus, a key aspect that distinguishes visual analytics from other related fields (InfoVis, SciVis, HCI) is the focus on *analytical reasoning* [73]. And indeed, the most widely accepted definitions of VA emphasize the analytical reasoning process as a crucial aspect of VA.

Analytical reasoning processes are just as important as final products in visual analytics [73]. These processes generate information through discovered *insight*s and how the users arrive at these insights (provenance). The visualization community's definition of insight differs from that of the cognitive community [28]. The cognitive science community has used the term insight "to name the process by which a problem solver suddenly moves from a state of not knowing how to solve a problem to a state of knowing how to solve it" [65] and posits that it specifically refers to what is commonly called an "aha" or "eureka" moment [61]. In the visualization community, though, the communication of insights is considered as the primary purpose of a visualization [22].

Visual analytics is also defined as a science-based activity to support *sense-making* in large, complex datasets through interactive visual data exploration. Through interacting with the system, users can explore possible connections, investigate hypotheses, and ultimately gain insight. This complex and personal process is referred to as *sense-making* [78]. In short, sense-making is the process of searching for a representation and encoding of data in that representation to answer task-specific questions [88]. The relation between sense-making and insight is that insight constitutes a step in the sense-making process. The sense-making process is described by Peter Pirolli and Stuart Card as follows [78]:

Information → Schema → Insight → Product

In their work, sense-making was explained through a process consisting of loops, among which two have special importance: the foraging loop, which focuses on the gathering and processing of data to create schemas, and the sense-making loop, which includes the processes involved in moving from schemas to finished products. Foraging refers to the process of filtering and gathering collections of interesting or relevant information. Then, using that information, users advance through the synthesis stages of the process, where they construct and test hypotheses about how the foraged information may relate to a larger structure, situation or problem. *Semantic interaction* combines foraging activities with spatial synthesis activities [43].

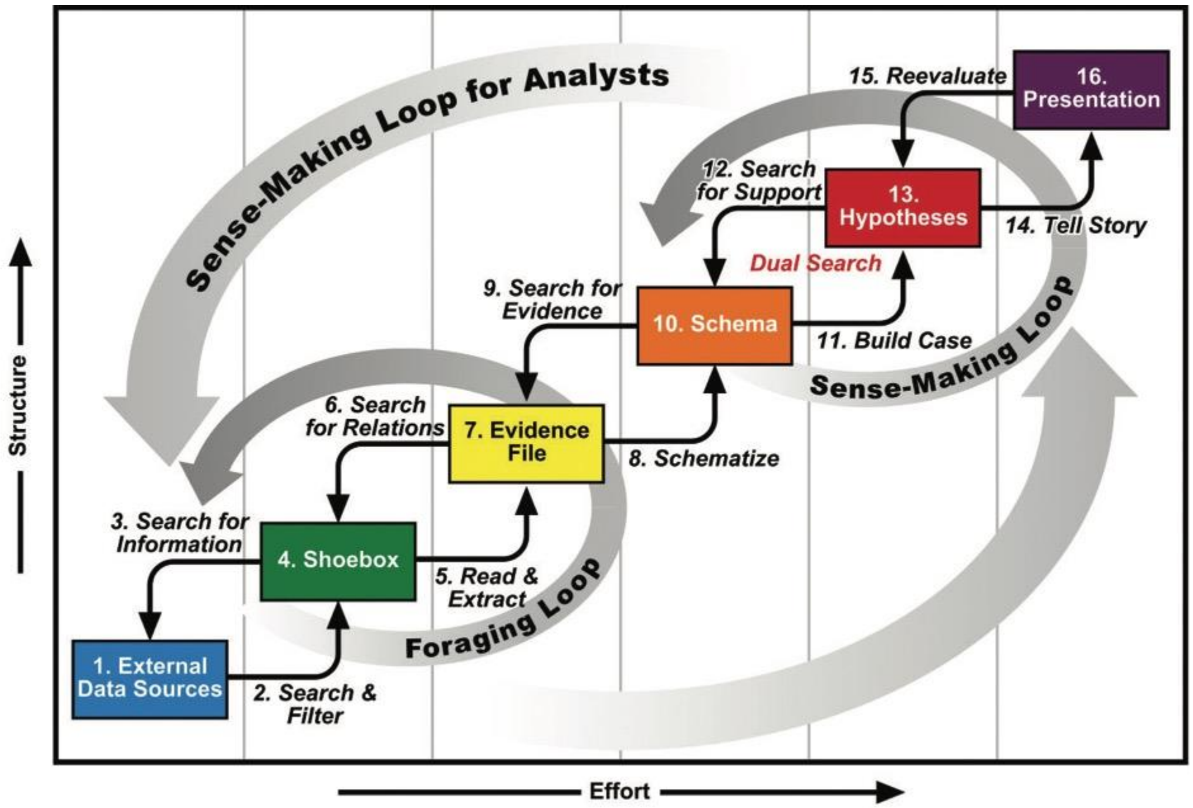


Figure 2.1. Sense-making loop

2.2. Motivation for Using Large High-resolution Display Systems in Visual Analytics

In a spatial workspace, users can interactively organize documents into spatial (visual) clusters to convey similarity or relationships [44]. To accomplish this, analysts of course require a spatial workspace in which to spatially organize documents into clusters and other visual structures [44]. The familiarity and flexibility afforded by a spatial workspace allows users to identify and make explicit – through spatializing them – the implicit relationships within the dataset [67].

The trend noted in Chapter 1 towards larger display areas in work environments has been continual over the past few decades. In earlier times when cathode ray tube (CRT) monitors were widely used, increasing monitor sizes, and using additional monitors for dual desktop monitor configurations or similar, was common [51]. Users of such systems reported benefits for their work [19,32].

Recently, display technologies have enhanced, display sizes have increased and their configurations have become more complex. I am investigating VA on large high-resolution display systems because I would like to improve VA task completion success and believe larger screens—LHRDs—can help. I also want to study how users exploit a large available space to organize their content since it is worthwhile to confirm whether the reported advantages of large displays extend to VA research. *Display* is defined here as the aggregate visual output intended to be treated as a single contiguous space [3]. This distinction is important because the display can be (and often is) composed of multiple tiled physical units, which could be projectors or monitors, possibly not contiguous. Several reports indicate that increasing size and resolution have positive effects on productivity in various tasks undertaken with such systems [2,13,32]. For example, *powerwall*s, high-resolution wall-size displays, can increase productivity [32] and LHRDs have been shown to support data analysis in comparison with normal sized displays [2]. Since the technology has improved and better configurations are becoming increasingly available, I was privileged to have available for my work a very large display system with uniform very high resolution.

Such displays improve productivity [2,13,32]; because they allow the user to visually access more information at once. Thus, comparisons can be done visually, rather than relying on memory and imperfect mental models. This can be clearly exploited; since a 17" desktop monitor, for comparison, covers only about 10% of the visual field [51]. This goes down to 1%, if head rotation is also allowed for the user to keep track of displayed information. Therefore, theoretically, the size of any information can be kept the same and the available display area can be increased to very large fractions of the visual field to reduce the context switching effects of a single (smaller) screen which rely on mental resources (e.g. increasing the number of monitors in a display system without any further changes, to have more space to accommodate data). Maintaining a high resolution would also increase the amount of information that can be accommodated, at least at a perceivable level. For example, the amount of text at a perceivable level that could be accommodated in a high-resolution setting of a monitor is higher than a lower resolution mode of the same monitor. On small or low-resolution monitors, we often encounter the trade-off between level of detail and the number of different objects that can be displayed on screen. The "normal" or detailed view is one in which the details of the document can be directly perceived (e.g. text can be read reasonably easily). While such a view depends on many factors, it generally takes a lot of space. Other representations include thumbnails, icons, and labels. Each of these requires progressively less room, but at the same time conveys less information [2]. While this is an issue on smaller monitors, on large display systems it is much less so.

In addition to displaying more information, large displays provide increased freedom in the placement of such information on the display system. Users can more easily exploit the power of spatialization by using semantic interaction on large displays by organizing their visual objects in the larger available screen space. On traditional desktop displays, spatialization is much reduced and users must use virtual methods such as tabs or detail-in-context methods to switch between documents, which makes it harder to encode the relations between objects into the visual (tab) structure. Using constraint-free spaces on very large ultra-high-resolution displays allows objects to be freely moved and thus makes it easier for the user to spatially structure the visual, enabling them to more easily encode relations between objects by their position in the space.

Larger displays have been proven to help user performance in various application areas, such as visualization [11,31,74,99,113] and geospatial tasks [98]. In addition, they have been shown to have a positive cognitive impact on their users, in terms of how users perceive and work with their information [2,18,99,106]. Users were reported to prefer larger displays, found them “convenient”, and thought that larger displays were useful when they were working with their information regardless of whether their task performances quantitatively improved or not. It seems clear that evaluating LHRDs is a much more complex issue than simply measuring performance with visualization tasks between subjects.

2.3. Systems Constructed for Large Display Research

To exploit the above-mentioned advantages of using large displays, many large display systems, targeted at a variety of applications, have been created so far. They have been built using different materials, with various design considerations and for different purposes. Here I briefly introduce some of the notable work in large display research.

In 2000, Bishop and Welch built a large display as a part of their project “Office of Real Soon Now” [19]. This very basic approach, shown in Figure 2.2, took no more than a few days to create and involves two mechanically adjusted parallel projectors throwing images onto the office wall. This approach eliminated two CRT monitors and freed up much desk space. Two projectors with 1024 x 768 pixels yielded a composite display with 2048 x 768 pixels. Images were projected on a flat display surface and were aligned approximately. No further details were considered. Even this simple change in their workspace was reported to provide better ergonomics and to improve social and technical interaction. They also reported that higher information content helped them in their work [19].

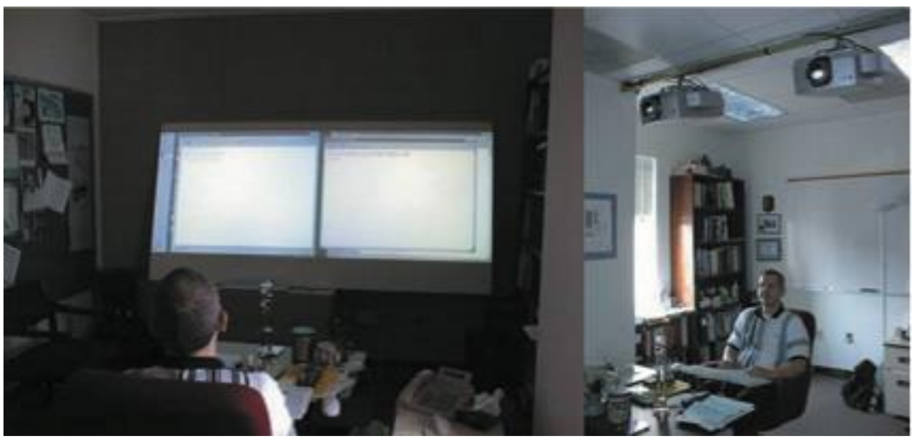


Figure 2.2. The “Office of Real Soon Now” [43] from back and front, showing display and parallel projectors.

Czerwinski et al. presented “Dsharp” (Figure 2.3), a large surface created by using three XGA DLP projectors at 1024 x 768 resolution, projecting onto a curved Plexiglas panel for an equivalent of a 3072 x 768 resolution display [32]. This system yielded a 12” high by 48” wide screen with a 4:1 aspect ratio curving around the user to avoid attendant perspective distortion. It was reported that there was a visible but very small (under 1/32”) seam between each of the three projectors [32]. A Microsoft natural keyboard and Intellimouse were used as the input devices. Microsoft Internet Explorer v. 6.0 and Microsoft Office XP Professional applications were used for the tasks. Their series of user studies demonstrated significant performance advantages of the large display system over 15” monitors [32].

Figure 2.3. User working on experimental Dsharp display.

In one of their studies Robert Ball and Chris North reported an observational analysis of the use of a large tiled display containing nine monitors in a 3x3 matrix, with a total resolution of 3840x3072 (11,796,480 pixels) [10]. The LCD flat panels formed a display area approximately 37” tall and 44.4” wide. They put the display on a standard desktop table approximately 30” off the ground. The tiled display was constructed from nine 17” Dell monitors affixed to a wooden frame (see Figure 2.4). A Dell Optiplex GX270 was used to drive all nine monitors. In addition to the dual head AGP video card that came with the computer, four additional PCI video cards were installed. All video cards were Nvidia GeForce FX 5200s. Figure 2.4 shows various aspects of the tiled display.

Figure 2.4. **Ball and North's [10] 3x3 tiled display. Top figure shows front view of display, with one monitor disconnected to show the underlying wooden frame. Bottom figure shows side view, and how monitors connect to the frame.**

In a study by Bi and Balakrishnan [18], a large high-resolution display was compared with conventional desktop monitors. The physical size of the display was 6' high and 16' wide. The display was created with 18 projectors with 1024 x 768 pixels-resolution each, in a 3x6 tiling, building a total display resolution of 6144 x 2304 pixels. Color and brightness calibration among the images coming through different projectors was minimal. A single computer with several multi-headed graphic cards was used to run all the projectors concurrently. The Windows XP operating system and standard software applications were used on the computer. While the interaction was not explained explicitly, we can assume it involved traditional interaction with keyboard and mouse. While not mentioned explicitly, the user always sat at the desk in front of the display. The users' distance to the display and visual angle were also not specified. A user working on the system is shown in Figure 2.5.

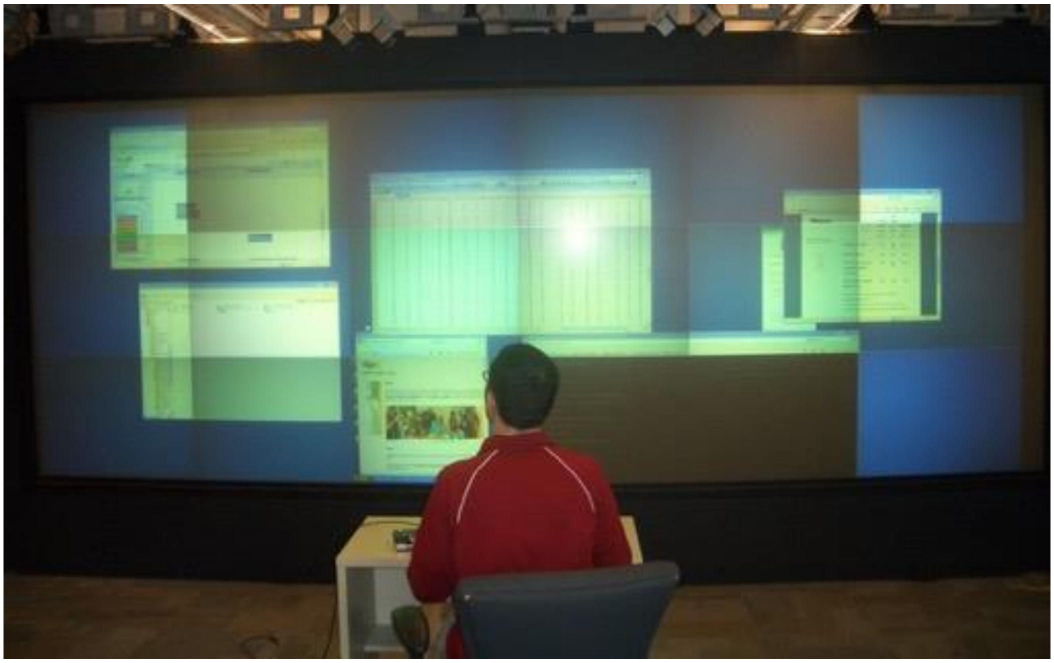


Figure 2.5. User working on a large high-resolution display constructed with projectors.

Tiling small monitors to construct a larger display system is a common approach when creating a LHRD. Shupp et al. presented a system with high pixel density and high pixel count (96 PPI, approximately 32 million pixels) that consisted of twenty-four 17” LCD flat panels and twelve GNU/Linux computers [99]. The monitors were paired, two to a computer, and set to their highest resolution of 1280×1024. Color, brightness and contrast were tuned to achieve close, though not perfect, matches between the monitors. The plastic casing around each monitor was removed to reduce the bezel gap between adjacent monitors to 2cm. Each column was mounted on a freestanding wooden support, allowing the columns to be moved independently. The twenty-four monitors were arranged in an 8×3 matrix approximately 9 feet wide and 3 feet high, with a total pixel count of 10,240×3072 = 31,457,280 pixels. It was a reconfigurable display system, that could be curved or kept flat, as seen in Figure 2.6. Although this is an effective way of building LHRDs with easily obtainable small monitors as building blocks, it has disadvantages such as many bezels, an issue that has been identified to create spatial distortion [3]. Moreover, the requirement for a cluster of machines to control all the monitors increases system complexity dramatically.

Figure 2.6. 24 panel, 31.5 megapixel display in the flat and curved configurations

The 34.5 million-pixel display wall at the University of Calgary, built with 15 rear projection screens [82], offered a solution to eliminating bezels and provided seamless images. Display size was 4.95m by 1.85m. Each projector had a 1920 x 1200 resolution, for a total of 9600 x 3600 pixels. Pixel density of the system was 50 PPI. The wall was operated by a single PC with dual Xeon E5505 processors, 96 GB RAM, and four NVIDIA Quadro K5000 GPUs running a Windows 7 operating system. Interaction was through a wireless keyboard and a mouse on a desk. The physical environment was described as a “quiet working space like an office or research space”. The user’s distance to the display was said to be roughly one meter. Users had tasks from different disciplines. The configuration is shown in Figure 2.7. However, projection-based solutions have their own set of problems, such as specific lighting requirements and lower image quality than monitor-based display systems. Furthermore, projection-based units have higher purchase costs, maintenance costs for bulbs, larger space needs and more complex projector alignment requirements [72].

Figure 2.7. The University of Calgary's 34.5 million-pixel display wall

"The hyperwall" designed by NASA was another significant approach in terms of extending large display usage in scientific visualization tasks [90]. The displays in the hyperwall are 18" Samsung 181T flat panel LCD monitors framed by a uniform black plastic bezel approximately ¾" in width. A custom designed mounting rack allows pitch and yaw adjustment of each monitor, as well as translational adjustment between rows and columns. In addition, each monitor can be moved independently up to 14 inches in "z", i.e., perpendicular to the frame, which enables nonplanar arrangements of the viewing surfaces, such as spherical or paraboloid surfaces. The rack was designed to provide all these degrees of freedom and to accommodate different viewing distances. In practice, the Samsung monitors delivered 170-degree omnidirectional field of view, which enabled rotating in each direction, so a wide range of adjustments was effective. Each display is driven directly by an Nvidia GeForce 4 Ti4600 graphics card, at 1280x1024 resolution. The aggregate pixel count for the entire display matrix is thus 64 megapixels, distributed over some 55 square feet of screen real estate.

Each graphics card is housed in a dual-CPU AMD Athlon MP2000+ rack-mounted slave node. The slave nodes each have a 100GB IDE disk, thus providing aggregate storage of 5TB. The slaves are driven by a similarly configured master node, and all communication is via Fast (100 BaseT) Ethernet, coordinated by a pair of Cisco Catalyst 2950 G- 48-EI switches [90].

Improvements in technology are now so extensive that several LHRD systems around the world now exceed a hundred million pixels. *Stallion* (Figure 2.8), the display wall in TACC ACES Visualization Laboratory at University of Texas at Austin consists of 307 million pixels on seventy-five LCD monitors [72]. The machine contains a total of 100 processing cores, 108 GB aggregate RAM and 46 render-node GPUs with 36 GB aggregate graphics RAM. The render nodes of Stallion are Dell XPS “gaming boxes”, each with two NVIDIA GeForce gaming cards. Of the seventy-five display tiles, fifty-eight share a GPU with another tile and seventeen have a dedicated GPU. These seventeen tiles are centrally located in the display, creating a “hot-spot” with increased rendering performance [72].

Figure 2.8. Stallion at the University of Texas at Austin

CAVE™ (CAVE Automatic Virtual Environment) is a projection-based Virtual Reality (VR) system that surrounds viewers in an immersive environment with four or more large display walls [75]. The Varrier display (Figure 2.9), a successor to this system is a 35-panel tiled system driven by a Linux cluster developed by Electronic Visualization Laboratory at University of Illinois at Chicago (EVL-UIC), the same group who developed CAVE as well. Two display panels are powered by one computation node via a dual-head nVidia Quadro FX3000 graphics card. One additional node serves as the master for the entire system. The 35-panel system is composed of 19 nodes, each containing dual Intel Xeon processors, connected by Gigabit Ethernet. Applications are built around the CAVELib™ platform. The display panels are mounted in a semicircular arrangement to partially encompass the viewer, affording approximately 120° to 180° field of view. The number of panels is scalable so that coverage up to 360° is theoretically possible. The net resolution is approximately 2500x6000, or 15 megapixels [89].

Figure 2.9. The Varrier display has 35 panels mounted in a semi-circular arrangement.

The Reality Deck, the first gigapixel resolution display in a fully enclosed setting, is an example of an immersive system. The Reality Deck is a visualization facility offering state-of-the-art aggregate resolution and immersion. It is a 1.5-Gpixel immersive tiled display with a full 360-degree horizontal field of view [76]. Comprising 416 high-density LED backlit LCD displays, it visualizes gigapixel-resolution data while providing 20/20 visual acuity for most of the visualization space. The Reality Deck offers a massive workspace of approximately 33 × 19 × 11 ft. The creators considered resolution, bezel size, display size, image quality and stereo support while picking the material for building the system. They also reported that at the time of construction, no commercially available display satisfied all the criteria; however, the Samsung S27A850D provided a good balance. It is a professional 27” PLS panel with 2,560 × 1,440 resolution with desired contrast, color saturation, and viewing angles. Unlike CCFL (cold-cathode fluorescent lamp) monitors, the S27A850D uses LED backlighting, which significantly reduces the weight and power requirements. Then, they modified the monitors with a custom mount to reduce the bezel to 14 mm. Figure 2.10 is a synthetic, to-scale view of the Reality Deck.

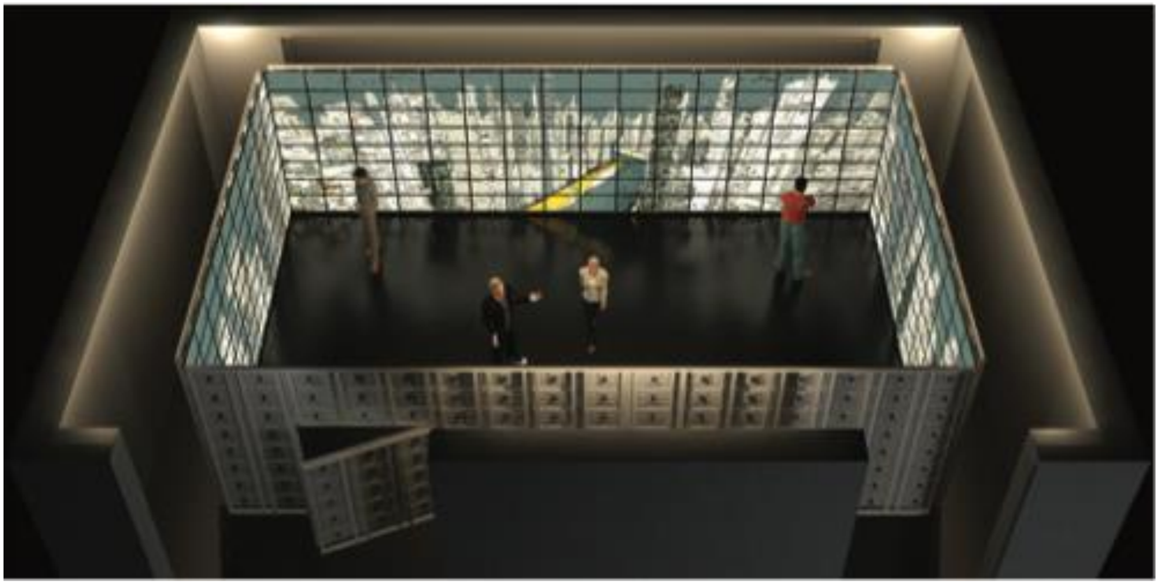


Figure 2.10. View of the immersive gigapixel Reality Deck facility displaying a geometric model of future New York City

The progression of the systems outlined above demonstrates the improvements in display system performance over time. However, improved system specifications do not necessarily mean improved effectiveness or usability which remains a central concern. When it comes to a specific use-case, several questions arise for many LHRD systems: For example, if the user of the system has a fixed position, problems of field of view and distance to center and sides of the wall occur. If the user is expected to move around, then fatigue due to potentially excessive physical navigation [3] and losing the full context by physical zooming can become issues. For groups of users a LHRD offers a platform for collaboration through state-of-the-art aggregate resolution and immersion. However, for a single user an LHRD might not be as ideal, an issue we explore in this thesis. Due to such factors, building large display systems is not a straightforward process – there are many trade-offs involved in the design. LHRDs have many specifications that must be considered individually. I can safely state that building such systems is a challenging act of balancing costs, features, and feasibility.

2.4. Closely Related Research

Large displays have been used for many purposes of course, not only for VA. However, since my study focuses on VA research with large displays I have singled out the most closely related previous work to examine more in detail. I will examine each of these in terms of their purpose in using large display systems and the tasks, physical characteristics of the system and the structure, research goals, and the results of the works.

Spatial representations of data on large displays are expected to aid understanding, the basis for my research. However, the question *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 [81].

They described two experiments. In the first, they investigated how learning performance and learner strategies are affected by the spatial distribution of information in a visual presentation, and interactive control over information viewing. They hypothesized that a spatially distributed layout would support superior learning performance in comparison to a non-spatial one, i.e., a slideshow layout, and that interactive, user-controlled viewing would improve task performance over automatic viewing, where the users are looking at static images that are depicting information about certain activities. Surprisingly the results of the first experiment did not support their hypotheses. Learning scores were significantly lower in the distributed layout than in the slideshow-style presentations, even though the participants were better able to remember associated locations for event cards with the distributed layout [81]. Also, there were no significant differences due to the viewing mode. In the first experiment, where they focused on studying learning differences due to varying levels of spatial distribution, the researchers did not maintain persistent visibility of content, by making the images disappear after showing for a while.

In the second experiment, the work studied how persistent visibility affects learning performance and learning strategies. The authors hypothesized that a distributed presentation (spatial layout) with persistent information visibility would allow learners to use the locations of the spatial layout to help organize information and aid recall [81]. This time, results confirmed that learning scores with the persistent-visibility distributed layout were superior to the automatic and interactive distributed presentations, as hypothesized.

In conclusion, the results of the study were quite surprising and different from the initial expectations. Interestingly, spatial distribution of the information did not help learning but persistent visibility did. My work uses a combination of both aspects, which will present an opportunity to compare my study outcomes with this research. Ragan et al. [81] planned to further study the effects of spatial information distribution with large display systems and future work includes considerations for the size of the data set, type of data representation, organization of information, and type of interactivity [81]. My study extends the results in visual analytics on large high-resolution displays (LHRDs) and contribute to spatial analysis of the information.

Another study from Virginia Tech investigated the effects of information layout, screen size, and field of view on user performance in information-rich virtual environments [80]. The paper describes an experimental evaluation of Information-Rich Virtual Environment (IRVE) interfaces. “Information-rich” virtual environments consist not only of three-dimensional graphics and other spatial data, but also include information of an abstract or symbolic nature that is related to the space [21]. They designed and evaluated two information layout techniques to support search and comparison tasks. During the experiments, two groups were tested: Those experiencing a single monitor condition and those that used a tiled nine-panel large display. 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?” However, the differences in time to completion across display configurations were not statistically related to task performance. They also found no statistically significant effect of display configuration on user accuracy. The authors mentioned that a potential reason for this was the reduced rendering speed, i.e., reduced frame rates. When the browser that the users were working on was enlarged to 1.5 x 1.5 screens or greater, the application switched to a software rendering mode which seemed significantly slower. Furthermore, changing the display mode did not yield the same results as the traditional single desktop monitor. Display size interacted with measured variables for accuracy. The best performing combination on a small display was the one called “Viewport Space”; however, it was outperformed by another called “Object Space” on the nine-screen display [80]. These results were not expected, and I will compare them with the results of my experiments in Chapter 7. I believe that the conclusion of this study might have been affected by unexpected and/or uncontrolled conditions. Successful transfer of an interface to a larger display is not simply a matter of scaling; many factors such as software design, hardware construction, interaction types and ergonomics require consideration.

2.4.1. Significance of Display Size and Resolution

Arguably the closest work to this thesis is a paper from Ni, Bowman and Chen, which suggests that increased display size and resolution improve task performance in information-rich virtual environments [74]. They designed a controlled experiment to evaluate the individual and combined effects of display size and resolution on task performance in an Information-Rich Virtual Environment (IRVE). They had three systems that they used in their studies. The results show that among a high-resolution small monitor, rear-projected screen, and a tiled high-resolution display system, users were most successful at performing IRVE search and comparison tasks on large high-resolution displays. In addition, users working with large displays became less reliant on wayfinding aids to form spatial knowledge [74].

While these authors investigated 3D spatial performance, which is fundamentally different from the VA activities that are the focus of my thesis, the research questions are very similar. Thus, I 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 desktop display (Figure 2.11, top left) is small and can be used at both low and high resolutions; a single projector (Figure 2.11, top right) provides a large, low-resolution display; and an array of projectors (Figure 2.11, bottom) produces a large, high-resolution display.

Figure 2.11. Apparatus used by Ni et al. [74] to study display size and resolution. Clockwise from top left: High-resolution IBM T221 LCD flat panel, Rear-projected screen, and VisBlocks tiled high-resolution display module.

In their experiments, the authors tried to keep *size* and *resolution* as independent variables. As users were sitting at the same distance, the larger display size yielded a larger physical field of view (PFOV Consequently, they adopted a 2 (size) x 2 (resolution) between-subjects experimental design, with eight participants assigned to each experiment condition (Table 2.2). To test the hypothesis that “users are more effective on IRVE tasks, if they have a wayfinding aid", half the users in each condition were provided a wayfinding aid in the form of a map that was on the displays, and the rest were not given.

	S-group	L-group
	Small-size (42.8° FOV)	Large-size (90° FOV)
Low-resolution (1280×720)	– Map	– Map
	+ Map	+ Map
High-resolution (2560×1440)	– Map	– Map
	+ Map	+ Map

Table 2.1. Experimental Design of Ni et al. [74].

Ni et al.'s experimental findings demonstrated the advantages of increased size and resolution [74]. As a general guideline, the large high-resolution display was the preferred choice for IRVE applications, since it facilitates both spatial navigation and information gathering. Interestingly, in their experiment, a large low-resolution rear-projected screen outperformed a regular-sized monitor with a higher resolution. However, I am not sure if this finding can be generalized and similar results might not be observed in our study, since the tasks in Ni et al.'s work differed from my proposed tasks and text was less legible in their case too.

A noteworthy contribution of Ni et al.'s study is their generalizable and reusable experimental design, which seems highly useful for large high-resolution display system research. Their design allows researchers to evaluate the effects of size and resolution independently, and any interactions between these two variables, if present.

Lastly, the authors also ask in their future work section, if there is an upper bound to content beyond which users will be overwhelmed by the displayed information. I investigate this aspect in my studies in terms of "information density".

2.4.2. Significance of Spatialization of the Content

In the first part of my study, I encourage users to cluster the information, then I observe and analyze how they spatialize the available content on large displays during their problem-solving session.

Previous research indicated that clustering information reduces the amount of visual search needed to find the elements required for problem-solving inference [88], which then translates to better performance in analytical tasks. Another study shows that a spatial contiguity effect applies to how deeply a user learns [97], 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 needs to be considered, since training is required of participants using a system for the first time [92], especially for analysis tasks that require tool understanding and analytical thinking patterns [66].

Endert et al. presented the results of a study, where users were asked to perform a spatial sense-making task on a large, high-resolution display in the LightSPIRE system (shown in Figure 2.12). Their study has also much common ground and shares methodology with the work presented in this thesis. The purpose of Endert et al.'s study was to analyze users' spatial clustering of information in a sense-making task. Study participants analyzed a textual dataset to understand and uncover a fictional terrorist activity using a simple spatial document organizational tool called LightSPIRE [44]. I am primarily interested in the analysis of spatial layout of information and in their research question about what structure exists within the user-generated clusters.

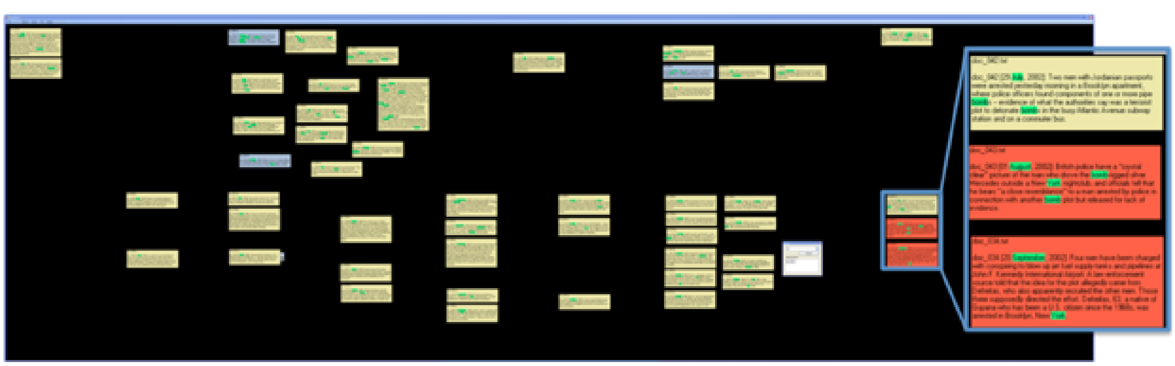


Figure 2.12. Endert's LightSPIRE [44], a large-display spatial workspace for organizing text documents.

For the study, Endert et al. used a large display system consisting of ten 17” LCD monitors in a 5x2 grid with a 13.1-megapixel workspace. The intelligence analysis training dataset used for this study consisted of 50 textual documents containing a hidden fictitious terrorist plot (VAST 2007 contest problem). The study involved 15 participants, which were all male undergraduate computer science students. The analysis of spatial layout on session completion consisted of characterizing the primary spatial layout and examining cluster structures, as both measures identify interesting aspects of users’ clustering approaches.

The analysis of overall spatial layout revealed three distinct patterns of spatial organization by users: Topical clustering, temporal clustering and hybrid clustering [44]. Most participants chose to organize their workspace primarily based on clusters of topically related documents; while a third of the users considered temporal information when organizing their workspace. A single participant used hybrid clustering to balance temporal awareness and understanding of topical relations.

Additionally, Endert’s study kept track of the total number of clusters created and the number of documents contained in each cluster [44]. Going beyond their work, there are other interesting characteristics of clusters that could be investigated, such as intra-cluster co-occurrences and transitivity. Transitivity expresses how connected different clusters are. I would like to observe if different clusters are independent from each other, or whether their positioning or the distance between symbolizes any relation. The experimental work in this thesis will consider those as well.

I believe it is important to observe how users interactively organize their information (unaided by any algorithm). Thus, my experimental designs do not provide any algorithmic clustering to users.

Spatial organization does not necessarily take place the same way on large displays [4]. One argument for this is that the required type of navigation is different as compared to conventional monitors. Previous work has shown that users organized content quite differently in a spatial way when there was a *physical workspace* (which requires more physical navigation then virtual navigation) such as a LHRD, as compared to a *virtual workspace,* where the users need to navigate through using virtual navigation techniques [4].

Andrews and North [4] explicitly examined the use of spatial sense-making techniques within two environments; one being a 33-megapixel large, high-resolution display and the other a small desktop monitor. Their study task used a basic analytic tool relying on manual spatial organization as the primary evidence marshaling technique[2]. 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 discussed 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. [37]”

The following quote summarizes Endert et al.’s findings [44]:

> *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. … 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…. (W)e 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.*

2.5. Methodology

Ball et al. [10] reported an observational analysis of the use of a large tiled display over the course of six months, which provides insightful feedback on how users do and do not use a large high-resolution display. They summarize the advantages and disadvantages of using tiled high-resolution displays and presented recommendations and guidelines for application designers [75].

[2] which was the same basic text analysis task that they used in their previous explorations of large displays for sense-making [2], and Robinson used in his studies [85].

In contrast to this qualitative approach, Tan et al. [105,106] conducted two studies to quantify benefits of physically large displays for individual users. They used controlled experiments, identifying independent and dependent variables explicitly while holding constant other factors. The angular visual field for each of two displays in their experimental design was the same. Participants performed spatial orientation tasks involving static 2D scenes. A significant performance gain was observed on the large display, even though the two displays created identically-sized retinal images.

My research benefits from both methodologies as discussed in the following chapters.

Chapter 3.

System Design

The system consists of *physical displays (V4-Space)*, *a software tool (DynSpace+)*, and *datasets* to be analyzed. Naturally all factors within the system that the user interacts with should be well understood and carefully designed.

3.1. Considerations for Designing the Display System: V4-SPACE

Transitioning to very large displays is not just a matter of adjusting the size of the monitors. Very large displays are a fundamentally different environment, for which technical aspects and usability considerations must be reassessed [3]. Using such a display system creates a new collection of design opportunities, issues, and challenges [3]. Tackling those is one of the side goals of my work.

There are many design aspects to consider when building a large display system. Different kind of systems have been constructed to meet different needs and design decisions. Surely, every system was expected to satisfy certain criteria for researchers. What advantages or disadvantages did previous systems have over each other? When we are aiming to build our own system with a specific purpose, what should we base our design decisions on? I will briefly discuss major design related questions to exemplify trade-offs when building a system for large display related research.

- How should we decide whether projections or monitor displays should be used? Projectors have advantages, such as seamless images that tiled monitor displays cannot provide. However, they have higher purchase costs, maintenance costs for bulbs, larger space needs and projector alignment requirements [72]. The screen material raises even more sub-questions. Rear projection or front projection technology? For display-based systems, should CRT, LCD, LED or OLED monitors be used? What should the size of each unit in the display system be? Small, traditional monitors can be obtained and carried around more easily, and smaller building blocks would make it easier to match desired physical dimensions. Yet this introduces an increased number of bezels, which would cause several problems such as spatial distortion [3], and a less coordinated system due to higher number of uneven display units since each unit is individually different from each other and cannot be expected to match exactly in terms of color, contrast and brightness. Moreover, a larger number of individual monitors increases system complexity.

- What degree of virtual and physical navigation techniques will be used? Will the user sit in front of the display and control the system remotely or will they be free to walk around and control the content more directly, e.g., through touchscreen interfaces? What will the input devices or navigation interfaces be? Physical navigation has advantages, such as better engagement with the system, immersion with the task, uses more intuitive ways of interaction, and ensures a focus by physical movements towards the display. However, drawbacks include increased fatigue and lack of generality of the cases it can be used in [11]. A moving user performing physical navigation faces the risk of not being able to have all the content within the field of view, and thus may not get a good overview.

- Physical attributes of the designed systems can vary greatly. A curved or a flat display surface? A curved display can bring all the displayed content to the same visual distance but can also be harder to build. If it is a display wall, how large should it be? Larger displays could be argued to bring extra power to visualize the content, though it might require additional degrees or amount of physical navigation, which potentially could slow the analytics process and hinder productivity. Fully enclosed, CAVE-alike systems would have advantages in cases where multiple users use the systems at the same time, though might affect one-person tasks negatively due to inefficient distribution of the content over the space.

All such design considerations are associated with trade-offs for system builders, and the decisions should be made carefully by evaluating all aspects and comparing those with the intended use cases. The motivation behind the design decisions for our large display system is shown in section 3.1.2.

After setting design goals in alignment with the expectations for the system regarding the tasks, the system gets designed and built. I prepared the following list of questions to be answered to aid in developing the system specifications. Specifications for our system, V4-SPACE, are then given in section 3.1.1.1.

- What is the physical size including the total dimensions in height and width?
- What is the resolution, and the aspect ratio?
- Is there a lower limit for desired pixels per inch (PPI) in the display?
- What is it made of; projection surfaces or monitors?
- What are the requirements for contrast, brightness, color?
- How many individual units does the display system consist of?
- If it is not a single display, how are the units separated?
- Are there any bezels or apparent joints, how large are those and is that distractive?
- What is the organization; flat, curved, fully enclosed?
- How are the monitors driven; by a single computer or a cluster of machines?
- What are the requirements for the machines that run the displays?
- What software is used?
- What are the means of interaction?
- Does the user have a fixed position relative to the display system?

- If yes, what is the distance to the displays and what is the angular field of view?
- Does the system require physical navigation and what degree, if it does?
- Does the system qualify as a *retina display*? While the answer to this question can be inferred from the above information, it is useful to directly list this criterion as it qualifies when the content becomes pixelated.

There are also various limitations that should be addressed when designing a large display system. Limitations can be technical, technological, physical, cost related and so on. For example, an nVidia Mosaic display system run by a computer is currently limited to 16384 horizontal and vertical pixels by current nVidia drivers and hardware. A MS Word file cannot currently be opened in full screen mode on displays with more than 8k pixels. A 4K image, which would be considered as a high definition image on traditional monitors, takes up a very small amount of display on a LHRD, and thus occupies a relatively small field of view (FOV), if the user is at distance to the display system that enables him/her to see the whole display at once.

Such examples show that the underlying technology is still not very supportive of large display systems to for various tasks including visual analytics. These limitations also determine what could be realistically expected from large display systems and differentiate possible solutions from ideal ones. A “perfect” system using a *single* curved very large UHD OLED (ultra-high definition organic light emitting display) monitor, say at 16k x 9k resolution, would probably be hard to build, order, and ship, probably exceeds current graphics hardware capabilities, and is less likely to be affordable, too. Therefore, it is crucial to compare realistic targets in a cost sensitive manner and make design decisions considering all mentioned factors.

In general, researchers should ask what the motivations for design choices are, what the resulting trade-offs are, what conditions have been considered when designing the system and what limitations apply to their design attempt.

3.1.1. The System

3.1.1.1 V4-SPACE at a Glance

V4-SPACE is a very large display system in the VVISE Lab at the School of Interactive Arts and Technology (SIAT), Simon Fraser University (SFU). It is a versatile display system environment designed to support multiple projects (Figure 3.1). Moreover, it is expected to be expanded with additional features, improved, used for comparative and other VA studies on LHRDs, and potentially even field studies. However, for the purposes of this thesis, the current state of the system has been designed for the VA study in this thesis.

My research uses V4-SPACE to support work on VA tasks. My objective was to create a good environment for an analyst by providing a system focused on supporting a single user. In such an environment, my goal was to observe and analyze how physical size, resolution, and the spatialization of the content on the available display space affect user performance on LHRDs during VA tasks. The goal was to provide a very powerful system to a VA user, at a cost level that will soon be within reach for companies or institutions that do high-value VA work.

To facilitate adoption, V4-Space is run by a single, powerful computer with multiple powerful graphic cards. This avoids complicated hardware configurations with a cluster of computers or complex software-based platforms. While the current instantiation costs more than CA$100,000 (cumulative cost of the system), we can expect that the falling costs of large display systems will make the system more affordable and cost-effective in the future.

Figure 3.1. V4-SPACE, the display system used in this study.

Description of V4-SPACE:

- A 1x7 array of large, tiled monitors, which form the main display. The monitors are 85” 4K Samsung Smart TVs, oriented vertically. An additional 21” monitor sits in front of the user below the view of the monitors, i.e., does not block the view of the main display, and is designed to be used for auxiliary tasks.

- The main monitors have stereoscopic display capability. However, this feature has not been used in this research.

- Currently the system is in a circular arc around the user such that the user is equidistant from each display unit in the system. Consequently, the system affords a uniform (perceived) font size across the whole display, which avoids information legibility issues due to non-uniform distances.

- Each 85” monitor displays an array of 3840 x 2160 pixels. Together the LHRD system has 15120 x 3840 pixels. In total, this makes V4-Space a 58 megapixel display system, with 58,060,800 pixels. While the aspect ratio of each single display is 9:16, for the whole system it is 63:16, very close to 4:1.

- Each display in the grid is 1.88 m tall and 1.06 m wide. This translates to 74" and 41.67", respectively.

- The whole system is 7.41 m by 1.88 m. As a single unit, it is a 7.65 m diagonal display with a display area of approximately 14 m^2.

- V4-SPACE has 52 PPI pixel density.

- One of the biggest motivations for using larger displays for us was to eliminate bezels as much as possible within the available budget and available technologies by using largest commercially available units of display. With our 1x7 grid display, we eliminated all horizontal bezels and have only 6 vertical bezels due to the outer covering frames of the monitors in the system.

- The system is run by a single computer, with an Intel i7-6700K processor at 4GHz and a Gigabyte GA-Z170X G1 motherboard, which features four PCI Express Gen 3 slots. To operate the large displays, there are two nVidia Quadro M5000 graphic cards, which provide four 4K outputs each. Hardware refresh across all outputs is synchronized through an nVidia Quadro Sync card. In addition, there is an nVidia Quadro K620 graphic card for the auxiliary desktop monitor.

- V4-Space relies on the nVidia Mosaic driver functionality, which presents all seven displays as a single display surface to the operating system. This greatly facilitates software development, thus the display functions like any other monitor.

- The Operating System is Windows 7.

- The system is designed for a single user who has a fixed position in front of the display system. As a limited form of physical navigation, the user may rotate a swivel chair to look at different parts of V4-Space. The system is not designed for the user to stand up and walk while using V4-SPACE for VA tasks.

- Interaction with V4-Space is through keyboard and mouse. Users are already very familiar with this type of interaction modality, which greatly reduces training requirements for using V4-Space.

- A standard keyboard is provided to the user. To support the high resolution of V4-Space, we use a high-performance mouse, a Razer DeathAdder Chroma 10000 PPI optical gaming mouse. The motivation for this choice is that regular basic mice are not sufficiently sensitive to afford efficient mouse navigation across 15120 horizontal pixels.

- The chair is placed such that a spot midway between the user's eyes is 3.3 m from the front face of the display (130"). Thanks to the curved positioning of the TV units in the display system, this distance is to the same for all units of the display.

- When determining the distance, I specified PPD, then calculated screen-to-user distance to achieve that PPD. This will be further discussed in section 3.1.2.5.6.

- From the fixed distance, the display system takes up approximately 131° of the user's physical horizontal field of view.

Finally, V4-SPACE is a highly re-configurable system. Details of the re-configurability aspects are given in Appendix B.

3.1.1.2 Advantages over Past Systems

Both experience from past research as well as recent advances in technology make more powerful display systems more easily available than they have been in the past. In general, and compared to the first efforts in the field, constructing large display systems takes less effort and fewer resources and involves fewer technical challenges. Moreover, past work informs today's researchers regarding effective ways of constructing displays and guiding them in designing such systems. Although large display usage in VA tasks has great potential, due to newer and cheaper technology, it is only now that we can build better systems than what was possible in the past which enables me to address some of the issues that existed in previously built large display systems that were discussed in section 2.4.

Due to technology advances, it is not reasonable to compare our system with very early attempts of larger displays. Bishop and Welch's system had low resolution [19]. DSharp [32] is in a sense comparable to modern curved 4K TVs. Chris North and Robert Ball's systems [10,99] used 17" monitors driven by clusters of computers. Not only does this increase system complexity, it also results in many horizontal and vertical bezels, as does NASA's 64-megapixel Hyperwall [90]. Bi and Balakrishnan's system [18], used 18 projectors with 1024x768 resolution which exhibited quite visible color pattern changes over the images coming from different projectors.

More recent systems became progressively more sophisticated. The 34.5 million-pixel display wall at University of Calgary was operated by a single computer and provided almost 10 square-meters area of display with 50 PPI density through rear projection technology [82]. However, as discussed earlier, projection-based systems have disadvantages and the Calgary system is no exception. One disadvantage is that bulbs for the projectors need to be replaced periodically, which increases maintenance effort and cost. Also, projector-based systems typically cannot match actual displays in terms of color, brightness and contrast levels due to the nature of the technology [72]. Projection displays also require specific environment and dark surroundings to provide the best quality. Finally, it is harder to reconfigure a projection-based system.

As discussed previously, a better display cannot be simply specified as the one that has the most pixels. Still, some systems reached quite impressive quantities of pixels. Stallion, for example, at University of Texas at Austin consists of 307 million pixels [72]. However, little was reported about use cases, how it was useful in specific tasks and what the advantages of the system were. Rather the system was a flat surface with very many screens brought together to create a system with great potential, but not converted into a specific solution to any specific research problems.

Lastly, some large display systems offer unique environments for users that do not necessarily address our use cases. The Varrier display (see section 2.3) emphasizes autostereoscopic technology to produce a VR immersive experience without requiring encumbrances such as glasses [89], targeting primarily 3D content. Similarly, the gigapixel Reality Deck aims to be a fully immersive system in an enclosed setting [76]. It targets groups working within the Deck and relies on physical user navigation to operate on the system. In contrast, V4-SPACE targets a single user who interacts with the system directly with the provided means of interaction, has the entire display all the time within their field of view, and does not require much fatigue-inducing physical navigation.

3.1.2. Design Challenges and Preferences

Design choices involve judging the available options, optimizing resources with respect to requirements and making decisions about known trade-offs. As Andrews et al. suggest [3], since the field started to explore LHRD systems, we are shifting away from the technical limitations of scalability imposed by traditional displays (e.g. number of pixels) to studying the human abilities that emerge when these limitations are removed. This exposes the potential benefits of large, high-resolution displays. A better understanding of the interaction between visualization design, perception, interaction techniques, and the display technology will help to identify sweet spots for LHRD technology. My focus is on studying LHRDs used for VA.

An extensive survey on LHRD technologies, techniques and applications by Ni et al. [75] developed a list of top ten research challenges (Table 3.1). The set of design issues we identified in section 3.1 for the creation of V4-SPACE mostly aligns with this list and we considered all the aspects as much as possible. Our decisions are discussed below.

Table 3.1. Top Ten LHRD Research Challenges listed by Ni et al. [75]

1. Truly seamless tiled displays.
2. Stereoscopic large high-resolution displays.
3. Easily reconfigurable large high-resolution displays.
4. High-performance cluster rendering.
5. Scalability.
6. Design and evaluate large high-resolution display groupware.
7. Effective interaction techniques.
8. Perceptually valid ways of presenting information on the large displays.
9. Empirical evidence for the benefits of large high-resolution displays.
10. Integrating large high-resolution displays into a seamless computing environment.

3.1.2.1 Use Case Scenario

The most important question regarding the system is how it will be used. What is the use-case? What is it built for? Who is it built for? How will it be used? Here we use the V4-SPACE for VA tasks. In past research and real-world practice, many different VA tasks have been used and/or evaluated. Some of the work targets individuals, others target collaborative work around VA. The types of collaboration, interaction and navigation differed significantly.

Other work has attempted to investigate large display system usage for data manipulation tasks [64], and the effects of display size and navigation technique on classification tasks [63]. However, these studies were performed in shared interaction settings. In real-life scenarios, VA processes consist of certain stages with different characteristics. When I was working as an analyst on a VA project, and even though other stakeholders were co-located with me, many VA tasks were handled solely by myself, and I spent most of the time working on the main VA task during the project. Still, there was collaboration at the beginning of the analysis, for reviewing progress and addressing problems during analysis, and for the discussion of results, mostly at the end of the task. Although such group sessions offer unique features and advantages for analytical tasks, in many real-life situations sustainable, long-duration analytical processes are associated with a single user setting [6,7,36] with a fixed, typically seated position. Consequently, even though our highly reconfigurable very large display system could also allow multi-user interactions, we decided to design and optimize the system to be used by a single user for this study so that the core part of the VA process, which is performed by a single user, is studied.

I decided to place the user behind a desk in front of the display system and placed the displays all at the same distance from the user. Since the system will be controlled by a single user, keyboard and mouse was used as input method.

To support the single-user scenario, the user of V4-SPACE sits on a swivel chair, so they can rotate to visually cover the whole area of the display system while performing the visual analytics task. With this limited form of physical navigation, the user does not need to get up and walk to the displays, since they are all at the same distance from the sitting location in a semi-cylindrical configuration. I also adjusted font sizes and the size of user interface elements to allow optimal reading conditions for this given viewing distances. I set the font size to 28 pt to match the perceived size of 10 pt font text on a page of paper at a reasonable reading distance (approximately arm's length). The user opens the visual analytics tool on the LHRDs in full screen mode to undertake visual analytics tasks. The tool and the tasks will be discussed in detail in the following chapters.

The auxiliary display is used to show reference documents and similar information which ensures that such documents and tools do not affect the workspace allocated for the VA tool.

The user performs VA tasks using the keyboard and high PPI mouse, which enables covering the large horizontal display space with ease. The user can work on the task for long periods of time since they will be sitting and not dealing with potential fatigue due to extended periods of standing. At the same time, the user uses eye, head and body rotations to cover the large viewing angle of the displays.

3.1.2.2 Navigation Technique

Navigation on displays can be either virtual, as in the traditional approach, or physical, which enables exploitation of embodied human abilities, such as spatial awareness, proprioception, and spatial memory [3]. There is a trade-off between virtual and physical navigation. In physical navigation, spatial understanding is higher, navigation is more direct, and the navigation UI can be simpler. However, the user must be able to change their physical location, e.g., by walking around. On the other hand, for virtual navigation, fatigue can be less of an issue, and many different interaction devices can be used [11]. Past research [11] indicates physical navigation offers some advantages over virtual navigation, especially for large display settings. Previous work on spatial visualizations argued that larger displays lead to more physical navigation, which offers improved user performance and reduces the need for virtual navigation [11].

Many strategies have been developed for displaying large amounts of information, including overview + detail, focus + context, and pan + zoom [29]. These techniques are used in virtual navigation, also some of them are occasionally used with physical navigation. Continuous, “intelligent” zooming techniques have been presented too, as in Figure 3.2 [35,94].

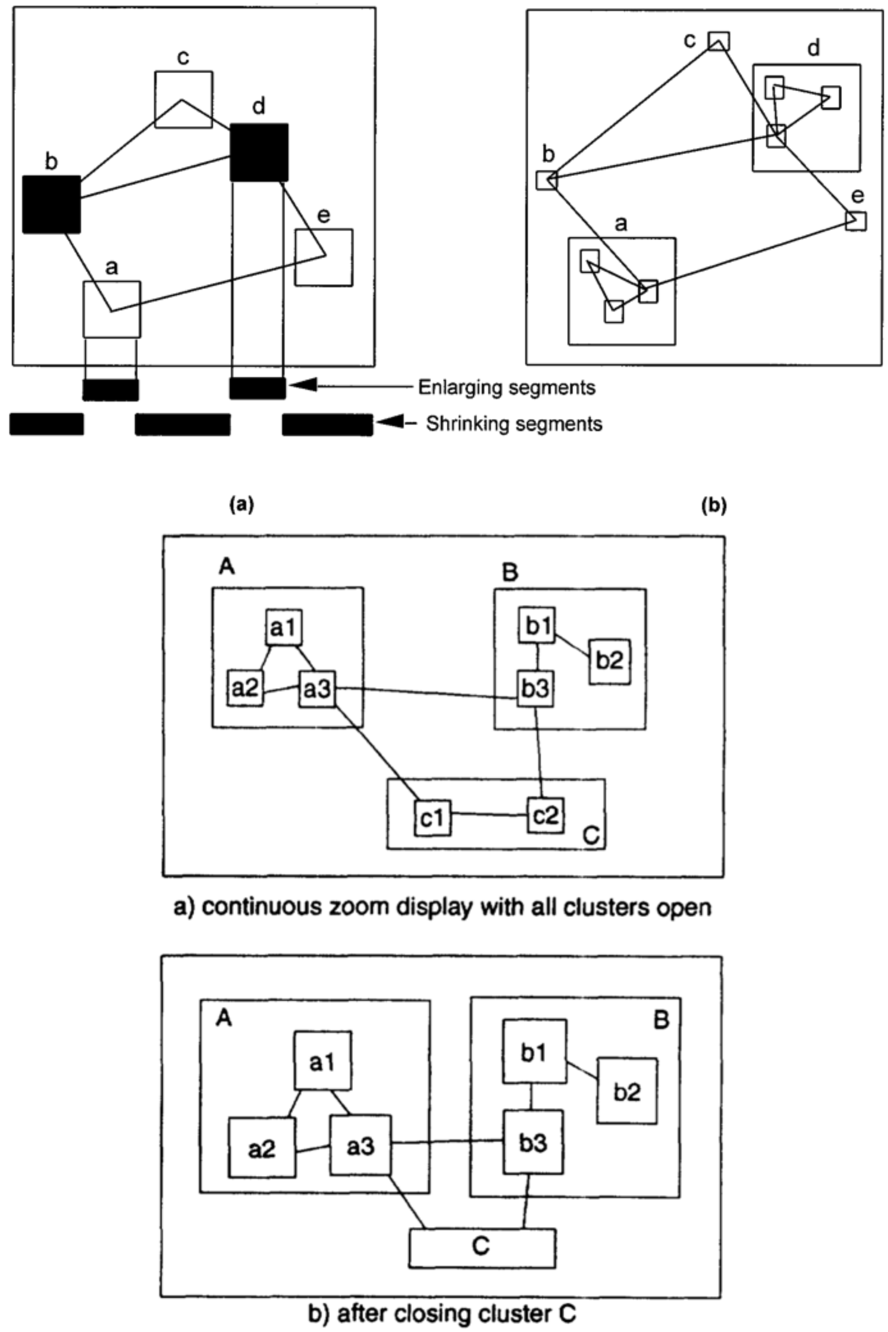


Figure 3.2. Intelligent zooming [35].On top, there is the example of basic operation; (a) network before zooming; (b) nodes a and d have been zoomed to show the subnetworks; other parts are shrunk as context [94]. At the bottom, there is the display of a network with one cluster (C) is open and closed in two consecutive views.

During a study by Ball and North [12] in which they investigated users' visualization task performance in a large 2D information space, users were given four conditions, designed with 2 independent variables and 2 levels for each. The systems that users worked on offered either physical or virtual navigation, and either a “context” or “focus” view. The results indicated that physical navigation was preferred over virtual navigation, as it promoted higher order thinking, such as investigating data points that held the most promise for solving a task, while virtual navigation promoted less efficient, simplistic algorithmic strategies. Physical navigation also lowered performance time.

Since V4-SPACE features a large area to lay out the information, more information can be displayed at a readable size, and less panning, zooming or scrolling takes place as compared to small monitors. Instead of using indirect techniques, V4-SPACE users can navigate through objects directly by simple physical movement, such as eye, head or body rotation, for scanning increasingly larger distances. There is no need for additional degrees of movement, which reduces the potential for fatigue. All standard interaction and navigation components on the displays, such as menus, scroll-bars, and buttons, can easily be controlled through the mouse rather than embodied navigation.

3.1.2.3 Input Devices

There have been attempts to improve interaction on large display systems, since the interaction becomes more difficult as displays get larger.

Bezerianos and Balakrishnan presented the design and evaluation of *the vacuum* [17], an interaction technique that enables quick access to items on areas of a large display that are difficult for a user to reach without significant physical movement. *The vacuum* is a circular widget with a user controllable wedge of influence that is centered at the widget’s point of invocation and spans out to the edges of the display. Far away objects residing inside this influence wedge are brought closer to the widget’s center in the form of proxies that can be manipulated in lieu of the original [17]. Results have shown that the vacuum outperforms existing techniques like a regular mouse when selecting multiple targets in a sequence, performs similarly to existing techniques when selecting single targets located moderately far away, and slightly worse with single targets located very far away in the presence of distractor targets along the path.

Some researchers have explored methods beyond virtual interaction. Khan et al. described a widget and interaction technique, known as a “*Frisbee*,” for interacting with areas of a large display that are difficult or impossible to access directly. A *frisbee* is simply a portal to another part of the display. It consists of a local "telescope" and a remote "target". The remote data surrounded by the target is drawn into the telescope and interactions performed within it are applied on the remote data [58]. Their results indicate that the *frisbee* is preferred over walking back and forth to the local and remote spaces beyond 1.37 meters (4.5 feet) [58].

In the current instantiation, interaction with V4-SPACE is through keyboard and mouse. With modern GUIs, this means that the user relies mostly on the mouse. Mouse control of very large display surfaces poses several challenges. Sufficient cursor size, movement speed, and easy long-distance targeting are the most notable ones. Previous research suggests that increased sizes of displays often affect how users handle their mice [14]. The problem arises from the need to increase mouse speed, as large displays feature many more pixels, which requires the mouse to cover larger distances, compared to the conventional monitors that the mouse was designed for. To avoid excessive clutching, users who want to get across the screen reasonably fast must set their mice to higher speeds or acceleration values, which introduces drawbacks. One problem is that higher mouse speeds make it harder for users to visually track the cursor and/or reacquire the cursor as it is approaching a target, such as a button the user wants to click. If users lose track of the cursor, they must spend extra time reacquiring it [14]. However, it is not only the cursor's speed that makes the cursor hard to track. Due to limitations of the operating system that renders the cursor or limitations in frame rate, it may appear to jump from one position to the next. This artifact, a form of temporal aliasing [33], becomes a problem as higher mouse speeds make the cursor jump farther.

Existing techniques to solve these problems, such as the Windows mouse trails, do not preserve the responsiveness of the mouse cursor on high-resolution display systems. Windows mouse trails were designed to enhance the visibility of the mouse cursor on LCDs with a slow response by leaving cursor images on the screen for two or more frames [14]. In my experience, they substantially slow interactivity down with very high resolutions, such as the 58 megapixel V4-Space display.

The general problem of long-distance target acquisition has been the subject of a much research on interaction techniques [23,41,71,100,114]. An important improvement, the *high-density cursor*, addresses this issue by filling in additional cursor images between actual cursor positions through temporal super-sampling [14]. The reported results of the corresponding user study indicated that the high-density cursor improved participants' performance on a Fitts' law task by up to 7% for target acquisitions across long distances. The differences found in the study were statistically significant but not very large.

To address the interaction problem on V4-SPACE I increased cursor speed and simultaneously increased the cursor size to address the drawbacks of the increased speed. To achieve the maximum possible size of the cursor in the Windows operating system I used “largest cursor ever” [125]. Since interacting with the system with a regular mouse did not yield an acceptable experience, we used a Razer DeathAdder Chroma mouse with a 10,000 ppi optical sensor [126] as the main input device.

3.1.2.4 Equipment Used

One of the central debates in very-large display system design is tiled display panels versus projector arrays. Some advantages for tiled LCD panels are: (1) they are easier to align and color correct, (2) maintenance costs are lower, and (3) they take less space (no throw distance needed). A disadvantage is borders between tiles due to the bezels, which we will discuss in detail in section 3.1.2.5.1. Bezels affect the display of imagery to a limited degree, but can affect the display of text more strongly [75].

Projector-based systems, on the other hand are typically cheaper to build in terms of square area of the display, but often cannot achieve the same image quality as LCD panels and need careful geometric calibration. In general, projection-based systems have also much lower resolution compared to LCD panels. Tiled projection systems also need additional system components to blend images from different projectors. Projection-based systems also typically have larger space needs in terms of the depth of the system.

My system uses very large display panels, which tackles all issues and minimizes the bezel effect. V4-SPACE consists of seven vertically oriented 85” displays, tiled horizontally. Due to the large vertical size of the displays, and in contrast to many other tiled LCD systems, this means there are no horizontal bezels. The monitors are Samsung 85” UHD TVs. UHD is the abbreviation for Ultra-high-definition television; which is also known as Ultra HD and UHDTV [27,39,55,122,128]. "Ultra HD" is an umbrella term selected by the Consumer Electronics Association [8]. There are two resolution configurations described as UHD, 7680x4320 and 3840x2160, referred as "8K" and "4K", respectively [129]. Our UHD TVs are 3840 × 2160 pixels tall (8.29 megapixels each), four times as many pixels as HD resolution, which is 1920 × 1080 pixels (2.07 megapixels). Our system provides 15120 pixels horizontally and 3840 vertically, which exceeds UHD and could be called "15K" resolution.

3.1.2.5 Physical Attributes

Many physical attributes must be considered. Here I use Andrews et al. set of such attributes to characterize V4-Space [3]:

size: measured in inches of the diagonal viewing area (e.g., 17 inches);

pixel density: measured in the number of pixels per inch [e.g., 96 PPI (pixels per inch)];

resolution: measured in a horizontal multiplied by a vertical pixel count (e.g., 1600 x 1200 px);

brightness: measures the amount of light emitted by the display (in candelas per square meter);

contrast: measured as the luminance ratio between the brightest and darkest color (e.g., 700:1);

bezels: the frames that surround conventional monitors;

display technology: the technology used to create the display, typically tiled LCD monitors or projectors (rear or front projected); and

form factor: the physical arrangement and form of the display.

We discuss our design considering these attributes in the following subsections.

3.1.2.5.1 Display Units

When the gigapixel display of the Reality Deck was being built, the researchers decided to use LCD monitors over projectors, due to factors such as lower maintenance costs, and better manageability and affordability [76]. Then they considered different types of LCD monitors based on the following criteria: resolution target, bezel number and size, unit size, image quality and stereo support [76].

Even though their number is limited in our system, bezels are the main issues with tiled multi-monitor display systems. While one study argues that larger bezels are causing small judgement errors for distances (around 5%) [110], another study from the same group focused on the fact that the additional costs associated with thinner bezels may not provide significant return on investment. They also found that bezels may act as visual anchors and be useful for placement of the interface elements [109]. Still, bezels might be useful under certain conditions, too. Other research has shown that bezels are sometimes actively used by the user for distribution of tasks among monitors and thus become useful in terms of dividing tasks into multiple subspaces [18,102].

Smith et al. present two experiments investigating bezels on large tiled displays used for fast-action gaming [102]. The first experiment examines bezel size independent of configuration, using simulated bezel sizes ranging from 0 (i.e., no bezels) to 4 cm in a 3x3 grid configuration. Results of the experiment indicate minimal effects for bezel size and compensation. The second experiment fixed bezel size at 4 cm and instead varies configuration from a single display up to a 3x3 grid of simulated displays. Results of this study indicate that while the 1x2 performed worse in certain metrics, globally, the effects of configuration were similarly small. The issue of bezels is still a great topic for future research.

Good image quality requires good contrast, backlight uniformity, and large viewing angle, and the high-end displays panels of V4-SPACE are state-of-the art in this respect.

3.1.2.5.2 Organization

Individual screens can easily be combined into wall-sized displays. Such combinations of displays could be straight, curved, or arranged in irregular patterns. With different organizations of the displays, one can even build a completely enclosed space, i.e., a CAVE. As there are many possibilities for arranging the screens, it is unclear what arrangements are “best” [62]. Moreover, there are many options for how to arrange content and applications across such large displays.

When we were organizing the individual display elements of V4-SPACE, we encountered certain trade-offs due to physical limitations. One can think of this as a puzzle, with 8 puzzle pieces (identical 4K TVs) of 85 inches with 3840x2160 pixels. One option was to create a 2x4 grid with horizontally positioned TVs, which would provide a 15360x4320 pixel configuration. The downside of this arrangement is that for a seated user there will be a horizontal bezel approximately at eye level. A bezel in such a prominent location is not desirable and thus we decided against this option. We could also have arranged the TVs vertically in a 2x4 grid to create a display with 8640x7680 total resolution. While this would create an almost square display surface, it would be very tall (almost 4 m high), which does not match the human visual field well and which also exceeds most room heights. Thus, we settled on a horizontal grid/row with the displays in portrait orientation. By installing the TVs on stands with casters, we can reconfigure the displays freely, which affords many degrees of freedom and increases the versatility to the system.

Once we had positioned the physical displays in the horizontal grid/row, we created a single, logical display using the nVidia Mosaic software. At this point we discovered a hardware/driver limit of a maximum of 16384 pixels (horizontally and vertically) for any Mosaic display. Eight displays exceeded this hard limit, so we used only 7 displays to get a display of 15120 horizontal and 3840 vertical pixels.

3.1.2.5.3 Shape / Curvature

The individual display units of V4-SPACE are organized along a semi-circular arc. Here we discuss why we chose this configuration.

Shupp et al. presented a study to analyze the effects of angled orientation on large displays on performance during 2D visualization tasks. They asked questions such as “How does curvature affect finding data in different locations on the display?”, “How does curvature affect comparing data at varying distances?”, and “How does curvature affect users’ abilities to reason about visualized data?” [99]. With a 24"-panel, 31.5 megapixel display in the flat and curved configurations, they found significant benefits of curving the display system. According to the authors, curving LHRDs provides the following benefits:

• Changes the amount and type of physical navigation, from translational movements to rotational ones. Translational movements decrease by about half, while rotational movements double.

• Further improves user performance time for some tasks, up to 24% faster.

• Eliminates users' significant region bias towards the left side of large flat displays.

• Changes the type of insights users gain from large-scale visualizations, to less overview insights and more localized detail insights.

• Reduces users' frustration levels on search tasks, and substantially increases user preference on comparison tasks.

A summary of the effects is given in Table 3.2 [99].

Table 3.2. Curved Displays versus Flat Display Surfaces.

Effect	Curved	Flat
Faster performance time (up to 24%)	X	
More turning (double)	X	
More walking (double)		X
Users change area of focus frequently	X	
No region biases	X	
More detail-level initial insights	X	
More overview-level initial insights		X
Less user frustration for search	X	
User preference for search	X	
User preference for comparison	X	
User preference for insight		X

Similar to their approach, I curved the display of our system so the user sits in the center of a semi-cylindrical display, and can maintain the same distance to information on all screens, which reduces or addresses several issues with large flat displays.

3.1.2.5.4 Display Size Viewing Angle

Previous studies about display size yielded interesting outcomes. A huge motivation for having large displays is the results of experiments which suggest that physically large displays, when viewed at identical visual angles as smaller ones (Figure

3.3), help users perform better on mental rotation tasks [106]. While the reasons for this are yet to be explored thoroughly, large displays do immerse users within the problem space and might bias them into using more efficient cognitive strategies.

Tan et al.'s study [106] is significant in this context. They kept the visual angle subtended from the user to each of two different displays constant by adjusting the viewing distances appropriately (Figure 3.3). They also held other factors such as resolution, refresh rate, color, brightness, contrast, and content as constant as possible across displays. Since the information content shown by each of the displays was identical, it would be reasonable to expect that there would be no difference in performance on one display or the other. However, quite interestingly, this was not the case, which suggests physical size is indeed an important display characteristic that must be considered as we create our display systems. Their results suggest that physically large displays, even at identical visual angles as small displays, increase performance on spatial tasks as well as mental map formation and memory. A visual representation of their experimental setup is given in Figure 3.3.

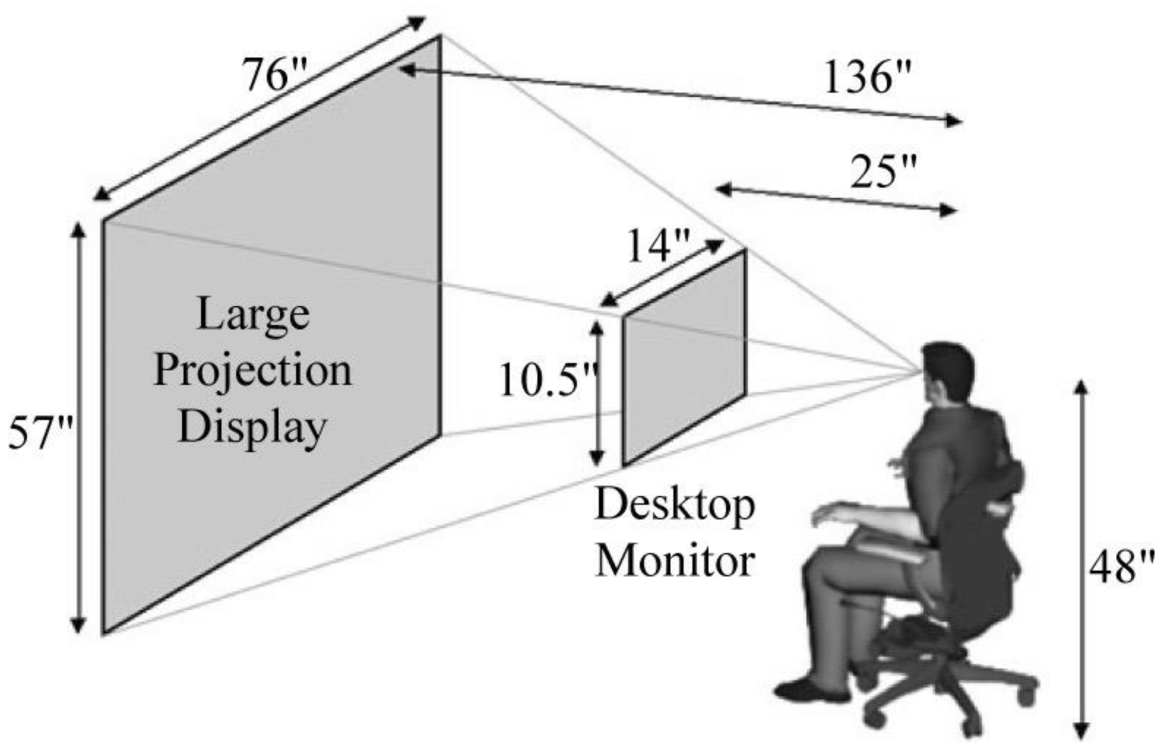


Figure 3.3. Experimental setup to maintain a constant visual angle between small and large displays [106].

This study provides strong motivation for some of my design decisions. As explained earlier, I keep all the information on the displays at a legible size (the perceived as a 10pt font text on an page at normal reading distance, at arm's length). Consequently, I do not rely on approaches that trade off space and speed, such as overview + detail, focus + context or pan + zooming. At this point, and if the perceived size of the content and viewing angle are going to stay same, one could ask why we did not place a smaller number of smaller high-resolution monitors closer to the user (say a couple of 27" monitors at arm's length), instead of larger displays at a larger distance. Our reason for this is Tan et al.'s evidence [106] that sheer display size, isolated from all other factors, improves user performance.

3.1.2.5.5 Field of View

Field of view (FOV) is related and connected to many attributes of a system. Geometrically, the horizontal FOV is the angle subtended at the human eye by the horizontal edges of the display [31], i.e. determined by the displays' horizontal physical dimension and the user's distance to the display. There is also a corresponding vertical field of view (VFOV), and diagonal field of view (DFOV), but for most situations only the horizontal FOV is of interest. Those are shown in Figure 3.4. Display manufacturers typically report DFOV numbers, as they are larger. However, as the main concern of my work is around horizontally wide displays, from here on I will use FOV to mean horizontal FOV.

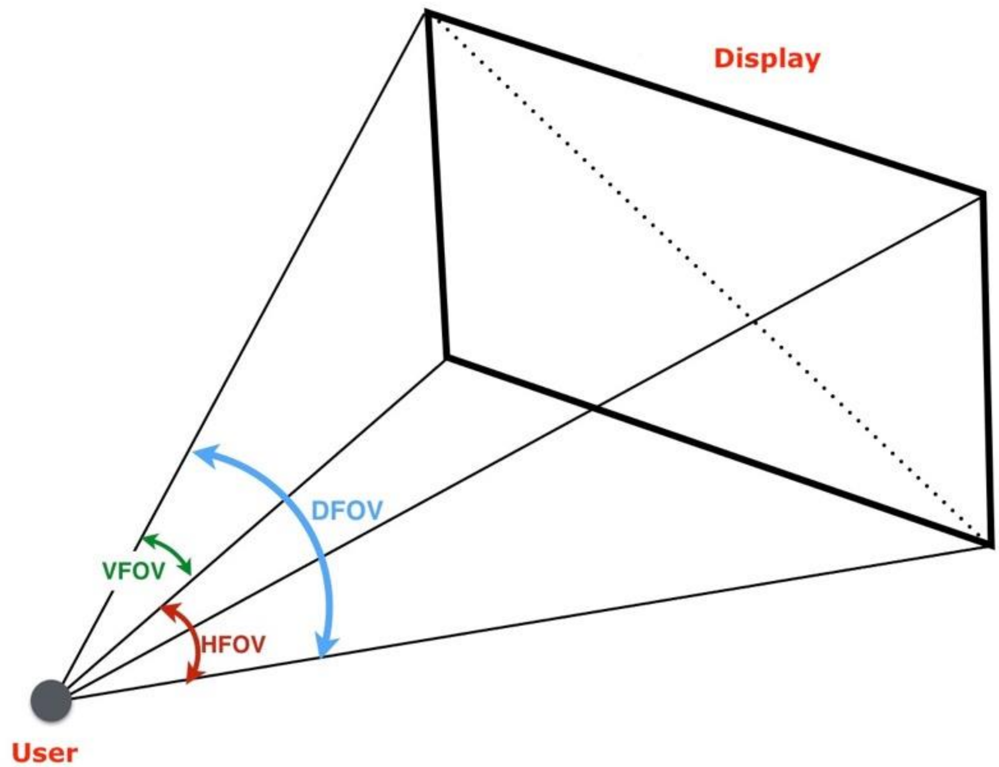


Figure 3.4. Vertical, diagonal and horizontal fields of view of the display.

The term FOV can also refer to physical field of view (PFOV) (or display field of view in other literature), or software field of view (SFOV). E.g., a display of 48” width placed 24” from a user’s eyes, will result in a horizontal PFOV of 90° [74]. The software field of view (SFOV) is used by 3D rendering software and *does not necessarily match* the PFOV, as the virtual camera may be at a difference distance to the display relative to the position of the user’s head. The effects of PFOV and SFOV on virtual environment task performance have been extensively studied [1,9,31,38,40,70,112], although much of the research was only concerned with virtual environments (VE). In my system, since my software sets the SFOV to be identical to the PFOV, I will not distinguish between the two terms.

Visual perception involves both the high acuity foveal vision and the wider overlapping peripheral vision. The role of peripheral vision in competent performance of adult visuomotor activities of walking, reaching, and forming a cognitive map of a room was examined in a study by limiting the field of view to certain degrees in each condition

[1]. In general, the results identify the benefits of higher degrees of PFOV, that is, large displays viewed by users in wide visual angles. However, there are still questions to answer, such as whether there is an upper bound beyond which users will be overwhelmed by the displayed information. In my study, I tried to find out if users of V4-SPACE prefer and perform better in the largest display setting, or whether they prefer to view the information at a smaller FOV.

The attributes of V4-SPACE that influence the user PFOV include the size of the display system, the curvature required to keep information at the same distance, and the pixels-per-degree that we want to preserve by placing the user at the correct distance, considering the resolution and display PPI. Across all seven displays in the system, V4-SPACE provides a PFOV of 131° to the user. With such a large viewing angle, I attempted to balance immersion and the constant visibility of all displayed information.

Considering the distance from the screens and the location of the user, the full width of V4-SPACE exceeds the *foveal* vision of any human, but not the peripheral vision. Thus, I want to vary the PFOV of the large display and use it as an independent variable in one of the experiment conditions in a study, i.e., change the field of view between experimental conditions, while the user remains in a fixed position. With this approach, I will be able to investigate the effects of the PFOV by comparing a larger FOV to smaller ones. To achieve this, I will be using 3, 5, and 7 screens in the display for displaying information in three cases (which will be providing approximately 56°, 94° and 131° FOV respectively), i.e., only the available display size will change, other attributes such as size of the information will be kept constant. In other words, the available display space will change with the number of screens, but everything else will stay the say.

3.1.2.5.6 Resolution and Pixel Density

Many researchers suggested that we need higher-resolution displays to better view scientific images that are becoming available with the higher-resolution sensors [72,117,118,119,121]. However, this trend is not limited to images alone. With the increasing availability of Big Data, any display system needs to display more information. In terms of Furnas and Bederson’s space-scale concept, a display with more pixels can show a greater amount of data space (more overview) or a greater depth in scale (more detail) [49].

V4-SPACE is a 58-megapixel display space with a 52 PPI pixel density. To make observations about effects of resolution on visual analytics task performances, I will also simulate a system with a lower resolution configuration by aggregating pixels in some experimental conditions. This will be explained more in detail in further sections.

3.1.2.5.7 Control Distance and Visual Acuity

As one approaches a fixed-resolution display's surface, individual pixels become increasingly discernible. Pixels per degree (PPD) is often taken as a measure of the spatial quality of a visualization. The visual acuity metric [76] correlates directly to the more commonly used *Snellen fraction*, which is used to quantify this aspect of vision. The fraction 20/X corresponds to "this person can see at 20 feet what a person of average vision can see at X feet." Based on this definition, 20/20 is regarded as normal/average vision, which corresponds to a visual acuity of being able to resolve details of 1 arc min, i.e., 60 pixels per degree. The term Retina Display is sometimes seen as a measure of the quality of a display. This term is a marketing term developed by Apple to refer to devices and monitors that have a high enough PPI so that a person is unable to discern individual pixels at a normal viewing distance [103], i.e., roughly half of one's arm's length for a smartphone.

A highly debated issue with LHRDs centers around visual acuity [93]. The question is if pixels are "wasted" in the periphery on large displays [113]. There is some research that supports the idea that there may be performance costs if visualizations are scaled up for larger displays, which could be avoided by drawing only what the user can perceive. The most relevant research deals with the eccentricity effect, which shows that performance gradually degrades as a target gets farther from our point of visual fixation and that the extent of this effect increases with larger display sizes [24]. Still, this argument assumes that the user does not move their eyes or head, which is essentially never true in practice.

Still, the visual acuity measure described above applies only in the center of the field of view of a single eye (the fovea) and drops off substantially towards the periphery. This is easily demonstrated as one cannot easily read text one does not directly focus on. We need to use the term "pixels-per-degree" from the user's eye at the viewing distance, to further discuss the perceived "image quality" displayed on the system.

In my case, the user sits 3.3 m (130") away from the display system. As my calculations [116] show, V4-SPACE becomes a *Retina Display* at 1.67 m (66"). Since another independent variable to track in this thesis is resolution of the displays, we will have three conditions (for each monitor):

- standard resolution (2160p, or widely known as 4K)
- half the standard resolution (1080p)
- one third of the standard resolution (720p)

For latter two cases, the required distance for matching the retinal criteria would double and quadruple to 133" and 199", respectively. The diagram in Figure 3.5. explains better.

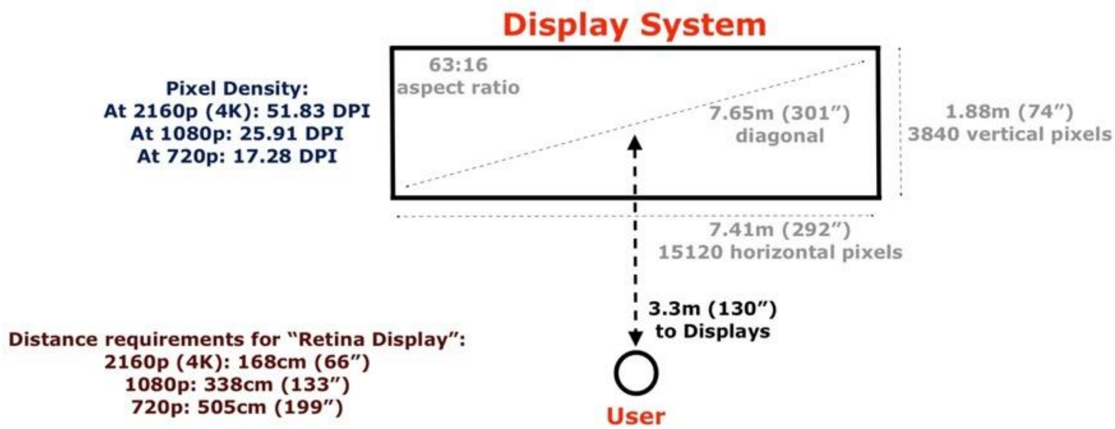


Figure 3.5. Diagram to explain Retina Display requirements considering the distances and the specifications of V4-SPACE displays.

The initial distance of the user to the display system at V4-SPACE was determined in a way that the middle case is very close to being "retina display". This gives us the chance to ask following questions:

- Does resolution play a role in performance if pixels become discernible? (lowering the resolution)
- Does resolution play a role in performance, even though the "retina display" condition is maintained? Is resolution still a performance factor

after reaching a certain PPD value? (increasing the resolution to 4K, even though the system already qualifies as *retina* at 1080p)

3.1.2.6 System that Powers the Displays

Since large display systems require much processing power, in the past clusters of computers have typically been used to operate LHRDs. But processor clusters are difficult to design, implement and manage - are they still needed?

In 2006, a 50-60 mega-pixel Powerwall display required a cluster of seven machines to drive it. In 2012 the same display could be driven by a single machine with three graphics cards, and in 2015 it could be driven by a single graphics card [86] (albeit sometimes at reduced frame rates). This demonstrates that computer systems are in general keeping up with display technology improvements, which motivated us to try to control all displays with a single machine.

A drawback of cluster-based display systems is that they typically require complex custom software to display data visualizations across the whole display area. Clusters usually require middleware software, which add an extra layer of complexity for building and operating large display systems. For example, the computer cluster that drives HiReD [86] uses middleware to eliminate problems of layouts transferred from conventional monitors. Thanks to the hardware system and the design of the visual analytics software that we are using on V4-SPACE, we do not require middleware to perform tasks in the intended way.

Our system consists of a single computer, which contains two nVidia Quadro M5000s [120] for the TVs in the display system and one nVidia Quadro K620 [115] for the smaller auxiliary monitor outside of the mosaic display. The system does not have visible and distracting delays, provides a consistent 60Hz refresh rate even under heavy, full screen GUI rendering, and has a 75 ms end to end latency.

3.2. Considerations for Designing the VA Tool: DynSpace+

In this work, I use *DynSpace+* as a VA tool on V4-SPACE. This software is an extension and a redesign of DynSpace (Figures 3.6 and 3.7), which is a VA tool for conventional desktop systems developed in the VVISE Lab. The plus sign “+” appended to the system name represents a modification for usage on large displays. DynSpace+ also differs in some aspects from the original, in terms of interface design, functionality and display preferences. This was motivated by previous research [3] that indicates that larger display systems have different sets of requirements, features and use cases, and therefore should use different interface paradigms.

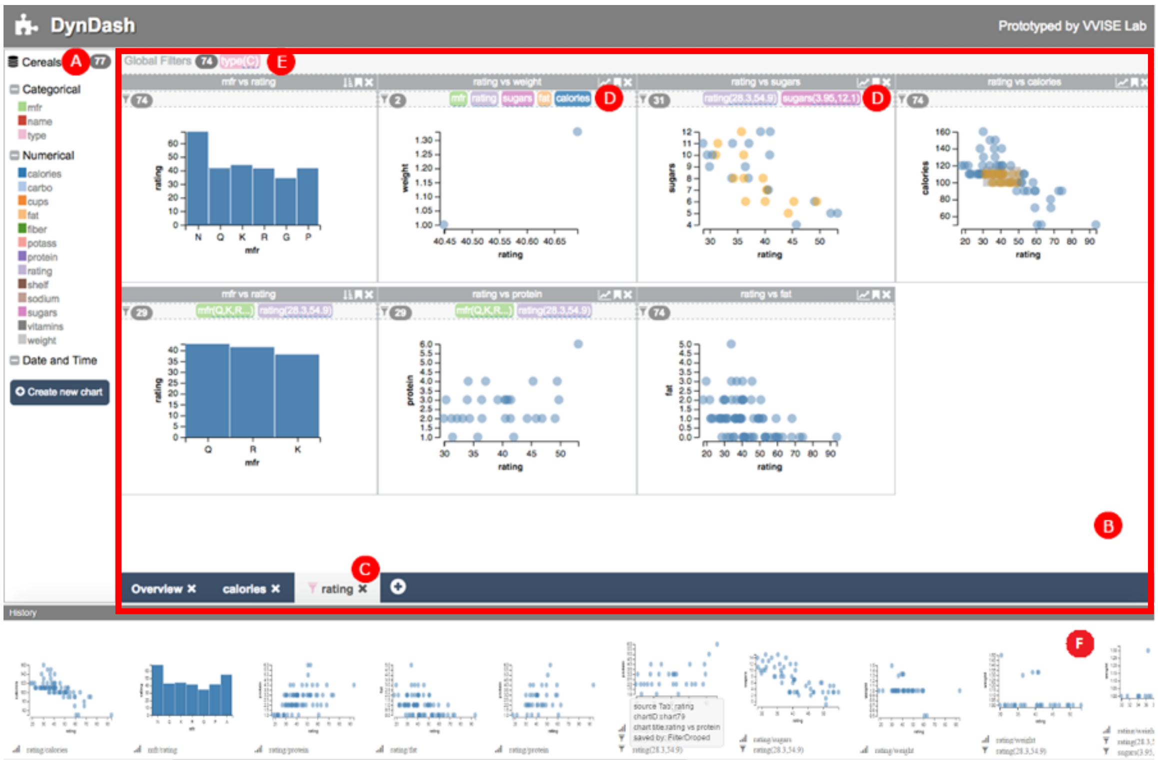


Figure 3.6. Desktop-version DynDash user interface.
The prototype contains three main panels, the Data Attributes Panel (A), the Dashboards Panel (B) and the History Panel (F). The Data Attributes Panel (A) displays the dataset name, total number of data records, and all the data attributes. Clicking on a data attribute name will highlight all charts that show that attribute. The Dashboards Panel (B) consists of multiple analytical dashboards, represented as tabs (C). Each dashboard can contain an arbitrary number of charts and all the charts inside a dashboard are linked (*charts* in DynSpace+ are explained later in this section). The Local Filter Bar (D) at the top of each chart shows all applied filters color-coded corresponding to the data attribute. These filters can be created, copied, or removed through simple drag-and-drop operations. Filters in the Global Filter Bar (E) at the top each dashboard apply to all charts inside that dashboard. The History Panel (F) consists of thumbnails of charts with their states maintained and tooltips showing details.

Figure 3.7. Original DynSpace in its environment

My thesis work took entirely place on large displays. In two separate studies conducted as parts of this thesis, two different versions of DynSpace+ were used. In the first study on exploring users' spatialization and information clustering approaches on large displays, a simplified version of DynSpace+ was used since the purpose was to emphasize semantic interaction and to remove functionalities that were not applicable so that they did not distract users' classification tasks. In the second study, all available analytical functionalities were offered to the user for supporting the sensemaking tasks that they were the focus of my second study. The views of those tool versions during two separate studies are shown below in Figure 3.8.

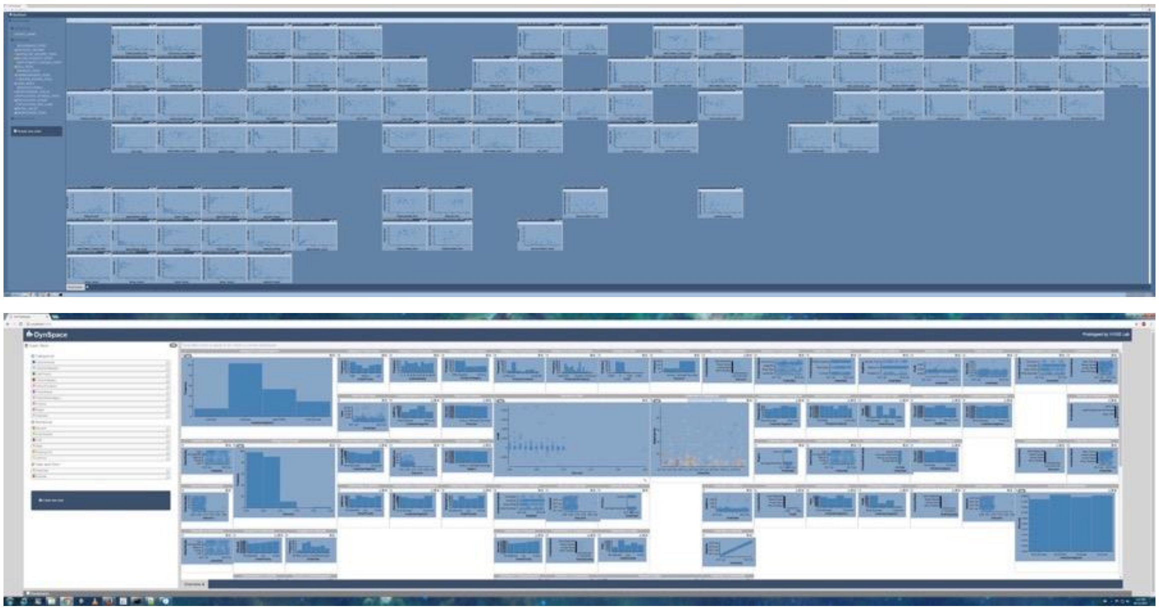


Figure 3.8. Simplified version of DynSpace+ used in Study I to observe content spatialization and how users cluster data charts (The upper image). The version of DynSpace+ used in Study II, with enriched analysis features (The lower image).

DynSpace+ is a browser-based VA tool developed with JavaScript. On the left side of the UI, is a *Data Panel*. The dataset's data dimensions are in this panel. Those data dimensions are classified as categorical, numerical, or date & time. Users can pick data dimensions, drag and drop them into charts in *Visualization Panel* for analysis.

The visualization Panel is the main panel containing data charts in which relations between selected data dimensions are explored. The visualization panel can also be considered as a *dashboard*. A dashboard is a "faceted analytical display" [46]: a set of interactive charts on a single screen, each of which presents a somewhat different view of a common dataset. DynSpace+ allows users to spatialize their data charts within the visualization panel, i.e. they can move, resize and easily compare multiple charts by moving them adjacent to one another.

A chart in DynSpace+ is a 2D representation of selected dimensions of a given data. Supported chart types include are scatterplots, histograms, bar charts and row charts (a type of bar chart in which there are horizontal bars instead of vertical ones). Charts are in rectangular panels, which are called chart containers. User can move, resize, add or remove charts through chart containers on DynSpace+.

- Moving: DynSpace+ has a floating layout, which enables the user to place charts (actually chart containers) in any part of the visualization panel. For moving the chart, users click on the top bar of the container, drag it and can drop it anywhere. DynSpace+ uses a grid layout manager [124] and a drop action will snap the moved chart container to the nearest grid position. This design decision was made to avoid overlapping charts and to guaranteed that all charts are visible. When there is a collision, existing chart containers are pushed down to make room for the moved chart in an animated sequence. For such collisions, only overlapping charts are pushed away.

- Resizing: Users can resize the charts in both horizontal and vertical dimensions. Resizing does not have to be uniform in both dimensions. To perform resizing on DynSpace+, the user clicks on the bottom right corner of the chart container, holds the mouse button and releases at the intended size. X and Y axes can be stretched independently, and the chart container adapts to the nearest grid location after the mouse release (while also making space to avoid collisions, as discussed above).

- Removing: This action only requires a click on the close button on the top right corner of the container.

- Adding new chart: For adding a new chart, the user clicks on the "Create New Chart" button in the Data Panel. This creates an empty chart container in the Visualization Panel. The user then needs to drag and drop data dimensions on X and Y axes to specify the data to be used. Depending on the data type, DynSpace+ automatically creates an appropriate scatter plot, histogram, bar chart or row chart.

Users can filter the charts. One way of filtering is by selecting data points on a chart and clicking on the "filter" button, which will create a new chart that involves only the selected data points. This is an effective way of focusing on details, separating parts of data, and/or removing outliers. Another way of filtering in DynSpace+ is to bring data dimension(s) onto selected chart to use the data dimension(s) as filters. Assume that a user is working with the “Bird Strike” dataset and is making a cost analysis of the bird strikes to aircrafts on a chart. For this, they can bring in other (related) dimensions and use them as filters by dropping them into a filter bar. They can also edit those filters by manipulating the data points or determining a range of values in those intended filters. For example, the user can specify “Originating Airport”, which is a data dimension, as a filter and can select only some of the airports from the corresponding dropdown list. Alternatively, the user can select the dimension “Time of the strike” and exclude “night” or “dusk” from the data points in that dimension. If the data is numerical in the selected dimension, users can specify a range. E.g., if they are using “speed at strike” as a filter, they can edit the filter and include only the ones that are higher than 20 knots and lower than 100 knots.

There are two types of filters: Global and local filters. Global filters apply to all existing charts, and local filters apply to only selected charts. For global filtering, the user needs to drag a filter to the “global filters bar”. For local filtering, the user drags the filter and drops on the filter bar for the individual chart. Filter name appears on the second top bar on the chart container, right under the name of the chart. All charts in the dashboard are coordinated through brushing and linking as well as global filtering. There are also various analysis functionalities, and study-specific features, which are explained in following chapters.

The intended use case for DynSpace+ is to analyze a complex dataset, which includes categorical and numerical data types. The user’s goal would be to gain insights, or answer questions regarding the details of data. Sample questions could be: “Do critics and ordinary users have usually similar opinions towards a video game?” (videogame dataset), “How high are the shipping costs with a delivery truck as compared to air shipping?” (superstore dataset), “Among the states in which females are less than 50% of the population, which state has the highest percentage of ESLs (people with English as a Second Language) in the population?” (US Census dataset).

To provide a starting point, *DynSpace+* displays a number of charts, generated automatically by picking random pairs of dimensions from the dataset input at the beginning of the analysis. The specific number of charts was adapted depending on the needs of each study, as explained in the respective sections. The main idea was that the automatically generated charts should take up approximately half of the analysis display space for the condition that uses the whole width of V4-SPACE.

As seen in Figure 3.9, users were provided in the beginning with a set of charts to help them to begin their exploration of the dataset. More specifically, DynSpace+ generates an regular array of charts with randomly chosen pairs of dimension, without any form of clustering (Figure 3.9.). Below this initial array, we left enough free space to enable the user to arrange the spatial layout of the content, e.g., by moving charts around, and encouraged users to create clusters. As the initially provided charts are only a subset of all possible charts, users were shown how to create additional charts, as needed.

Figure 3.9. The initial views of DynSpace+ versions in Study I and II, respectively.

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 [52,54,59,87,95]. However, as Schreck et al. summarizes [95], the unsupervised nature of the clustering algorithm can be disadvantageous for certain applications. Depending on the initialization and data characteristics, cluster maps (cluster layouts) may emerge that do not comply with user preferences, expectations or the application context [95]. Or, worse yet, bias the exploration of the user in inappropriate ways. As the generated clustering in our research should conform to the user’s own mental mapping of the information, we intentionally provided *no* initial clustering, keeping the clustering process entirely manual and up to the user.

In summary, the initial organization of the randomly chosen charts is completely non-clustered. Charts are simply placed next to each other. The users are asked to cluster their charts freely. As the experimenter, I did not assist them in this aspect, since I did not want to intervene in their thought process and thus potentially influence their own clustering strategy.

3.3. Data Sets for the VA Evaluation on V4-SPACE

Picking appropriate data sets for an evaluation is not an arbitrary choice and the quality and quantity of the data need to be carefully considered. Of course, if the amount of data is too small, analysis becomes trivial. At the other end of the scale, too much data may overwhelm participants, unless they are (some of the few people who are) used to deal with huge amounts of data. Enrico Bertini discusses *when data is too much* [15].

In general, (and assuming a constant pixel pitch, i.e., constant PPI) more pixels usually provide more space for visualizing data; but the number of pixels also imposes a limit to how much information can be displayed. However, this is not the only factor and we cannot necessarily expect a linear increase in displayable information with a linear increase in pixels, e.g., due to the limits of the human visual system. Clearly, as displays get smaller, the amount of information that can be displayed decreases too, which then increases cognitive loads, as the user has to remember information that cannot be shown. We can increase the display size to increase the number of pixels, but that will reach limits in terms of the field of view for any given eye and viewer position. For large displays, navigation costs also generally increase with an increase of size. There are further issues, such as the type(s) of navigation and limits to the amount of data one can cognitively manage. Thus, in general, for every data visualization scenario, there are lower and upper limits to the quantity of data – and information - that can be managed.

Potential consequences of working with very large datasets are cluttered displays, performance drop, information loss and limited cognition [15]. To make sure that the datasets in the evaluation were reasonable to use on V4-SPACE, I tested several datasets in DynSpace+, the main one being the 2016 US Census Data set [130]. Other datasets used within this work were an aircraft-bird strike dataset [123], a videogame sales dataset [127], and a superstore dataset [131]. All these datasets are too large to be displayed in their entirety on V4-SPACE, but (at least parts of) the datasets are still “small” enough to be comprehensible by university students within a time frame appropriate for a user study. Specifications of the datasets that were used in measured analysis sessions are given in Table 3.3.

Table 3.3. Following datasets have been used on V4-SPACE for VA evaluation.

Data Set	# of Data Records	# of Data Dimensions	Categorical Data	Numerical Data	Date & Time Type Data
US Census Data	51	18	1	17	0
Videogame Sales	200	16	6	9	1
Superstore	200	18	10	6	2

DynSpace+ supports histograms, bar charts, row charts and scatter plots in 2D (Figure 3.10).

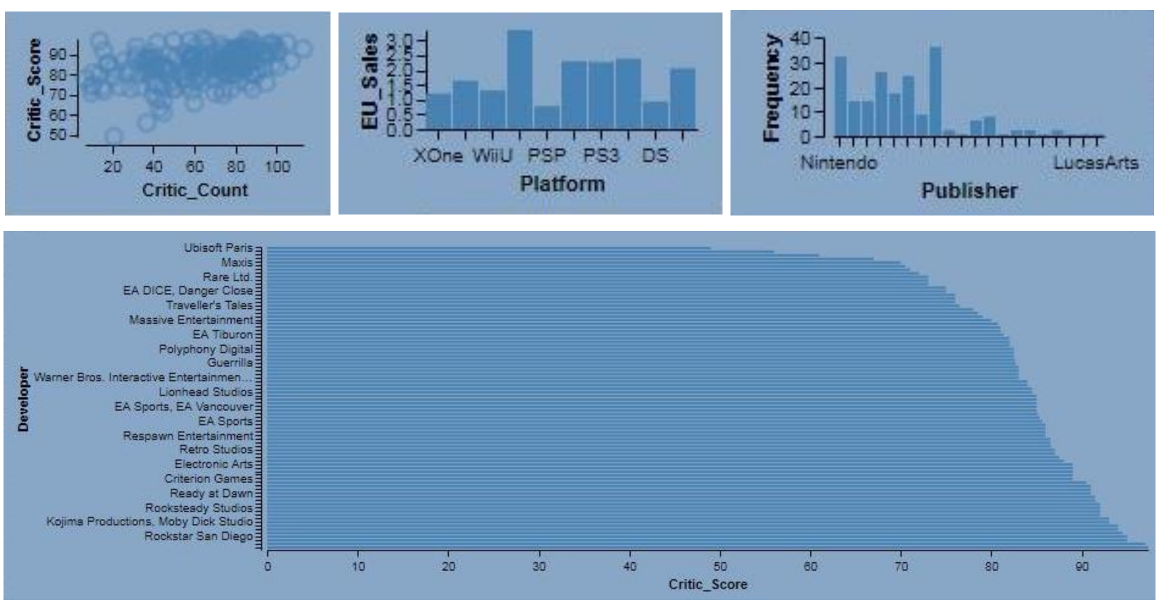


Figure 3.10. DynSpace+ charts: Top left is a scatter plot, automatically created when a chart is created using two numerical data dimensions. Top middle is a bar chart. Top right is a histogram. Bottom is a row chart (horizontal bars).

Daniel Keim created the VA mantra [56], as a variant of Ben Shneiderman's visual information-seeking mantra [96], which suggests a core process as follows: Analyze First, Show the Important, Zoom, Filter and Analyze Further, Details-on-Demand. In alignment with this, I simplified the data first (explained in the following study chapters). Initially, users are given an overview consisting of an initial random set of charts, but also as the opportunity 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 having users being distracted 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".

3.4. Design Overview

In summary, my experimental system builds on and extends previous research through careful consideration of the impact of display size, resolution, field of view, distance to display, and spatialization on VA task success while using LHRDs. Building on the findings of previous work for various design alternatives guided me in my design choices towards optimizing single user performance for VA tasks on a large high-resolution display system.

I conducted two studies to observe users' spatialization and clustering mechanisms, and the effect of display size and resolution on visual analytics task performance. During those two studies, I used two versions of DynSpace+, a VA tool developed in the VVISE lab and which was redesigned for the LHRD system. To facilitate the desired analytical reasoning process, I used several sufficiently large and complex data sets for my studies as briefly described above. A user on V4-SPACE running the VA study using DynSpace+ can be seen in Figure 3.11.

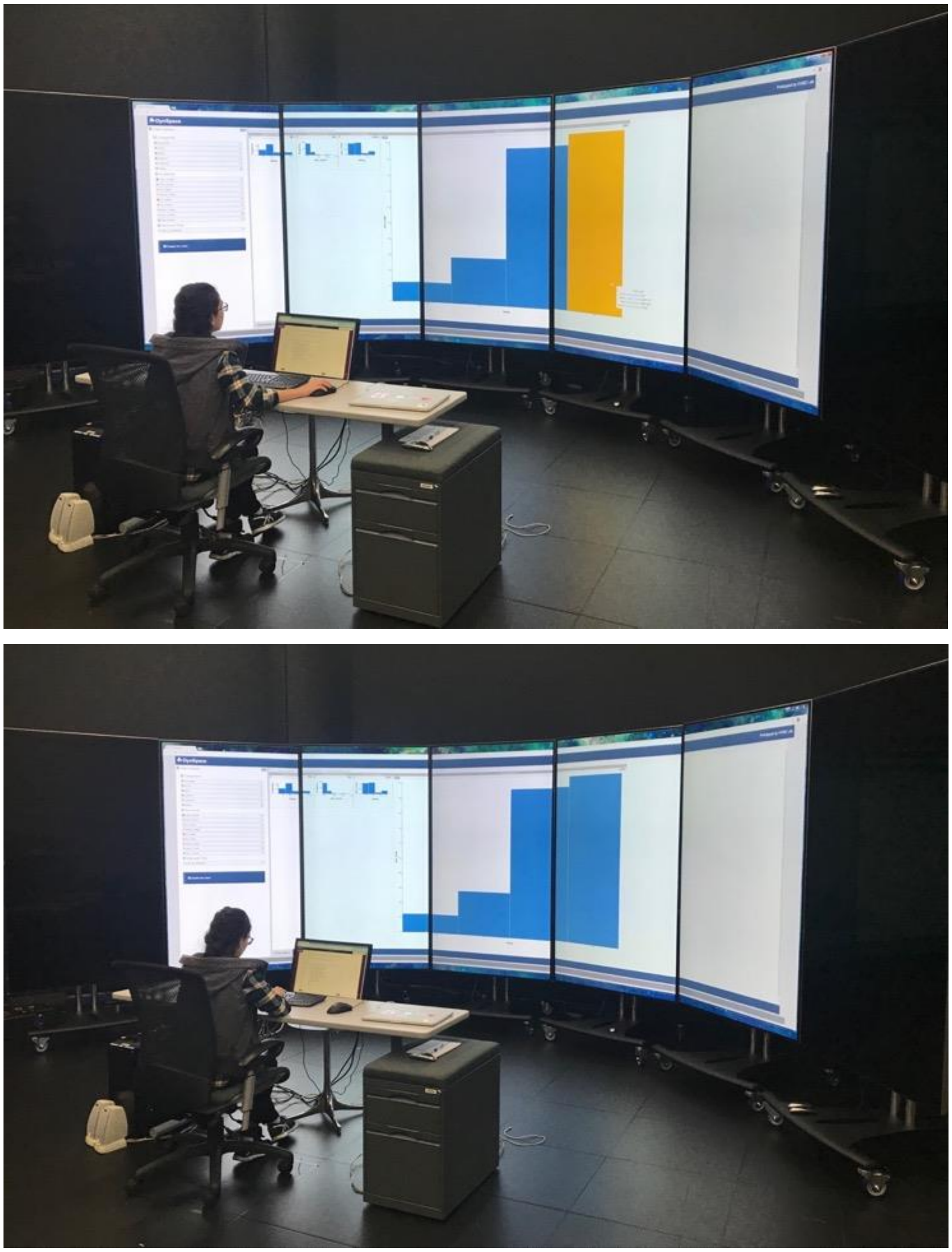


Figure 3.11. A user taking part in Study II, operating DynSpace+ on a specific size and resolution setting of V4-SPACE (in this case 5 out of 7 displays). In the top figure, the user is interacting with the data charts. In the bottom figure, the user is typing answers to study tasks using the keyboard after having fond some results in their analysis. The tasks are displayed on the auxiliary display shown in the figures, but the latter does not interfere with the visibility of the large display system.

Chapter 4.

Study I – A Classification Task

This first study addresses RQ1: How do people use the (large) space of an LHRD to cluster their items while trying to solve analytical problems that require re-organization of the content? What are the factors that users consider when clustering; topical relations, visual similarity, common dimensions of charts? 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? Do those patterns lead users to different types of navigation on the system? 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 behavior solving a classification task. The spatialization task used a simplified version of DynSpace+ to display many 2D plots visualizing relations among a subset of the 2016 US Census Dataset. This study took place on the largest possible V4-SPACE configuration, a 7-display mosaic in a single row.

Users were given a workspace populated with charts as shown in the top view of Figure 3.9. In this study, only scatterplots were used. While DynSpace+ can support multiple data chart types, my objective was to keep it simple and to focus on a classification task for study I. Thus, I used only one type of data (numerical) and one type of charts (scatterplots). During pilot studies, I observed that novice users' classification behaviours could be based on the type of charts, which has no relation with the visualized data. By using the same chart type, I wanted to make sure that classification decisions were not affected by such factors, as all charts were of a uniform type.

On the given workspace, users could create new charts, delete existing charts, resize the charts as needed and move them around freely. When resizing, they did not need to maintain the initial aspect ratio, i.e., they could resize charts in the x- or y-dimension independently, or in both with a different proportion. One noteworthy feature of DynSpace+ is that when the user clicked on a data dimension on the data panel on the left, charts that were using that dimension in the visualization panel were highlighted, as in Figure 4.1.

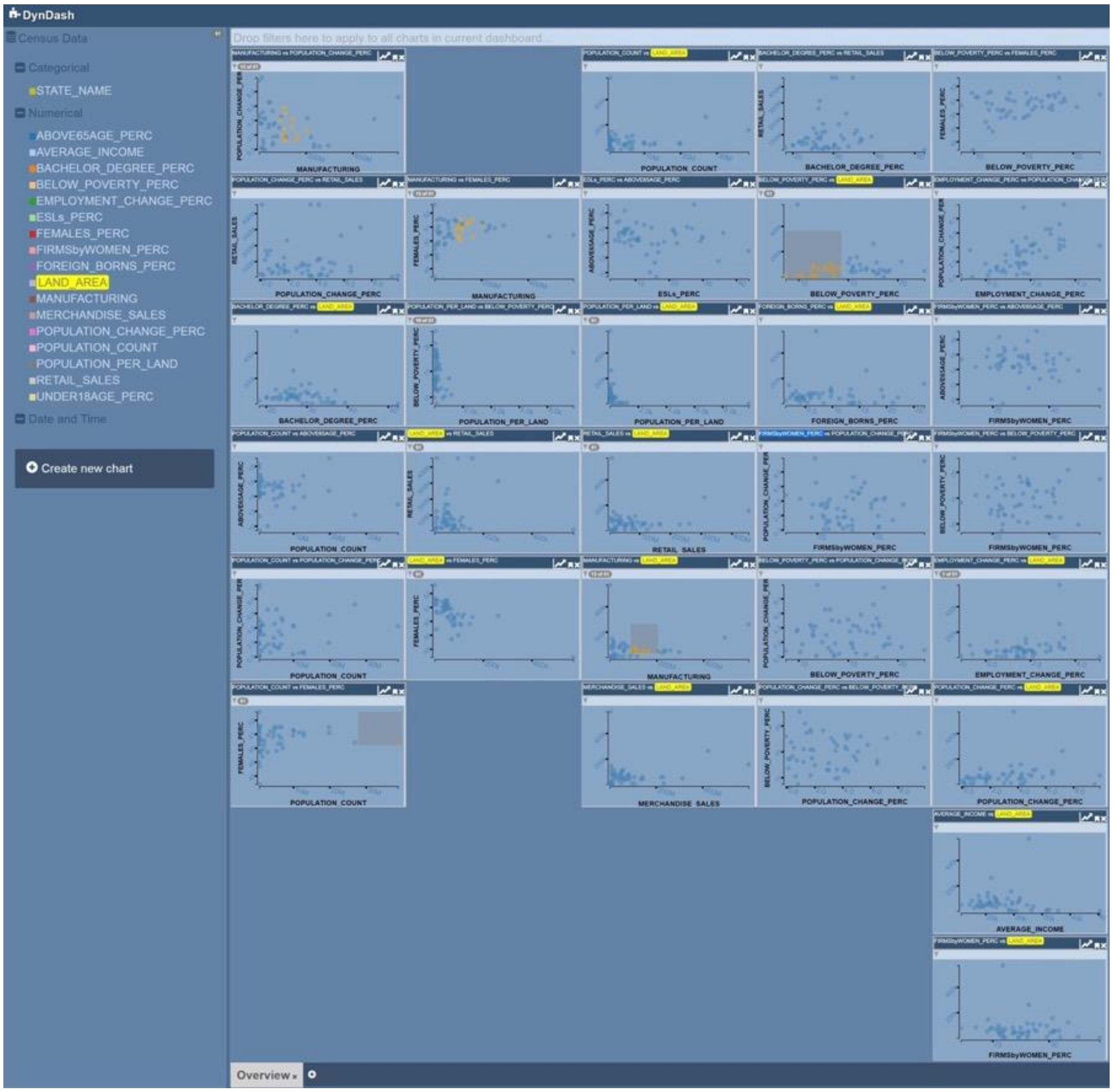


Figure 4.1. **Figure showing when a data dimension is clicked within the data panel on the left, all corresponding data dimension labels in the visualization panel are highlighted to identify where that data dimension is used.**

This study examined how users organized the 2D plots into clusters when given the instructions below. I put a lower limit to number of groups that they needed to create in the first task, namely that they needed to create at least 5 clusters. However, participants were free to create more than 5 clusters.

Task 1: Grouping

(a) Please arrange the charts into groups based on similarity. You can consider any similarity criteria. Group ALL the charts. The number of groups is up to you, as long as it is at least 5. You can have as many charts as you want in each group.

(b) Please explain the motivation for the way you grouped the charts in part (a). Try to give a detailed response about all the aspects you considered. Use the text box below.

Task 2: Highlighting

(a) Highlight each dimension in turn. Pay attention to any visual groupings.

(b) Identify those [dimensions] that do and do not correspond well to the grouping.

In Task 2, users could highlight dimension by clicking on each data dimension in turn on the data panel on the left side. When a user clicks on a data dimension, the tool shows the instances of that dimension as in Figure 4.1. The purpose of this task was to ask the users to highlight dimensions after they finish with grouping, and see if their grouping criteria matched a "common dimensions” factor, which means plots having the same dimension in one of their axes. The users needed to complete Task 1 before being asked to highlight. After they observed how highlighted charts were aligning with their own grouping, they were required to type in their observations to the given text field.

Endert et al.’s work [43] presents the concept of semantic interaction that seeks to enable analysts to spatially interact within their analytical workspace, which was the main inspiration for this study. Grouping the information content and generating clusters of objects during analytical tasks is a part of sense-making, which is the process of searching for a representation and encoding data in that representation to answer task-specific questions [88]. It takes place in the early stages of visual analytics processes, and is a part of preparation stage, which can impact task efficiency in the next stages, especially on large displays. For this reason, my objective is to study clustering on LHRDs as a part of this thesis in this first study.

4.1. Apparatus

The system consists of physical system, a software tool, and a dataset to be analyzed.

4.1.1. Physical System

The physical system consisted of the LHRDs, an auxiliary monitor, keyboard and mouse, and the computer that runs system, and additionally a comfortable swivel chair allowing only rotation and height adjustment. Translational motion was restricted to maintain a constant distance to the displays. For details, see section 3.1.1.1.

Study questions were answered by the user using the auxiliary monitor positioned so that it did not visually overlap with the analysis space on large displays. Having this auxiliary display enabled me to provide continuous visibility of the actual analytics material. The use of the auxiliary display is justified in section 3.1.2.1.

In this study, the distance of users to 1 x 7 array of large displays in a circular arc around the user, whose position was fixed at 3.3 m (130”) from the displays.

4.1.2. Visual Analytics Tool

I used a simplified version of DynSpace+ as explained (section 3.2) and illustrated (upper image in Figure 3.9) above.

A screenshot during a participant’s analysis is given in Figure 4.2.

Figure 4.2. A view from DynSpace+ during Study I. Left panel (1) shows dataset and its dimensions; analysis takes place in the main panel (2) where user arranges 2D data charts into clusters.

4.1.3. Dataset

In this study, I used a custom dataset generated from the 2016 US Census dataset [130], which can be found in the data division of the United States Census Bureau website (American Fact Finder). This Census dataset can be downloaded as a MS Excel file and contains over 50 columns (dimensions), related to population, economical and geographical facts. It also has over three thousand rows, one for each state and county of US.

My custom version is available at http://bit.ly/2zkupn7 as a JSON file, the file type that DynSpace+ works with. Usually, displaying tens of thousands of records simultaneously causes cluttering and dozens of data dimensions contribute to performance issues. Some strategies for dealing with visualization of large datasets are sampling, aggregation, tuning/tweaking and segmentation [15]. I was reluctant to sample data points, fearing information loss, as randomly filtering out data points might have caused the loss of useful data. Also, specific to my case, having some of the states in the dataset and not having others would not make so much sense and such a design decision would have been difficult to explain. Instead I aggregated the data. I simplified the data set to make it more appropriate for VA/LHRD studies, by aggregating county data at the state level.

I also limited the number of dimensions to 17. Following that, I shortened data dimension labels so that they would fit in the chart containers at a readable size. If, for example, the label was “The percent of persons that are over 25 years old, in between years of 2009 and 2013, who have Bachelor’s Degree or higher”, I made it “BACHELOR_DEGREE_PERC” (Due to a technical problem, I could not include a “%” sign or “ ” character).

The 17 dimensions of the dataset used in Study I can be seen in Figure 4.3.

Figure 4.3. Metadata for subset of 2016 US Census Data used in this study. At the top is the dataset name and the number of rows of the dataset (corresponding to states) in the top right corner. Data dimensions are given under the respective data types. The names of the states are categorical, while the 17 data dimensions are numerical.

4.2. Participants

There were nine participants, P0 to P8, four of whom were male, providing a balanced gender distribution. Ages ranged from 18 to over 40. Some participants were undergraduate students who participated for course credit, others had at least a bachelor’s degree and were individuals from my personal connections who volunteered to take part in the study. A table for detailed user responses for the motivation survey, which was a part of pre-study survey, is given in Appendix C.

There were seven questions in the list of terms and concepts that I gave participants during the pre-study survey to identify how familiar they were with the context. On average, participants were reportedly familiar with 4.4 of them. The summary of the concept familiarity data for this study is given in Appendix D.

The participants of this study were mostly novice users, since the study did not require VA skills other than basic skills that would assist them organize the content provided to them in a workspace. 5 participants out of 9 did not have any experience in visual analytics. One reported less than a year of experience and three reported 1 to 3 years of experience. From a software point of view, seven of them have never used any visualization or VA tools, whereas one had used D3.JS and another used R in a statistics course for a short period.

To ensure that the displayed information made sense to participants, in the pre-study survey I asked them whether they could interpret data from scatter plots. They ranked their likelihood 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 scale used.

I also made sure that participants had not participated in previous studies of our research group, since the visual analytics tool used here was the same as some other VA studies in our lab and I did not want familiarity bias to occur. Also, users were asked if they had perfect or corrected vision, to make sure that they could read the text on the displays from the specified distance.

4.3. Experimental Design

4.3.1. Procedure

To prevent procedural errors and to remove potential biases, I carefully designed the experimental protocol and documented the entire process in advance. The whole procedure is given in Table 4.1 below.

Table 4.1. The steps of the formal procedure followed in the study.

1. Introduction of the environment
2. Pre-study background survey on auxiliary monitor (discussed above)
3. Start DynSpace+ on V4-Space
4. Verbal introduction to study procedure
5. Questionnaire on auxiliary monitor given (user tasks, discussed above)

6. Start screen capturing on V4-Space
7. Main task 1: Clustering
8. Main task 2: Highlighting
9. Screen capturing ends
10. Post-study interview (audio recorded)

The 1st and 4th steps required giving verbal introduction to users, regarding the system or the tasks. To ensure that every participant got the same introduction and instructions, here is the sequence I followed: During step 1, I first informed participants about the length of the study and advised them to get comfortable in the swivel chair by adjusting their seating posture, as needed. I told them that they could rotate/swivel the chair to see the full displays, but that their chair had to remain in the spot until the end of the experiment, since I wanted them to retain the same distance to all displays in the system. Leaning back and forward was allowed, as needed. Then I gave them basic information about the display specifications and the system configuration, to give them a basic awareness about the system that they were interacting with. I observed that most participants found V4-SPACE very interesting, which helped to motivate them to perform the subsequent experimental tasks. Next, I trained them in basic yet frequent operations: keeping track of cursor, how to find the cursor when they lost track of it, and switching between LHRDs and the auxiliary monitor through moving the mouse cursor to the top right of V4-SPACE and back.

During the 4th step, I started an instance of DynSpace+ and introduced them to the VA tool, how it works, what they needed to do during the study and some practical tips regarding the usage. For each of the participants, my introduction was roughly as follows:

- This is a visual analytics tool. It allows analysis of data. I am doing studies on this process of analysis on large high-resolution displays.
- A dataset is loaded into the visual analytics tool. Data has various dimensions. You can see those dimensions in the left panel. The data we are using right now is census data of the United States from 2016 that gives various information about the population. [I give examples here, from random data dimensions on the screen.] We also have scatter plots in the main panel, where we can see and explore the relations between data dimensions. There are 51 points in each scatter plot and each point represents data of a different American state for the given data dimension. For example, if I am looking at the scatter plot visualizing the relation between AVG_INCOME and BACHELOR_DEGREE_PERC, each data point on the plot is showing where each state stands in terms of those two data dimensions. You can resize

charts, delete unneeded charts, or create new charts using data dimensions of your preference. [Here I demonstrated each activity.] If you click on any data dimension on the data panel, it will highlight all data charts in visualization panel that contain this data dimension.

- You will start your analysis with the given dataset using the VA tool. For this analysis, I am asking you to organize your workspace by grouping your data charts. The reason I am asking you to move chart containers around is that, for the analysis process, you will need to have several distinguishable groups of charts containing different kinds of information. By moving charts that are relevant to each other closer to each other, you need to group charts together on your workspace. You are asked to create a minimum of 5 groups, but there is no upper limit. It is up to you what you choose as your relevancy criterion for grouping. E.g., you can go by how the charts look, common data dimensions, or can be any kind of similarity or commonality that you can think of.

- Feel free to use the given charts if you find them useful. You can delete charts if you do not think that you will need them. You can add new charts as needed, if they are important for your analysis.

- There is no right or wrong answer for this task, the study is about your preferences and considerations. Charts are draggable. You can “grab” them by the darker top bar of the container and move them around. If you want to insert your chart into an occupied position through drag-and-drop, the tool will push the other charts away to provide you with a free spot.

- Ask me whenever you have questions or get confused and let me know when you are done.

I followed the steps of the procedure in the given order for consistency. I assured users that they could ask questions at any time during the experiment and that they should communicate around the tasks, their thoughts or any points where they experienced confusion. With pen and paper, I noted down any direct verbal feedback from the users during the tasks, as well as indirect physical or approach-based responses. I collected this data to be analyzed later.

4.3.2. Tasks

In this study, participants were asked to complete two tasks, consisting of two subtasks each. In the first task, participants grouped their scatter plots. They 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 the charts. The number of groups was up to the user as long as it was at least five (5), and they could have as many charts as they wanted in each group. After completion of the grouping, they were asked to explain their motivation for the way they grouped the charts. I recorded their detailed responses about all the considered aspects via the text box provided.

In the second task, they used the highlighting facility through clicking on the data dimension in the left panel, to see where these dimensions exist in charts in the main visualization panel, and to observe any patterns that might become apparent. They highlighted each dimension in turn, and observed if the highlighted charts were mostly together in their grouping or not. They shared their observations about the highlighted dimensions in relation to their own grouping by typing their response in the study questionnaire on the auxiliary display.

4.3.3. Data Collected

Users of the study completed a pre-study survey after signing a consent form. The online survey responses were stored on Simon Fraser University's servers. During the tasks, users completed a study questionnaire. The answers were recorded through free-form text boxes and the responses were stored in the same survey system.

Users' clustering activity was tracked in two ways: First, I screen-recorded their activities. Since the study used a very large ultra-high-resolution displays, I could not record the activity in real time as a video. With more than 58 million pixels on the display, each screenshot takes up more than 3 MB in the default *JPG* format. If we tried to capture the screen at full resolution with a 60 FPS video, we would a disk system capable of a sustained write speed of up to 200 MB/s, and would save more than 11 GB of data every minute, which means hundreds of GBs for each participant and terabytes of disk space. Additional space savings, typically an order of magnitude less, can be expected due to the fact that the screen video does not change for every frame. While it is feasible to handle such volumes for recording, the real constraint is the compression effort. A quick test convinced me that the CPU load for the continuous compression during the screen recording would severely slow down the rest of the system, which would render the system too sluggish for normal usage.

Thus, I decided to use a program called "Chronolapse" which performs periodical automatic screen captures, where I captured a screenshot every 120 seconds. The motivation was that observing the grouping actions of the participants did not require high frame rate, and 120 seconds would still be frequent enough to record major changes that the user did on the graphical layout during the tasks. I additionally recorded the positions and information about each chart for each final participant's grouping. This data enables me to apply clustering algorithms on them after the fact.

At the end of the study, I did an audio-recorded post-study interview with each participant, which typically took less than 5 minutes. This interview was semi-structured, and I asked them a list of questions regarding the tasks they completed, the software tool, the system, what they liked, what was challenging and/or confusing, and whether they had any further feedback. The results of those interviews are discussed in section 4.6.

At each step of the procedure, I watched the user carefully and noted down important observations, to be able to better understand the raw screen recordings, e.g., through recordings of frustration at a specific point in time, which in turn further informs the analysis of the experimental results.

4.4. Results

4.4.1. Clustering

Participants were expected to organize charts by manually grouping the plots into clusters. Screenshots of the final output for each participant are given in Appendix G.

Number of Clusters

Users were instructed to construct at least five groups; yet there was not an upper limit specified for the maximum number of groups that they could have created. With these instructions, three participants created five groups, the minimum allowed number. Out of the remaining six, four clustered their plots into 6 groups, one into 8 and one into 16 groups.

Clustering Criteria

Participants used various criteria when they were clustering the scatter plots. My analysis of these criteria is based on my nodes from the observations of each participant during the experiments, their written answers to the study tasks, and my interpretation of their final group layouts (through the screen recordings) after the experiments.

I observed three main approaches, which can be classified as (i) visual appearance that reflected correlation of the plots, (ii) topical relations of the used data dimensions, or (iii) commonality in axis labels.

- Visual appearance: Some participants interpreted the scatter plots as “pictures” and simply grouped the ones that “looked similar” together.
- Topical relations: Some participants grouped charts with similar and/or related data dimensions together. E.g., they created clusters with women-related dimensions, or sale-related dimensions.
- Common axis labels: Some participants picked a specific data dimension and clustered all charts that involved that dimension. E.g., they created a cluster of LAND_AREA versus every other data dimension, regardless of whether that comparison makes sense or not.

Each of these approaches corresponds to a primary clustering criterion that determines which chart goes into which group, which implicitly separates clusters across the display. Some users used only a single criterion for grouping and did not consider anything else. Other participants used a secondary criterion to determine either the positioning of charts within a cluster or to determine the relative positions and distances of clusters. This includes applying one criterion for constructing the clusters and applying another criterion for in-cluster arrangements and depicting intra-cluster relations. For example, P6 stated that he grouped charts that had common x-axis labels, as the primary clustering criterion. However, P6 deliberately chose the positions of charts within each cluster, by considering “*how they visually appear*”.

Three participants based their decisions on how the plots visually appeared (Figure 4.4). They grouped the plots so that those that were showing similar trends were clustered together. Among those three, P1 mostly relied on a higher level, pictorial observation of the plots. P1 said; *“… Most dots in Group C were staying in middle of chart. In Group D, most of dots are [at the left side] of the chart*”.

P2 on the other hand, considered whether there were any patterns in the data or not. She responded to the questionnaire as follows: "*Some data of [the charts] is very concentrated. People may find [rules/correlations] from the [charts]. But some data of [the charts] is very dispersed.*" It is interesting to observe that she clustered strong positive and negative correlations together, since she was focused on how usable the charts were, in terms of reflecting strong trends in the data.

P7, the other participant with the same approach, exhibited a behavior similar to P1, but seemed to be more aware of the concept of correlations and did not just interpret the plots as pictures. P7 created groups with specific kind of correlations, and an additional group with no visible correlations. Interestingly though, P7 started the task by topical clustering: "*When starting to organize the graphs, I first looked at the topics that were being compared.*" However, P7 appeared to be a bit uncomfortable with reading text on V4-SPACE, even though this participant did not report any vision problems and was wearing glasses at the time of the experiment. At the beginning, P7 leaned forward to be able to better read the text. However, it seems that the readability problems remained. Thus, P7 changed their clustering approach to avoid any problems/discomfort when reading headings and labels.

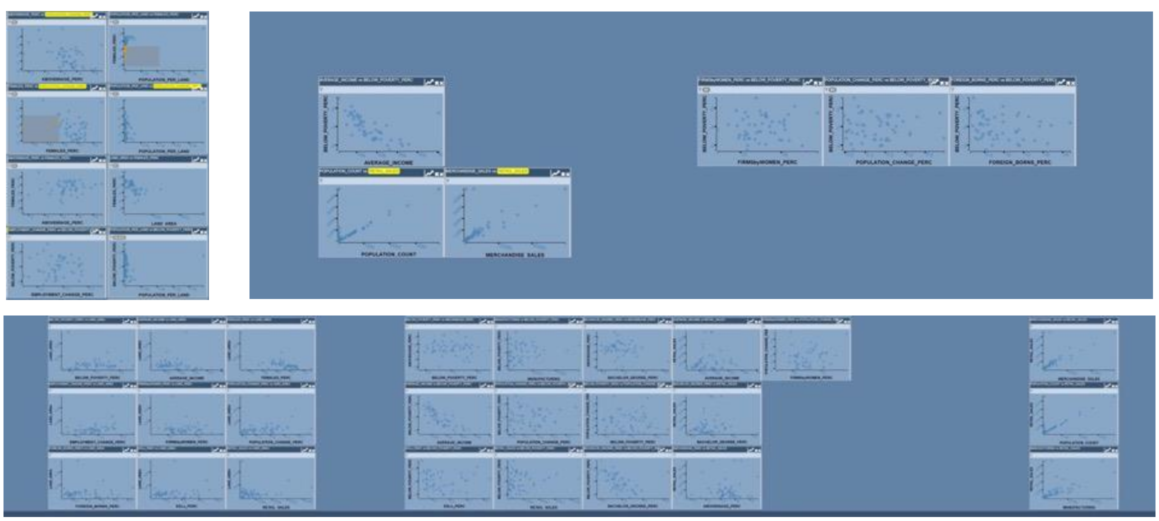


Figure 4.4. Excerpt images from P1, P2 and P7's grouping. All 3 participants based their grouping on the visual appearance of charts. Top left is from P1, top right is from P2, bottom figure is from P7.

Above examples of the different approaches for grouping are shown. P1, where an excerpt is shown at the top left of Figure 4.4, stated that this was a vertical grouping, although the plots are placed next to each other. All entries in the column are visually similar for this participant. A part of P2’s chart layout is shown in the top right of Figure 4.4. Here the grouping criterion is if there is (any) correlation or not. On the left side, negatively and positively correlated scatter plots are grouped together. On the right, charts have no correlation. Finally, P7’s groups are at the bottom of Figure 4.4. In those 3 clusters, the participant grouped three kinds of scatter plots: On the left, data points are spread mostly along the X-axis. On the right, there are charts with a roughly linear positive correlation. In the middle, there is no visible pattern.

Three of the remaining participants clustered scatter plots with the strategy of picking one common dimension, either x or y. Interestingly, all three participants who followed this strategy chose the vertical, i.e., the Y-axis as the common data dimension, and grouped all scatter plots that have the same dimension together (Figure 4.5).

P0 explained this behavior as follows: “*I chose [vertical axis] as my choice instead of [the horizontal one], because visually [the vertical axis] stands out.*” P4 explained *“… this way will be easier to read it*”, and P5 did not provide any feedback. P4’s approach in particular was unusual, as this participant decided that there were too many plots. Thus, P4 picked five key data dimensions, created groups around them, and deleted plots that she thought she would not need. It is noteworthy to state that all three participants who adopted this strategy as their primary clustering approach seemed to agree on the selection of dimension for common properties. However, this approach did not align with P6’s comments, who adopted a different approach but made use of labels on axes as well, where he commented that labels on X-axis were more readable.

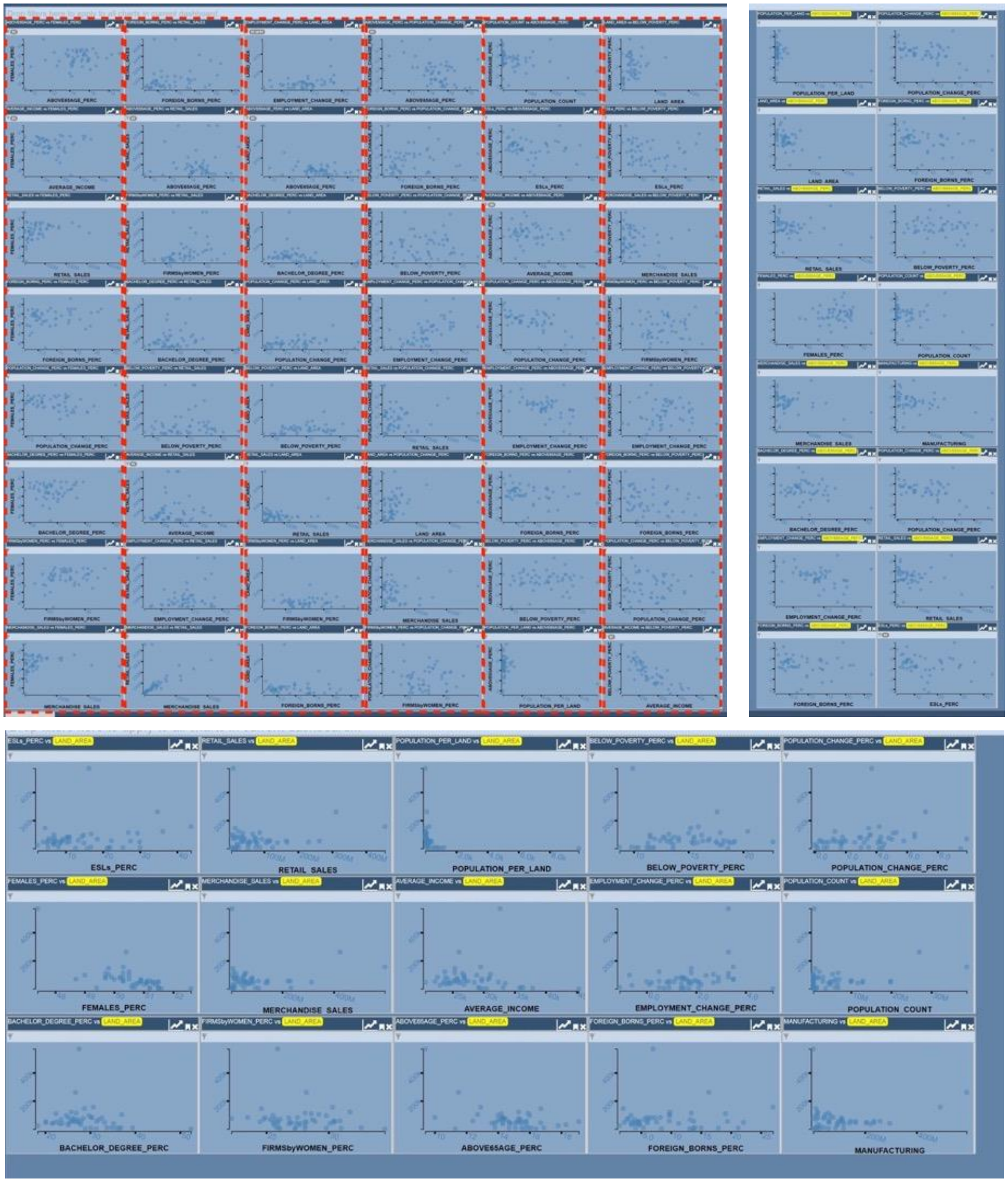


Figure 4.5. Excerpts from P0, P4 and P5's organization are shown above. All 3 participants based their grouping on common axes (the Y-axis). Top left is from P0 (6 groups indicated by red dashed line), top right is from P4, bottom is from P5.

As shown in the top left of Figure 4.5, P0 had all the groups positioned side by side, with no gaps in between. I added the red dashed rectangles in the image to make the groups more distinguishable. Each of the six column (groups) has a different, common data dimension. As visible in Figure 4.5, P4 (top right) and P5 (bottom) followed the same approach. Yellow highlights (applied by the users) visualize the common dimension for the group (ABOVE65AGE_PERC for P4 and LAND_AREA for P5).

The remaining three participants used a topical selection strategy when they were constructing their clusters, i.e., they grouped the same topics together. Different from the previously discussed participants, they did not restrict the clusters to only a single dimension; instead they put charts with a few related data dimensions into the same cluster. Part of P3's organization exemplifies this approach, as shown in Figure 4.6.

P3 said "*I tried to organize the charts using their axis names. Some of them were about average, some about population and some about females (women). I did not care about the data inside them. I just grouped them by the [labels] I would see, mostly horizontal but sometimes vertical ones.*" In this participant's final layout, the groups consisted of charts with related axis labels, either the horizontal or vertical one, and an extra group for uncategorized plots. P3 said that they did not quite understand what those plots were representing, which is why they kept them in a separate group.

When clustering the plots, P8 considered topical relations as well and put plots that represent related data dimensions, such as different sales parameters and manufacturing stats together. Also, P8 created another cluster with "FEMALES_PERC" and "FIRMSbyWOMEN_PERC" grouped together and placed it next to sales and manufacturing data.

P6 spent a lot of effort on the task and gave a detailed response to the written questions regarding their grouping approach. This participant had various criteria while creating clusters, but prioritized the topical relations in the data.

> *I used horizontal axis texts to group charts. And in each group if dots are grouped [close], I stretched [the] graphs x or y axis, so I can see better and makes them easier to catch [with] the eyes. I created [a] virtual grid in my mind, which divides [the] screen [into] three parts horizontal. And I tried to list group elements in some random groups in different order. For example, all group elements [have a common] vertical or horizontal [axis]. Purpose of this is to break the order so I can focus. I tried to avoid screen [edges], they make it hard to read; but not a crucial problem. One last thing; I tried to collect groups to the right, because [the] left [panel] distracts me.*

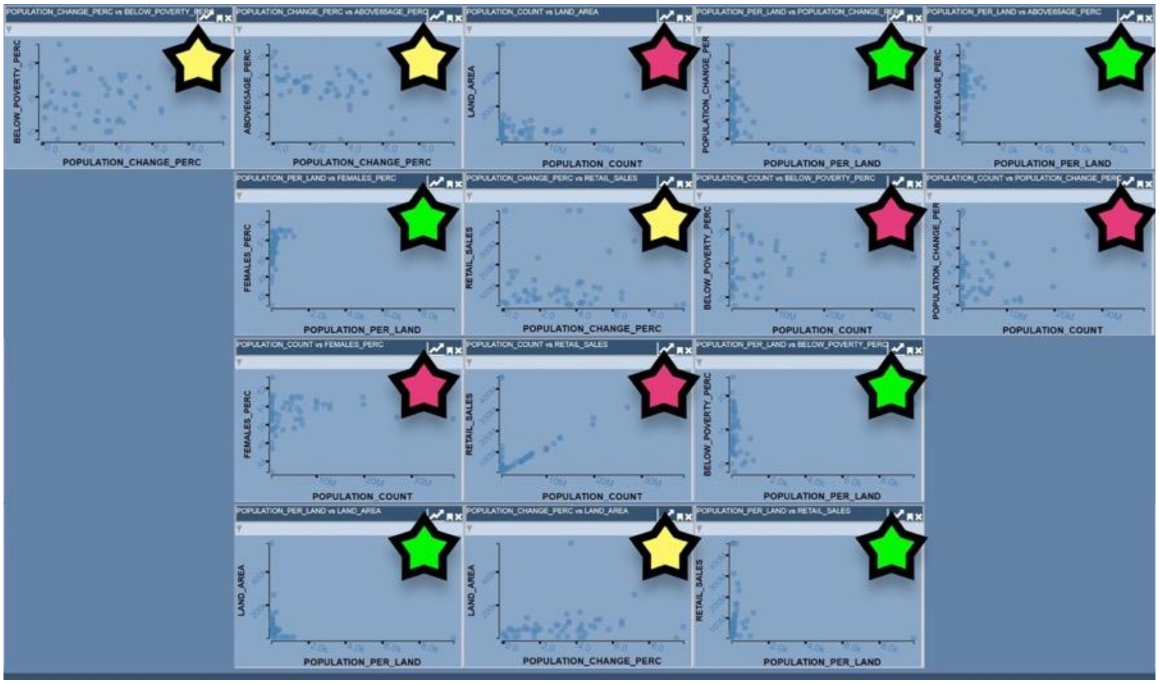


Figure 4.6. Above is a cluster that P3 created. All scatter plots have population-related data dimensions in their x-axes: POPULATION_COUNT: red star, POPULATION_CHANGE_PERC: yellow star POPULATION_PER_LAND: green star.

The full responses from participants around their motivation for the way they grouped their charts are given in Appendix E.

After the data dimension highlighting task, their responses around which highlighted axis labels do and do not correspond well to their grouping are given in Appendix F.

Cluster Shapes

The clusters that the participants created were mostly either linearly arranged in long, narrow clusters, or had a roughly round shape around an arbitrary center point. However, some variations occurred, which combined those two distinct characteristics. That means also that clusters had widely varying aspect ratios.

P0 and P1 generated similar arrangements, where the clusters were organized into vertical columns of charts. Such an arrangement is notable, as vertical “stacks” of items on V4-SPACE are an inefficient use of the whole display system, which is horizontally much wider (almost 4:1 aspect ratio). With DynSpace+, participants could quickly scroll up and down vertically via the mouse wheel to see other items in the cluster. P4’s approach was also somewhat similar, with the difference that P4’s clusters consisted of “double columns”, i.e., a two-charts wide tall group, which decreased the amount of vertical space needed, and thus did not require vertical mouse scrolling. Despite creating these vertical double columns, P3 also created a linear horizontal (!) group, which they used for uncategorized plots. Different from aforementioned participants, P3’s linear group was horizontal, and was covering all the horizontal space of the display system in a single row.

The remaining five participants did not create linear groups, and charts in their clusters were thus closer to each other. For P2 and P7, no visible patterns stood out, as their cluster shapes were mixed, following no defining shape. On the other hand, P6 and P8, created clusters that were approximately circular. Finally, P5’s approach used horizontal row-shaped groups of plots, which use the horizontal space of V4-SPACE efficiently. In other words, P5’s approach was complementary to the vertical arrangements of P0 and P1.

If I classify cluster shapes through their aspect ratio, where numbers larger than one signify horizontally-wide clusters I get the following results:

- Much higher than 1: P5.
- Much lower than 1: P0, P1, P4.
- Close to 1: P2, P3, P6, P7, P8.

Examples for each behavior are provided in Figure 4.7.

Figure 4.7. A vertically-dominated clustering with aspect ratio << 1 is shown in the top left (P4). Top right shows a roughly round arrangement with ~1 aspect ratio (P6). The bottom shows the horizontally-dominant clustering of P5 (aspect ratio >> 1).

Intra-Cluster Relations

I also analyzed the arrangements to identify any potential semantic patterns *between* as well as *within* clusters. By "relation between clusters", I mean if the position of the charts and distance between clusters have a discernible meaning in the layout.

Such relationships were only visible for a subset of the participants. For four participants, I could not observe any such relationship, nor found that they self-reported relations between different clusters. For those four, clusters appeared to be clearly separate, and were randomly positioned. P0 and P1 did not have any space between clusters, and they were arranged in a random order. For P4 and P5, clusters were spatialized with some distance between. However, the spacing seemed quite uniform and clusters looked no different than one another.

On the other hand, some participants used the arrangement to convey some meaning. For P3, cluster shape changed with what the clusters represented. While the main groups looked quite similar, their cluster of ungrouped plots looked different from others in terms of shape and the distance to other clusters (Figure 4.8).

Figure 4.8. For P3 categorized clusters were generally rounder with an aspect ratio closer to 1. However, the cluster of uncategorized plots (at the top, marked with red) had different visual characteristics from the others. The difference in state of the groups (categorized / uncategorized) also determined the cluster shapes.

For P2 and P7, who both clustered plots in terms of visual appearance, clusters of strongly correlated plots were placed far away from other clusters, likely to further emphasize their difference. Though, other clusters, whose charts were visually more similar, were positioned closer, as explained under Figure 4.9. These participants used distance to symbolize the “closeness” of the relations they used. In other words, if members of different groups were completely unrelated, groups used to be far away. If they had similarities with members of other groups, they appeared like #4 of Figure 4.9.

Figure 4.9. A snapshot from P2’s grouping. Among the numbered clusters, #4 is a (meta-)group of clusters that are close to and transitioning into each other, due to the relations between the charts in them. Clusters #1, #2 and #3 are distinct and they are clearly separated. In #1, there are plots in which data points are dense around an axis. In #2, there is either a positive or negative correlation between data points. #3 consists of data points which are scattered around the chart.

Finally, for P6 and P8, the same effect was seen with the topical relations. Most groups were located depending on what information they revealed. Examples include "RETAIL_SALES" and "MERCHANDISE_SALES" being on top of each other and population-related plot clusters being side by side. While these were still different clusters, they were closer in terms of topics.

Distance between Clusters

Among the participants, I observed four different patterns around positioning and distance between clusters. P0 and P1 did not put any distance between their vertical columns of charts. These columns were right next to each other, which made it hard to distinguish the groups from each other in the beginning, as in the top left screenshot in Figure 4.5. For P2 and P7, the distances between clusters varied by the strength of relation that existed between them (Figure 4.9). Here, distance represented the strength of the semantic relations between clusters.

I configured DynSpace+ so that the canvas panel permits the user to organize all charts into a fixed, rectangular grid. Each rectangle can contain a single chart with its default (minimum) size. The reason behind this design decision was to prevent the user from overlapping charts, which ensures that all content is visible. P3, P5 and P8's grouping strategies reveal that they wanted to have some distance between clusters, but that they also tried to keep that distance minimal: the majority of their clusters were separated by single-chart-sized gaps.

A significant observation from the study is that bezels played a role in users' clustering and spacing habits. A previous study [75] already had observed a bezel adaptation strategy employed by most users when working with tiled LCD surfaces. Their work showed that bezels tend to help users to quickly separate multiple applications and tasks, which significantly decreased context switching. In my study, I observed similar behaviors where participants attempted to make use of the bezels, mostly by using it as an aid to separate clusters. P4 and P6 made use of bezels by placing inter-cluster gaps where the bezels were. P4 in particular seemed to adapt their entire grouping strategy to the bezels since they had created two-column lines, which separated by a single-column gap where the bezels were.

On the other hand, bezels have disadvantages too. I noticed on a few occasions that users missed some detail on a chart since the chart was placed across a bezel and the continuity of the image was distorted. Then, participants could either not see a diagonal trend properly, or could not categorize the label name correctly since the text was running over to the next screen.

4.4.2. Space Usage

Moreover, I observed how much horizontal space was used, how much vertical space used, and how the content was oriented on the workspace (Table 4.2).

If the user runs out of free space, DynSpace+ permits vertical, but not horizontal scrolling. Vertical scrolling is easier with the mouse wheel, while horizontal scrolling is less accessible in today's graphical user interfaces.

Out of nine participants, seven used the entire width provided by the system. P0 and P1, though, used only about 2/7 and 3/7 of the space respectively. P0 reported that the large space distracted them, so they wanted to limit the content to a smaller PFOV. Due to this trade-off, P0 and P1 needed much more vertical space than available, 210% and 170% respectively. P8 also used about 150% of the available vertical space and thus also needed to scroll up and down to reach all the charts in the workspace. However, in P8's case, this was mostly due to non-ideal arrangement of the content that did not fit a rectangular workspace well. Other users used almost all the visible space to lay out the same amount of content and did not have to resort to virtual navigation through scrolling.

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 slightly biased towards the left side. One potential explanation is the data dimensions panel that they were interacting with was on the left. In contrast, P8's content was concentrated more towards the right side, as also confirmed in that participant's self-report. Other users' clusters seemed to be distributed across the whole display space.

Table 4.2. Users' horizontal and vertical space usage is given below.

	P0	P1	P2	P3	P4	P5	P6	P7	P8
Amount of used horizontal space	2/7	3/7	7/7	7/7	7/7	7/7	7/7	7/7	7/7
Use of vertical space (% of visible)	210%	170%	100%	100%	100%	100%	100%	100%	150%

4.4.3. Navigation Techniques

Earlier studies reported a preference for physical over virtual navigation [4,11,12], as discussed above. However, in my study, I observed mixed results around navigation.

P0 and P1 clearly preferred virtual navigation over physical navigation in my study. They used nearly 37 degrees and 53 degrees of PFOV respectively, leaving the remaining space blank (Figure 4.10). During clustering, both users ran out of vertical space, and relied on 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.

P8 used both virtual navigation through vertical scrolling and physical navigation by body/head/eye rotation to cover the complete 131° HFOV. However, it seems that for this participant vertical scrolling was more like an undesired outcome of an unsuccessful attempt at organization, caused by inefficient usage of space. The participant stated that they preferred physical navigation.

Figure 4.10. The final state of P0's organization. The scroll bar is on the right edge, emphasized with a two-headed arrow. Most of the space was left empty, which caused the user to run out of vertical space early on. This participant still preferred vertical scrolling instead of using the empty screen space on the right for grouping. Later the participant explained that they wanted to work close to the data panel on the left, and did not want to turn around so many times to check the right side of the screen every time.

The rest of the users used physical navigation techniques during the tasks. However, P5 stated that he did not like navigating physically, and perceived it to be very time consuming. P7 was observed to lean forward as a different type of physical navigation, to be better able to visually focus on the displays. P3 reported that they enjoyed physical navigation and prefers that over virtual navigation: "*even if [they] had needed to navigate virtually, [they] would have rather used tabs than vertical scrolling*".

4.5. Discussion of Study I

In study I, charts were uniform, i.e., all charts were scatter plots. I wanted to avoid displaying multiple chart types, since this decision might affect users' data classification. For example, users could have grouped a scatter plot and a bar chart depicting data dimensions from the same domain differently, just because "they appear different", which would not be an data-based classification decision. And indeed, I observed chart classification behaviours based on visual appearance, see the results detailed in section 4.4.

4.5.1. Clustering

For further analysis, I identified five different topics to investigate clustering strategies that participants used and attempted to group users through similarities in their clustering strategies. To make sure that my observations reflect how users actually clustered, after the sessions I explicitly asked them about their clustering, e.g., how many clusters they had and what they considered while clustering. While we were looking at the final version of their clustering attempt, participants explained their groupings, showing them and explaining what they considered while doing such. Each user's clustering behaviour seemed to be unique. Thus, I was not able to observe strong patterns or commonalities among users. A detailed breakdown of the clustering criteria I used is given below.

Number of Clusters

The number of clusters that participants created can be classified into three groups: Low (P0, P1, P2, P3, P4, P5, P7), mid (P8), high (P6) number of clusters. The majority (7 of 9) preferred to create six or less clusters, likely due to the lowest allowed number of clusters having been five.

Clustering Criteria

The participants seemed to have used one of the following three strategies when clustering: Similar visual appearance (P1, P2, P7), commonality in labels (P0, P4, P5), or commonality in topics (P3, P6, P8).

Cluster Shapes

When building clusters, charts were either arranged horizontally or vertically (P0, P1, P4, P5), around a center (P6, P8), or in a mixed way (P2, P3, P7).

Inter-Cluster Relations

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 (but not for other users).

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.

One significant observation from the cluster analysis is that bezels affected some users' clustering strategies. The influence of bezels on VA tasks will be further discussed in Chapter 6.

4.5.2. Space Usage

Two of the participants used space in a radically different way. In essence, their approach was orthogonal to the characteristics of the display system and the features it provides: while we had a horizontally wide display system that allows the user to view information on the large display surface just through physical movements, these two users' approach (P0 and P1) depended completely on vertical scrolling with the mouse wheel. As explained above in section 4.4.2, these two participants gathered all charts together, and did not leave any distance between them to spatially encode different clusters. Moreover, as they used very limited horizontal space, their navigation through the charts required frequent vertical scrolling using mouse interaction.

There could be multiple reasons for this behaviour. First, some users might not prefer physical navigation, in contrast to what the ample previous literature suggests. In Study I and also in Study II, introduced below in Chapter 5, there were users who stated that they wanted to keep everything within their central, i.e., foveal, vision. 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 foveal vision) that they can monitor without missing any content.

Another possible explanation is that the advances in technology are pushing users towards certain paradigms. As smartphones are pervading everyday life, we are getting more and more used to scrolling (vertically) through content. It is possible that this form of accessing content has formed navigation habits that shapes how users arrange content, even on wide screens.

4.5.3. Navigation

Users exhibited widely different navigational habits and preferences. Some preferred physical movements over scrolling in a virtual space, while some preferred the opposite. Some even preferred tabs over scrolling, even if the navigation needed to be virtual. In summary, although more users preferred physical navigation, I could not observe the clear, dominant preference over virtual navigation that is suggested by previous work [4,11,12].

4.6. Limitations

Due to the nature of the task (working on large displays, having 100 charts), participants needed to perform a large number of “drag” interactions. Most of the users found this to be a very iterative and tenuous process and expressed a wish to make this grouping process faster. Eight of nine participants asked if there was a feature for selecting multiple charts for drag and drop. Based on this strong desire, I believe that a future version of the software should allow multiple-chart selection, together with the ability to move multiple charts simultaneously.

Chapter 5.

Study II – LHRD impact on VA

In this chapter I describe Study II, which targets RQ2 and RQ3:

- RQ2: Do larger displays help the user to do better in visual analytics tasks? How does task success change as screen size is increased? Task success is here defined as shorter task completion time and better accuracy in visual analytics tasks, such as answering fact-based questions or finding insights in relatively complex datasets among categorical, numerical or time data using scatterplots, bar charts and histograms.

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

The purpose of the study was to determine if display size and resolution had any impact on VA task success, analysis approaches or user preferences. I was primarily interested in the effect of display size (RQ2). After an initial training phase with a separate data set, each participant used our VA tool to analyze three datasets, using one of three display sizes with each dataset (without repetitions).

Also, to investigate the effect of resolution, each participant was randomly assigned to one of three groups, where each group was associated with a different screen resolution.

5.1. Apparatus

The study took place on V4-SPACE, using a newer version of DynSpace+, working with four datasets introduced in section 3.3.

5.1.1. Physical System

V4-SPACE was described in sections 3.1.1.1 and 4.1.1. The same physical system was used for Study II as was used for Study I, but with different numbers of screens, i.e., different physical screen sizes, and varying resolutions.

V4-SPACE is a 1x7 array of large, vertically oriented 4K-resolution displays. Since each of 4K displays have 3840x2160 configuration, the pixel configuration of the entire system with vertical 7 displays is 15120x3840. Participants of Study II worked with 3-, 5- and 7-screen conditions in each task of the experiment (1x3, 1x5 and 1x7 arrays). To create the 5-screen or 3-screen conditions, the outer two or four displays were temporarily disabled.

All participants were randomly assigned to one of 3 different display resolution configurations:

- 3840x2160 (original 4K resolution)
- 1920x1080 (1080p or half of original resolution)
- 1280 x 720 (720p or a third of original resolution)

Refer to Table 5.1 for the experimental design. Figure 5.1 shows a user working on V4-SPACE with a 3-display condition and resolution R2 (HD 1080p).

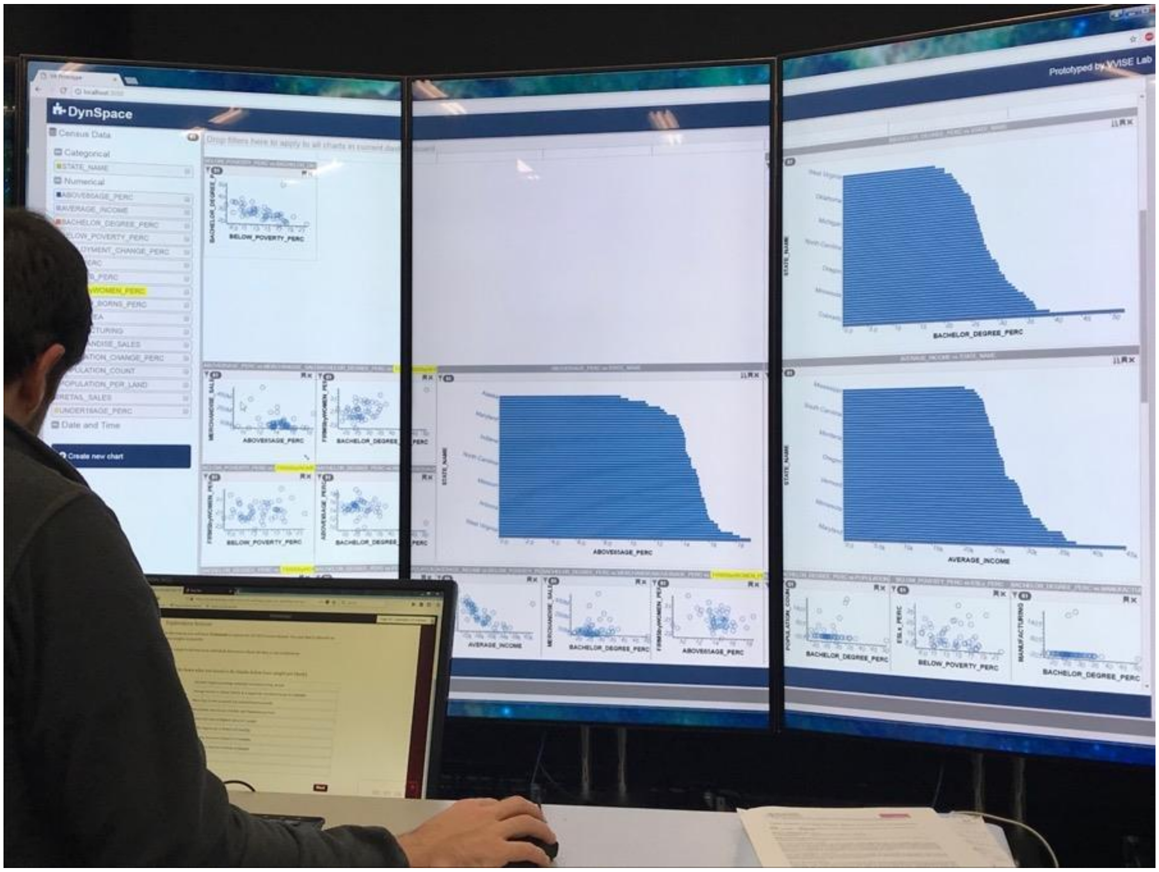


Figure 5.1. Participant working with the 3-screen configuration (S2) and 1080p resolution (R2) conditions of V4-SPACE, using the US Census data on DynSpace+. Tasks are displayed on the auxiliary display, visible in the foreground. Responses are entered on this screen, while analysis is done on large displays.

5.1.2. Visual Analytics Tool

An updated version of DynSpace+ with small improvements, bug fixes and adjustments considering the task was used in this study (bottom image in Figure 3.9). Smaller improvements were based on the fact that this VA tool is still a prototype and is being continuously enhanced, especially for LHRDs. Several more advanced features were added to this version, due to differences between the two studies reported in this document. While a simpler version of the tool was used for the classification task in Study I, for Study II, the tool enabled performing a more complete set of VA operations, such as drag & drop, brushing & linking, and data filtering.

DynSpace+ for Study II differs from that used for Study I as follows:

- Check boxes were added next to each data dimension in the data panel on the left. Users can select those dimensions that they want to work with, which limits the number of initial charts shown in the main visualization panel.

- Histograms and bar charts were added as chart types.

- To support the analysis better, charts could show three dimensions simultaneously through the use of color for the third dimension (Figure 5.2).

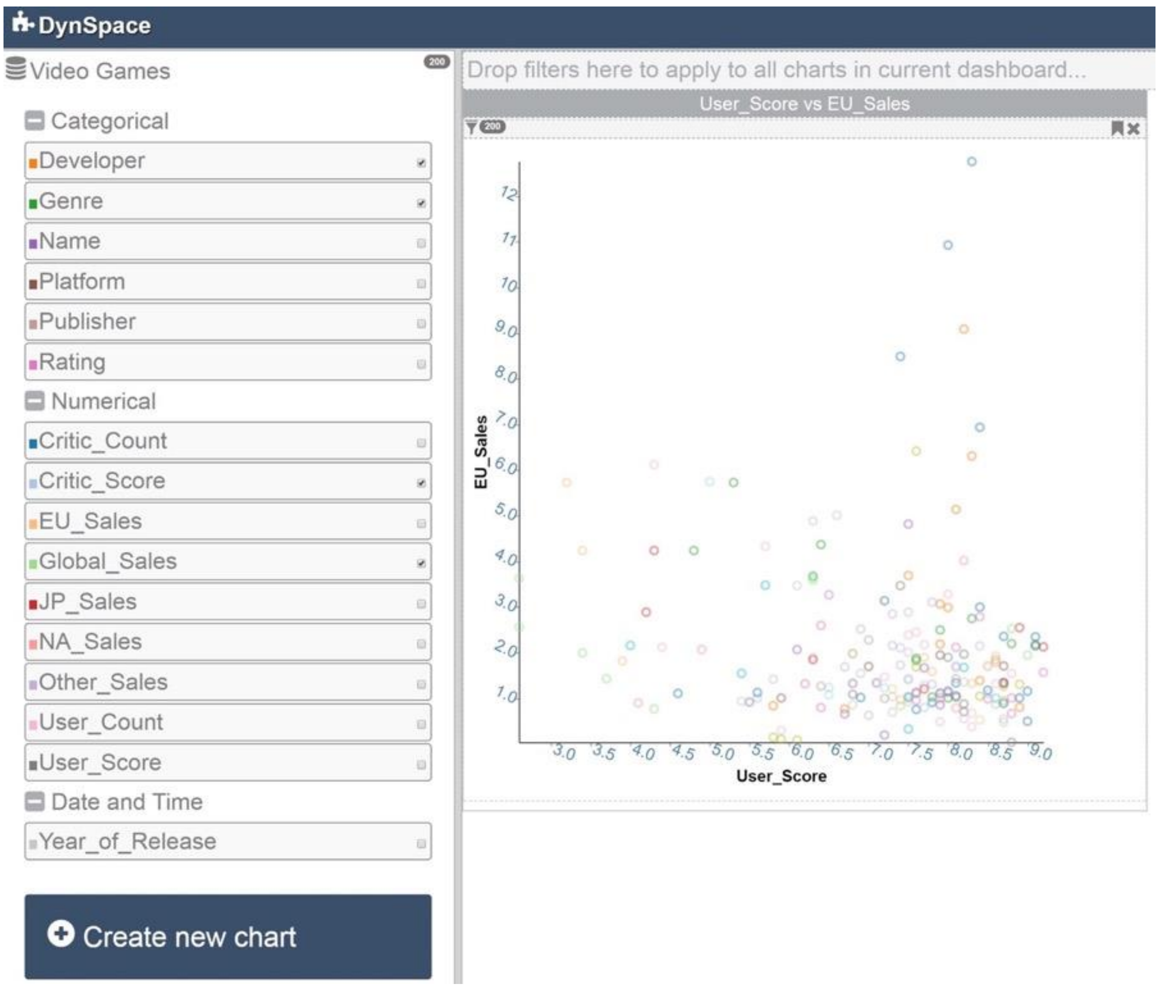


Figure 5.2. Correlogram from videogames dataset. Each color represents a different developer.

5.1.3. Datasets

Four datasets were used in this study. The “Bird Strikes 1996 – 2000” dataset [123] was used in the training session and the analysis of that dataset was not measured. For the three tasks where I collected measurements, I used the “Superstore Sales” [127], “Video Game Sales” [131] and “2016 US Census” [130] datasets.

All datasets consisted of following data types: Categorical, numerical, date & time. The data were shown through histograms, bar charts, row charts (horizontal bars), and scatterplots, as appropriate for the data type and number of dimensions.

5.2. Participants

All nine participants (P0 to P8) were graduate students at SFU SIAT. Three of the nine were females. Seven students were 23 to 28 years old. Two students were 29 or older. Participants were offered a cash stipend for their participation.

There were several criteria for the selection of participants. To eliminate any potential familiarity factor with the software and the display system, participants who participated in any of my pilot studies, Study I of this thesis, or any VA studies conducted by my research group were not allowed to take part in this study. Moreover, as in the first study, participants were required not to have vision problems (I relied on self-reports here) and had to be able to read text on the displays, which was the size of 10pt font-sized text on an Letter format paper at reading distance.

Display resolution varied between participants. Participants were put into one of three groups, referred to as G1, G2 and G3:

P0, P1 and P2 were assigned the lowest resolution condition (G1: 720p),

P3, P4, P5 experienced medium resolution (G2: 1080p),

P6, P7, P8 saw the display at their native, high resolution (G3: 2160p4K).

From an experience point of view, two participants stated that they had more than three years of experience in VA and one participant also reported an experience of more than six months in the field. The remaining six students had less than six months of experience. Typically, their experience consisted of taking one graduate level visualization course.

In the beginning of the study, I also asked participants about their knowledge and experience about VA concepts and tools. The questions and participants' replies are listed in Appendix H.

5.3. Experimental Design

5.3.1. Conditions

There were two variables in the study: Display size and display resolution. As noted above, the display resolution factor was investigated *between* participants, by using: groups G1, G2 and G3 with resolutions R1, R2 and R3 respectively.

Size factor was observed *within* participants by having each participant solve a different problem on each of the three display sizes (S1, S2, S3). To avoid any ordering effect, I applied the *Latin square* method to order the three conditions. A Latin square arrangement ensures "each display condition to occur once for each participant and once in each order" in my 3 x 3 arrangement of three participants in each experiment groups and three display size conditions. Table 5.1 summarizes my experimental design.

Table 5.1. Groups, participants, displays sizes for each of the three tasks, and display resolution conditions for Study II.

Groups	Participants	Task1 Size	Task2 Size	Task3 Size	Resolution
G1	P0	S2	S3	S1	R1
	P1	S1	S2	S3	
	P2	S3	S1	S2	
G2	P3	S1	S2	S3	R2
	P4	S3	S1	S2	
	P5	S2	S3	S1	
G3	P6	S3	S1	S2	R3
	P7	S2	S3	S1	
	P8	S1	S2	S3	

Each task used a different data set, to reduce any potential learning effects. Datasets were chosen to be similar in size and complexity. Based on this, I assumed the tasks to be similar in terms of how challenging they were, since I designed all tasks to require equivalent effort: First, a simpler question that could be answered right away by interpreting the relation of two dimensions on a chart, followed by a question that required sorting, then a question that focused on digging into details, and subsequently a final question that required a sequence of multiple analysis steps such as filtering, comparisons, and information. All other factors were kept identical across users.

5.3.2. Procedure

Initially, I briefly introduced the experimental setting to each participant. Participants could adjust their swivel chair and move rotationally but were told not to get out of the chair, as they needed the keyboard and mouse for interaction, or move the chair around, as that would change the FOV and PPD.

Next, I gave the participants the questionnaire for the study, which was a single-page pre-study survey (see section 5.2 and Appendix H).

After that they saw a practice task. Before the task, I verbally introduced the VA tool in 10-15 minutes by demonstrating the main features in a predefined sequence. Then, I asked participants to complete the training task on their own. They were permitted to ask questions if they forgot about a given functionality, or if they got confused. They were required to finish all steps of the training task until they could achieve the tasks on their own, i.e., were reasonably comfortable with the analytic capabilities of the tool. I did not start the actual tasks until participants ran out of questions, which I considered as them “being ready to start the study”. This phase was not time restricted. More information from the practice session is provided in Appendix I.

After the practice phase, the three formal study tasks were given to participants, as discussed in section 5.3.3. Following the VA tasks, a post-study questionnaire was completed by the participants (Appendix J).

5.3.3. Tasks

For each participant, there were three different tasks. For each task, participants worked with a different dataset on a different display size condition. Each 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 identify (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 [91] (Figure 5.3). In the second part, users were asked a specific set of questions about the dataset; participants could find the answers by using the VA tool to analyze the dataset. A sample subset of questions is given in Figure 5.4. The complete set is available online [132].

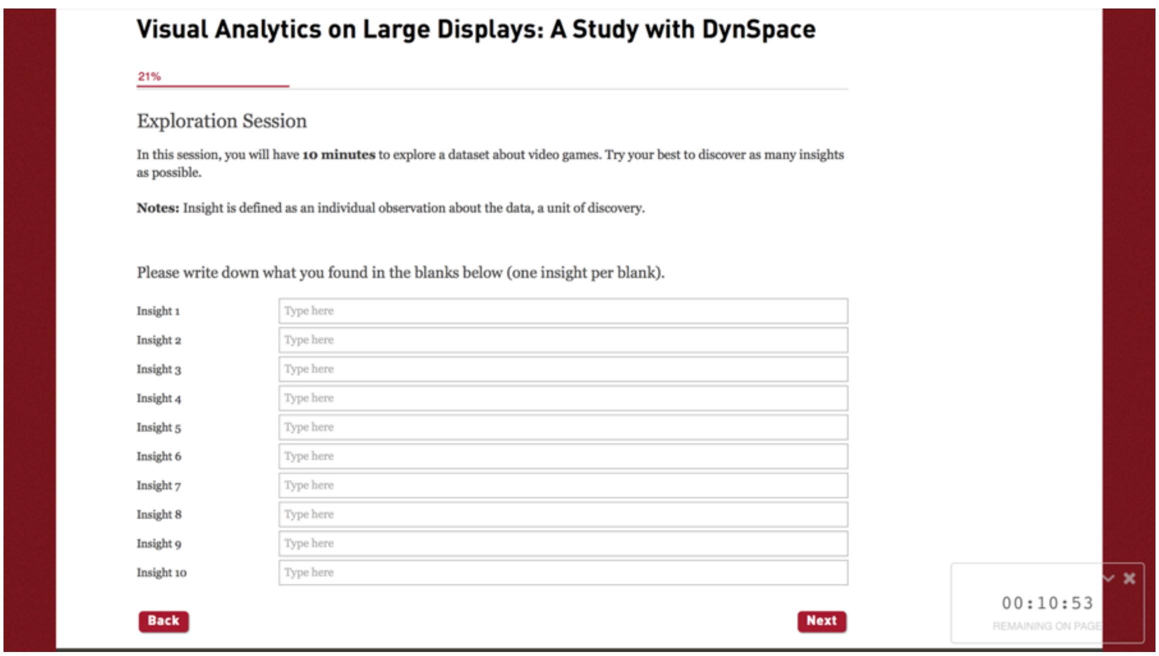


Figure 5.3. Form for the first (exploratory) part of the analysis session with each dataset. Insights are open ended. Users could type in any finding from the dataset they thought was valuable.

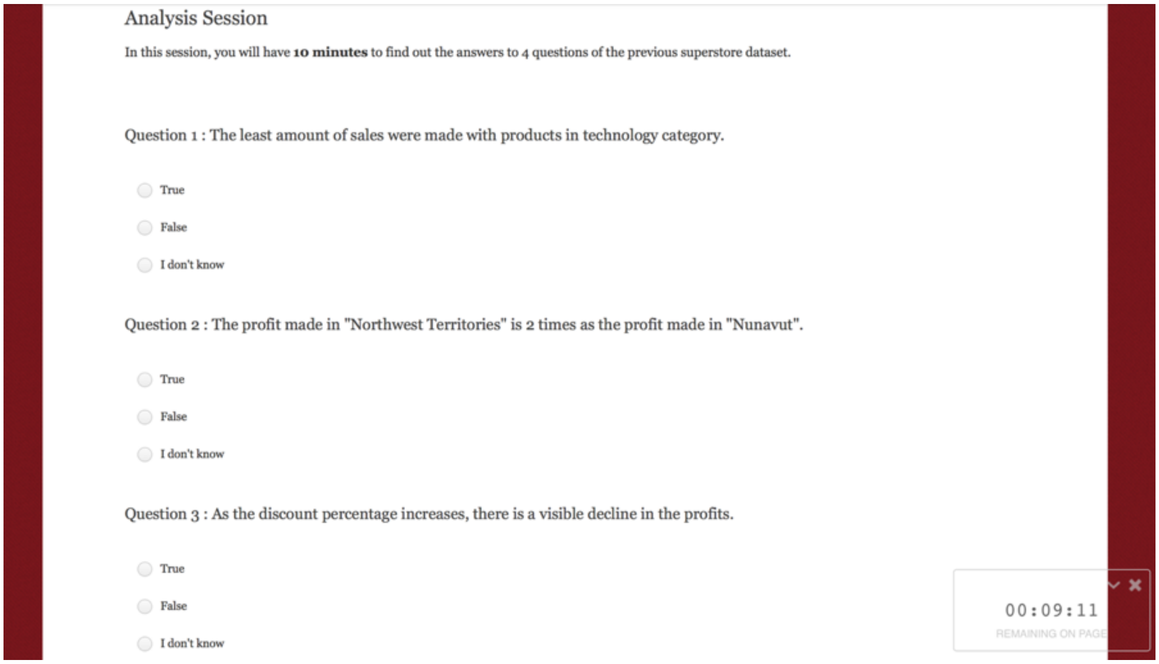


Figure 5.4. Form for main VA tasks with structured questions (three of four tasks are visible in the figure). This measures the participant's ability to follow a pre-determined set of reasoning steps to identify facts about the dataset. The timer visible in the bottom right corner of the screen is for all four questions.

Time was a constraint for all parts of the tasks. Users were given eleven minutes for each of 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, 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.

5.3.4. Data Collected

The questionnaire that the users of Study II completed before, during, and after the study were recorded in a single online form. Their responses to all stages of the study were saved, and a report file was generated by the SFU FluidSurveys online survey system.

In addition to user responses, I observed all notable user behaviours during all experimental stages for every participant and took notes. My notes included observations about their analysis approaches, task behaviours, operations they used, challenges they encountered, and solutions they discovered. I did not adopt a structured way of taking notes, noting only those observations I considered important. Typically, my notes were two A5 pages per participant. Those notes informed the discussion of the study in section 5.5.

Finally, I screen recorded the study sessions for each participant by taking an automatic screenshot every few seconds.

5.4. Results

5.4.1. Quantitative Results

Task Scores

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

To analyze the effect of display size (within subjects) and resolution (between subjects) I created a scoring system to obtain a numerical value for task performance.

The scoring system assigned each participant a single point for each insight they noted down in the first part of the analysis. Time was limited giving participants approximately one minute per insight, which made this somewhat of a challenge. A time bonus for recording a complete set of ten insights before the time expired was applied as an one extra point. Consequently, the highest possible score from the first part of each dataset analysis was eleven points.

The second part of the analysis consisted of four questions on each dataset, usually requiring the user to perform a sequence of operations to find the answer using DynSpace+. Each correct answer received three points. If users selected “I don’t know”, they were unable to answer the question, which resulted in zero points. Wrong answers were penalized with a single negative point. The highest possible score from this part was twelve points.

In summary, each participant was given three different datasets, each with a different problem and a different screen size. They could score up to 23 points on each task. The results of the scoring are shown in Table 5.2.

Table 5.2. Study scores. See text for detail.

Participant	Condition	Score	Resolution	Dataset	Order	Gender
P0	S3	15	R1	C	O3	M
P0	S5	19		V	O1	M
P0	S7	16		S	O2	M
P1	S3	14		V	O1	M
P1	S5	17		S	O2	M
P1	S7	17		C	O3	M
P2	S3	20		S	O2	M
P2	S5	23		C	O3	M
P2	S7	20		V	O1	M
P3	S3	15	R2	S	O1	F
P3	S5	15		V	O2	F
P3	S7	19		C	O3	F
P4	S3	20		C	O2	M
P4	S5	23		S	O3	M
P4	S7	15		V	O1	M
P5	S3	20		V	O3	F
P5	S5	10		C	O1	F
P5	S7	19		S	O2	F
P6	S3	15	R3	S	O2	M
P6	S5	18		V	O3	M
P6	S7	11		C	O1	M
P7	S3	20		V	O3	M
P7	S5	23		C	O1	M
P7	S7	15		S	O2	M
P8	S3	9		C	O1	F
P8	S5	19		S	O2	F
P8	S7	15		V	O3	F

S3, S5 and S7 represent the number of displays in the system. Resolutions R1, R2 and R3 represent 720p, 1080p and 2160p respectively. Dataset, order and gender are only shown for reference. In the dataset column, C is the census data, V is videogames and S is the superstore data. Order is the sequence in which the users were given the datasets and the screen condition.

Task success per display size and resolution conditions are further analyzed in the next section. Task success by total number of pixels is given in Appendix L.

Self-Reported Evaluations

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

- How likely that a display condition was to be their overall favorite,
- How well that display condition enhanced 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, with some preferring each of the three sizes equally. In general, the 5- and 7-display conditions were considered as the best ones in terms of efficiency and effectiveness. However, the 3-display condition—the smallest—was preferred for ease of use.

Participants' overall favorite seemed to be the large (7) display condition, though this was not a significantly different preference. If user ratings are classified as detractor, neutral and promoter ratings, out of nine participants the 7-display condition had four promoter and four detractor ratings for the overall favorite display condition. The 5- and 3-display conditions had two promoter ratings along with five respectively six detractor ratings, respectively. Each condition was considered as the top choice by some participant(s), which indicates a large variation in users' qualitative evaluations. Screen resolution did not seem to have any impact on users' self-reported evaluations.

Table 5.3 shows actual user responses, with ANOVA results for display size conditions are attached under.

Table 5.3. Users ratings of each display size condition out of 10. P6's overall favorite was S3 (10-5-0), P0's favourite was S7 (0-5-10) and P7's choice was S5 (2-9-4).

Participant	Size Condition	Overall Favorite	Efficiency & Effectiveness	Ease of Use
P0	S3	0	8	9
P0	S5	5	9	7
P0	S7	10	10	10
P1	S3	6	5	9
P1	S5	7	6	8
P1	S7	9	8	7
P2	S3	10	7	10
P2	S5	8	8	9
P2	S7	4	4	6
P3	S3	6	4	6
P3	S5	6	6	8
P3	S7	4	4	6
P4	S3	0	0	0
P4	S5	5	10	10
P4	S7	10	10	10
P5	S3	7	6	10
P5	S5	9	9	8
P5	S7	8	9	6
P6	S3	10	8	7
P6	S5	5	6	5
P6	S7	0	2	3
P7	S3	2	3	7
P7	S5	9	9	9
P7	S7	4	4	4
P8	S3	2	5	6
P8	S5	5	7	7
P8	S7	10	10	10

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Condition	2	28.22222	14.1111	1.0311	0.3719
Error	24	328.44444	13.6852		
C. Total	26	356.66667			

5.4.2. Qualitative Results

Direct Observations

Participants' analysis strategies were quite different from each other. Interestingly and based on my observations and notes, different display sizes seem to have encouraged/forced them to make changes to their strategies on the fly during analysis.

P0 usually followed the routine of picking a chart, resizing, hovering on a data point to see pop-up details, and finding an insight. He also did dimension filtering, filter editing, local filter copying, brushing, used global filters, created new charts and asked questions regarding the functionality.

After switching from the S7 condition to S3, P0 commented: "*I was using (all) screens. Now I feel trapped."* To overcome space restrictions, he limited the data to a smaller subset of dimensions to reduce the amount of the data.

P1's strategy was to clear the space first, by filtering most of the data dimensions out and leaving only the ones that he was going to work with. When he started working on a chart, he always resized it first to make it larger. If he was working with bar or row charts, he sorted the bars/columns so that the values of data points appeared low to high, likely to be able to make better comparisons. Characteristically, P1 always worked on a single chart and made it extremely large (Figure 5.5). "*Can we clear the space?"* was P1's first question when he started the analysis, as he did not seem to find it useful to have many visible charts at once. As he had the habit of working with a single, oversized chart in all cases, changes in display size did not seem to affect his work.

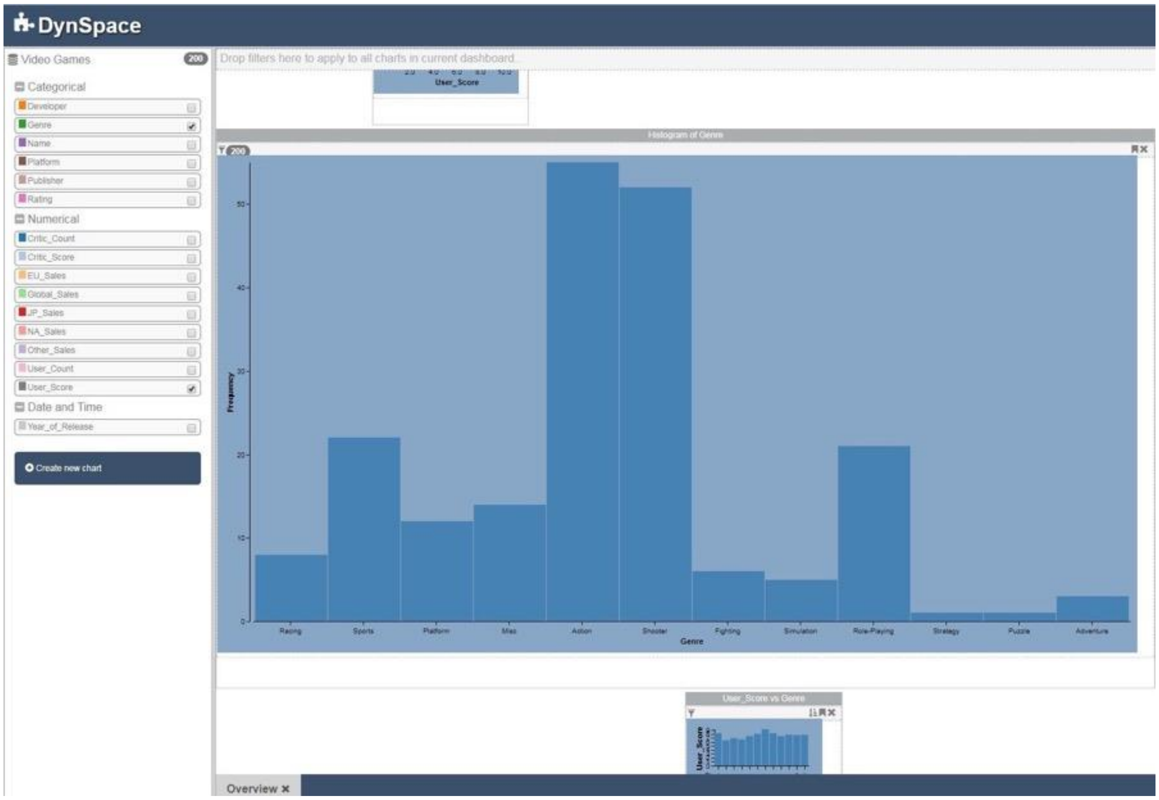


Figure 5.5. Example of the giant charts that P1 used for analysis (the chart above is taking up approximately 3.3 m^2 space on displays, which is nearly 70% of the vertical space).

P2's analysis strategy consisted of resizing chart(s) first, thinking aloud, making analysis comments and then reaching insights. His analysis seemed to have been affected negatively by two factors: vertical display bezels and the small, fast mouse pointer on large displays. Trying to avoid the bezels and occasionally losing the mouse pointer slowed down this participant's analysis process. Bezels also impacted task success, by misleading this user and causing errors in reading during the analysis. For example, at one point of the analysis, P2 was reading data points, which were: 10-12-14-16. P2 overlooked part of the information as it was displayed across a bezel and read it as 10-14-16, which yielded a totally different answer to the task question.

P2 preferred smaller screens as he commented during the analysis: "*[The] 3-screen setting is better for me. Everything is in focus*."

I observed with some participants that some discoveries about how they could use the system in the earlier part of their session may have affected the rest of their analysis behaviour. For example, P3 sped up their analysis when they discovered the strategy of not filling up the space and working on fewer data charts. Over time, this approach helped improve P3's success rate.

P3 commented that "*I guess I filled [the space] up too much*", when they started scrolling up and down, through the content that was nearly 10 times the display height. Surprisingly, though, P3 did not mind scrolling, probably due to being very used to vertical scrolling on most of the devices that P3 was familiar with. P3's favourite was the middle case [S5] and that participant explained that "*[S3] was too small and in [S7] there was too much data at the periphery.*"

I observed similarities between P4's approach and the approach of P3 and P0. A main characteristic of P4's analysis routine was to move charts often in the S7 condition. When P4 switched to S3, he also switched to vertical scrolling from horizontally moving charts around, which he did not seem very happy with (in contrast to P3). P4 felt restricted by the display size like P0, and just like P0, adapted a strategy of picking fewer data dimensions to tackle display size restrictions.

P5 seemed relatively less interested. This participant interacted with the data less than other participants. Resizing and moving charts seemed to have helped other users to get "involved" with the data but P5's behaviour involved little resizing. Related to this decreased amount of interactivity, P5 sometimes did not follow instructions to identify answers.

P6's approach was similar to that of P5. As this participant lacked VA experience, they struggled in the beginning. As P6 started to understand more and get better at analysis they started to interact more with the data. The frequency of resizing charts increased and a larger variety of actions (beyond just looking at charts) was used, which yielded a better task performance.

P7 had a mixed approach, making use of most of the introduced capabilities of the VA tool. This participant had a productive session and identified most of the answers. However, for P7 fatigue seemed to impact performance as time advanced, as this participant also commented on during the tasks.

P8 is another example to improving performance significantly after developing a new analysis strategy on the fly, similar to P3. This participant figured out that one could drag a data dimension (say, X) onto an existing chart (say, Y vs Z) to switch the dimension with one of the existing data dimensions (X substitutes for Y). This enabled this participant to identify information on the new chart (X vs Z) and was then repeated with a new data dimension. P8 thus discovered that it becomes much faster to look at correlations this way as compared to creating charts from scratch. Moreover, and similar to P2, dealing with mouse pointer loss issues negatively impacted P8's analysis timewise.

Self-Reported Results

In the end of the analysis sessions, users were asked to answer questions regarding their experience. Here, my purpose was to learn more about users' opinions about display sizes, their space and content management strategies, navigational preferences, and overall analysis approach.

Display Size

Mostly, users preferred a wider analysis space over a narrower one. P0 said: "*With the full width, I could forget about the rest of the display and focus only on the data, while both others left me wanting to expand the windows."* P8 said: "*I like it when I used the largest display. The small one was too small for me."*

Furthermore, P1 points out a relation with multitasking: "*Wider works best for me, since it allows me to solve multiple questions at the same time."* P3 also prefers wider space "*to be able to compare different graphs on full scale at the same time."*

In contrast, P2 commented: "*The 3- and 5-screen setup is so immersive, making multi-dimensional data analysis easier*." P5's response stated that: *"Generally, the widest screen was harder to control, and I felt the need to move the charts in front of me*." P6 considers the time that a person spends on looking at charts: *"[I prefer] narrower. Less time [is] needed to look for the data, especially when there are 20+ charts opened (that are very similar visually)."*

Finally, some participants realized advantages and disadvantages of both ends of the display space spectrum: P7 summarized it as follows:

> *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 users managed to use the horizontal space efficiently. P4 said: "*On a few occasions, I put together multiple charts to see insights*." Others did not. P1 said: "*I usually expanded the plots in the center of the screen, while there was extra white space [on the] far-right end*." Also, P5 added: "*I usually moved the charts I am interested in to the middle screen and enlarged them to see better*."

Content Management

Initially, DynSpace+ provides of charts right from the beginning to the user, as explained in Chapter 3. Most of the users like this feature, including P2: "*I feel the system was effective at getting a bird's eye view of the trends and correlations*." P3 stated: "*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 only look the graphs I needed*." P7 mentioned: "*The initial figures that are generated automatically made the process and exploration faster. The filters and sorting were quite useful to find a specific detail about the data.*"

However, P6 felt that having too many charts open at the same time was a little bit confusing at times. P4's opinion was similar: "*Although there were ready-to-use charts, they seem overwhelming. Therefore, I relied on creating new charts every single time.*"

Navigational Preferences

Study participants were asked the following question: "Did you prefer virtual navigation (i.e. mouse scrolling vertically), or physical navigation (i.e. turning around, moving the head or body to see different parts of the display)?"

P2, P5 and P7 preferred virtual navigation. P5 explained as follows: "*I prefer virtual navigation because it feels more effective and faster that way*." P6 responded with "*a mix of both*". However, the majority of the users still preferred physical navigation. P4 said: "*With physical navigation, it was easier to create a mental map of where exactly a chart is located at. When I am working with a bigger screen, I tend to lose the same* *mapping if I rely on virtual navigation*." P3 said that physical navigation was better, as using mouse was a bit tiring for this large display.

Analysis Approach

Analysis approaches varied much among the participants. Every user had their unique way of exploration. P3 summarized their approach as follows:

> *I first tried to see the numerical values in terms of the categories I felt [that were] important. For the questions, I filtered the variables that were wanted and then scaled up the graph to better see what was presented. A crowded scene where I was exposed to all the graphs was not really providing me an advantage, so I tried to eliminate everything I did not need. I used the filters as I needed to. I guess I could have used the creating new chart feature, but I haven't. Bookmarking was also something I barely used.*

5.5. Analysis of Results and Discussion of Study II

5.5.1. Discussion of Quantitative Results

Two-way Mixed ANOVA Results for Display Size and Resolution

To analyze the quantitative task results, I applied two-way mixed ANOVA on the data. The results showed that there was a significant effect of the number of screens ($F(2,8) = 4.14$, $p= < .05$). According to the post-hoc test, the S5 condition significantly outperforms S3 and S7. Resolution did not have a statistically significant impact. I also verified that there was no significant interaction between the two factors. Scores per each condition are in Figure 5.6. The means and statistical characteristics for each condition are provided in Table 5.4.

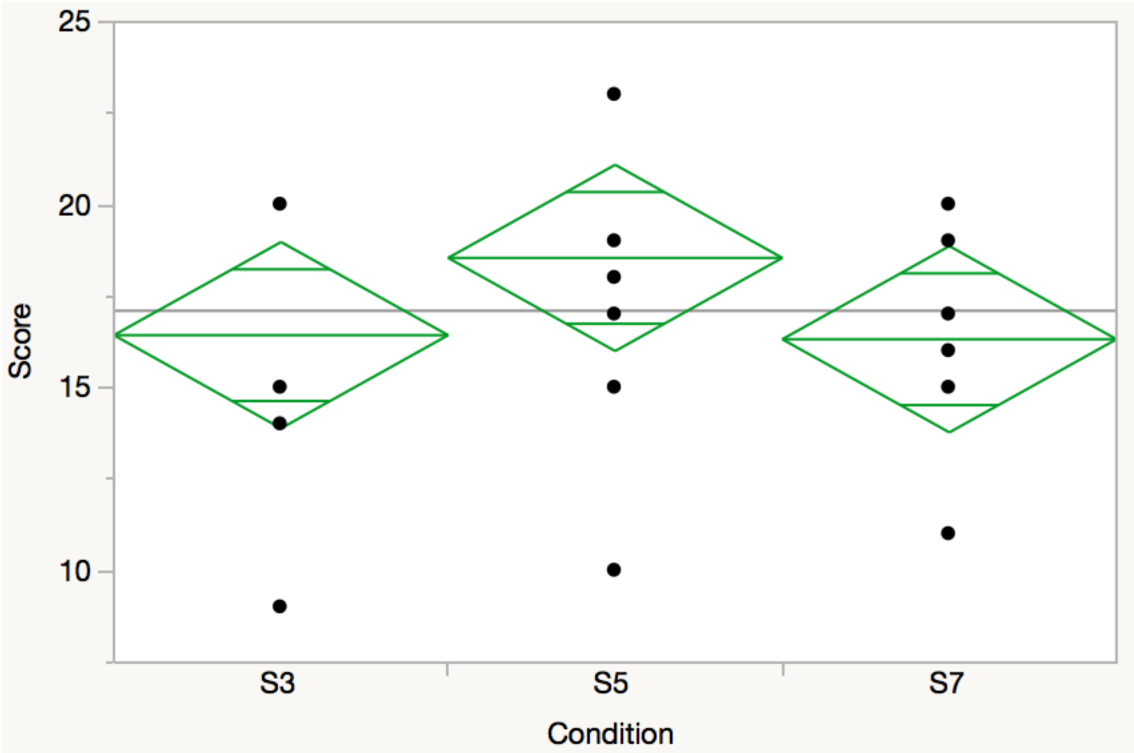


Figure 5.6. **Task scores in in the S3, S5, S7 conditions. The conditions were given in Latin square fashion.**

Table 5.4. **Data for within-subject display size experiment.**

Level	Mean	Std Error	Lower 95%	Upper 95%
S3	16.44	1.23	13.90	18.99
S5	18.56	1.23	16.01	21.10
S7	16.33	1.23	13.79	18.88

Looking at the whole data there is a statistical trend for the S5 condition to perform visibly better than S3 and S7 (Figure 5.6). When I looked closer at the data I noticed that previous user experience seemed to play a role. To eliminate the effect of task familiarity and experience between users, I normalized the data and ranked task successes out of 10, giving the maximum points to each users' best attempt. Rescaling the scores allowed each user to score 10 with their best attempt, which eliminates individual differences between users. 7 out of 9 users achieved a score of 10 in the S5 condition. Although S5 was looked to dominate the other conditions – it attained the top score 7 out of 9 times, there was no significant difference. Looking even closer, I identified that this likely was the consequence of not having removed outliers. I noticed that for each condition, there was exactly one score that seemed much lower than the remaining data (55% or lower success rate, once in each 9 attempts). I removed these outliers from each of the condition and considered only the remaining 8 of total 9 scores for each of the condition. With the normalized data and outliers removed, S5 outperforms the other alternatives significantly. The graph and the statistics are given in Figure 5.7.

Summary of Fit

RSquare	0.505729
RSquare Adj	0.242117
Root Mean Square Error	0.981552
Mean of Response	8.908333
Observations (or Sum Wgts)	24

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	8	14.786667	1.84833	1.9185
Error	15	14.451667	0.96344	**Prob > F**
C. Total	23	29.238333		0.1317

Parameter Estimates

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Condition	2	2	7.9830159	4.1430	0.0369*
Resolution	2	2	1.1070091	0.5745	0.5749
Condition*Resolution	4	4	5.6923659	1.4771	0.2584

Figure 5.7. Two-way mixed ANOVA results of normalized task score.

The results already indicate that S5 outperforms the other two conditions significantly (p < .05). Still, its effects on the performance can be even sharper than how it already looks. To avoid a reduction in task success rate, users appeared to have adapted to different display size conditions by changing their strategies. This adaptation behaviour likely made it harder to observing quantitative impacts. Therefore, I believe that the qualitative results and user feedback are also important to help understand the impact of display size in VA on LHRDs.

Display Size Analysis through User Ratings

I investigated effectiveness and efficiency, ease of use and overall popularity of each display size condition. None of the conditions were statistically significant, and variance was very high. Graphs are in Figure 5.7 and further statistics are in Table 5.5.

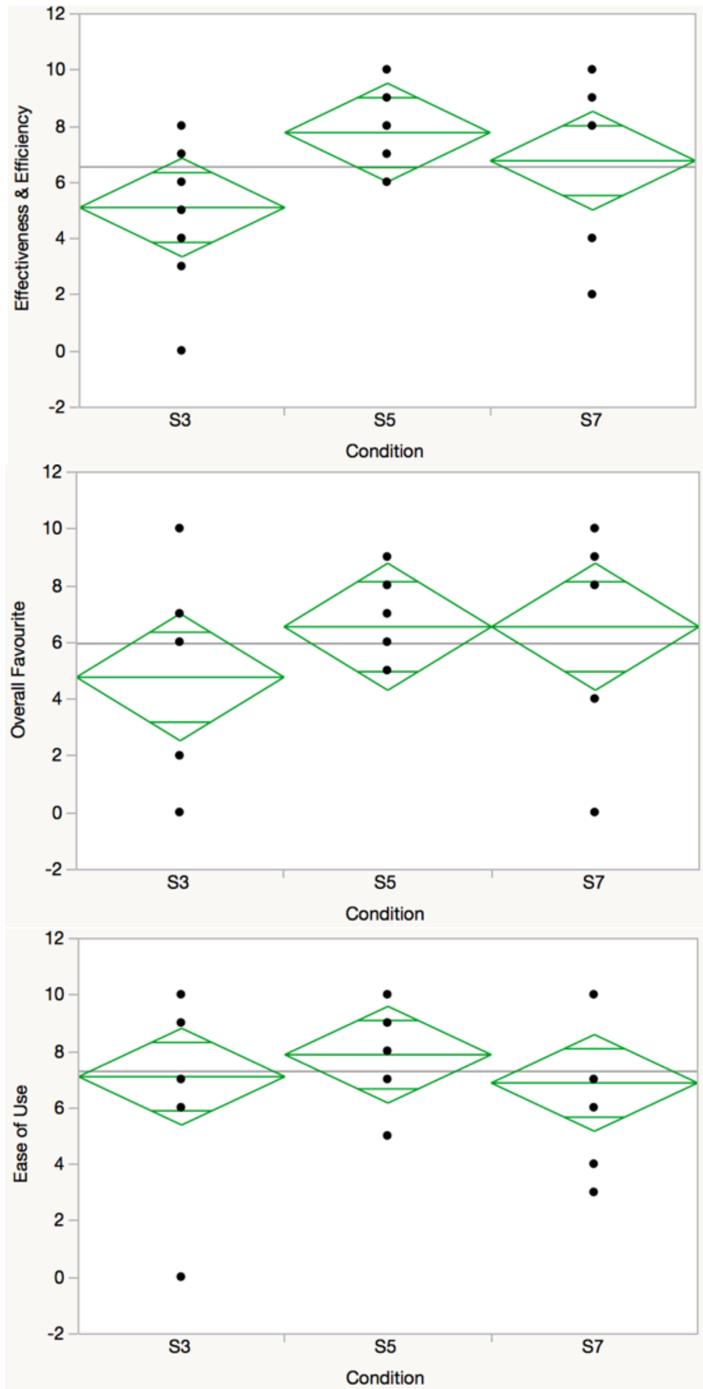


Figure 5.8. Plots of user ratings for 3 categories.

Table 5.5. Data for "effectiveness & efficiency", "overall favourite", and "ease of use" in order of "goodness of fit".

Level	Mean effectiveness	Std Error	Lower 95%	Upper 95%
S3	5.11	0.85	3.36	6.87
S5	7.78	0.85	6.02	9.53
S7	6.78	0.85	5.02	8.53

Level	Mean	Std Error	Lower 95%	Upper 95%
S3	4.78	1.08	2.54	7.02
S5	6.56	1.08	4.32	8.80
S7	6.56	1.08	4.32	8.80

Level	Mean	Std Error	Lower 95%	Upper 95%
S3	7.11	0.83	5.40	8.82
S5	7.89	0.83	6.18	9.60
S7	6.89	0.83	5.18	8.60

For the data shown in Table 5.5, none of the p values are significant (.102, .422, and .674, respectively).

In summary, I can only state that users seem to have different preferences on *how large a large display should be*. Opinions were spread across the spectrum, and not a single criterion seems to be useful to decide between different display size conditions for each participant. This fails to replicate some previous work [4], which demonstrated a clear domination of "physical navigation over virtual", and implied that largest displays will be preferred by most users. If generalized, I can suggest that S5 and S7 seem to be one step ahead of S3. However, the best would be to accept that there are different types of users who prefer to view their content differently.

Display Size Analysis for Different Experience Levels

According to their VA experience levels, I considered users in 3 groups: Not experienced users, users with little experience and experienced users. I observed that not experienced users preferred S7 and experienced users preferred S5. There was no clear pattern for users with little experience. When considering the other two groups, for not experienced ones the decision could have been affected by "how exciting large displays look". Experienced users might have considered trade-offs between advantages of larger size and disadvantages of navigation speed on larger displays, which might have caused them to opt for the middle case. However, further studies are required to identify the precise motivation of those preferences of users with different levels of VA experience.

Display Resolution Analysis

A secondary goal of the study was to analyze the effect of different resolution conditions (R1, R2, R3) through three participant groups (G1, G2, G3), and to study how task success changes between subjects as the display resolution changes, which addresses my RQ3. Figure 5.8 and Table 5.6. show users' performance in terms of the scores they obtained with different resolution settings.

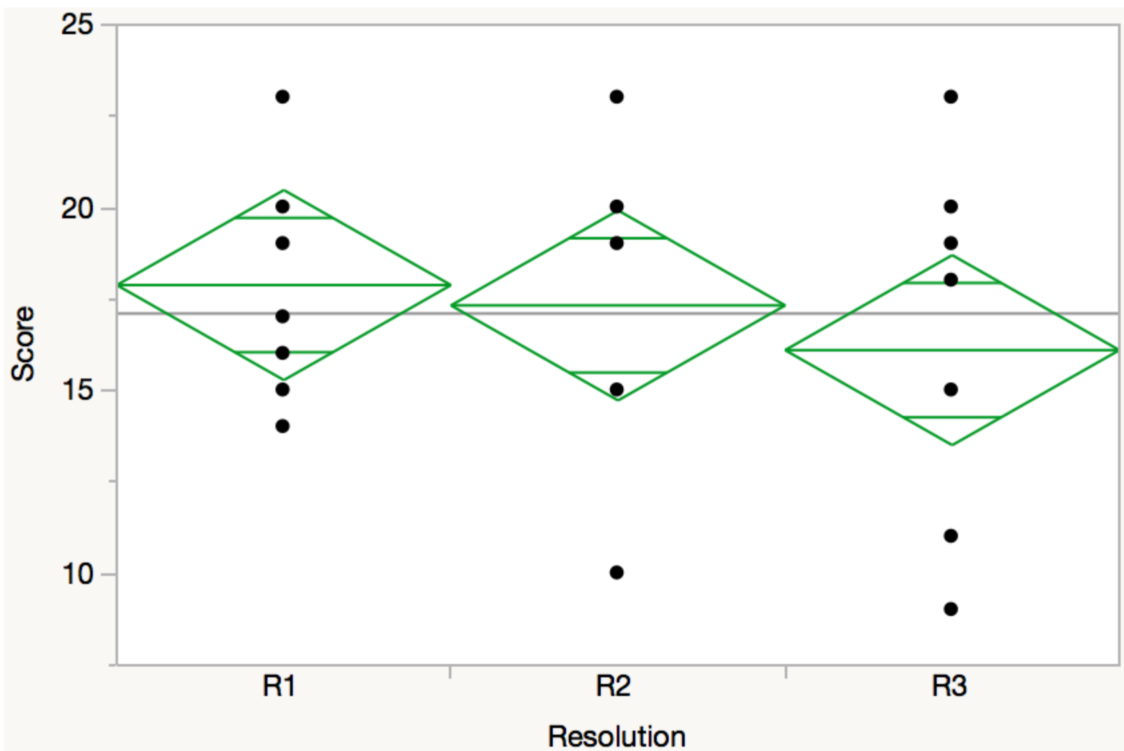


Figure 5.9. The score and resolution graph. R1, R2 and R3 are 720p, 1080p and 2160p respectively. The difference in score was not statistically significant (p=.6).

Table 5.6. Scores (out of 23) per resolution condition.

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
R1	9	17.8889	1.2579	15.293	20.485
R2	9	17.3333	1.2579	14.737	19.930
R3	9	16.1111	1.2579	13.515	18.707

As seen from the figure above, the impact of resolution on quantitative task success measures was not significant. In fact, the trend is counterintuitive, i.e., opposite to what I had been expecting, since R3 is the highest resolution, which had the lowest average result.

Based on my observations, I speculate that quantitative differences due to different resolutions might not be significant, if one considers that participants could tackle resolution-related issues in many ways, even if that create extra effort. For example, if an axis label was not readable at 720p with the initial chart size (due to pixelation issues), users resized the chart to be able to read better. Such requirements surely created discomfort and extra effort but should not result in quantitatively lower scores.

On the other hand, the lowest average scores for R3 is still counterintuitive to me. The only reason I can currently see is that my study had only few people (3 users) in each group.

In Appendix K, further graphs and statistics regarding some other criteria (gender, dataset, task sequence) are given, but were not directly investigated here. They are either assumed or verified to make no difference.

5.5.2. Discussion of Qualitative Outcomes

Using the qualitative results of the study, it is possible to discuss the effects of resolution, size, navigation techniques, usage recommendations, and potential performance inhibitors.

As discussed in section 5.5.1, display resolution did not seem to affect task success, at least directly. Still, I observed that the text around charts, such as chart names, axis labels and data values, become more readable as the resolution increased. P5, P6 and P8 did some of their analysis without enlarging charts. All other users resized their charts to at least two or three times larger than the initial size before starting to work on a chart. P5, P6 and P8 were all in either G2 or G3, who worked with 1080p or 2160p resolutions, where the displayed content was at retina display quality or better. This gives the users the advantage of being able to fit more charts, and therefore more information, into the same space and enhance their analysis. However, it is also important to mention that only some of the users used the whole space effectively, and that thus resolution could be considered as an advantage for only a subset of the large display users.

Concerning the size of the displays, users expressed substantially different opinions. While P0 indicated that *he felt trapped* when he was switching to S3 from S7, P2 said that S3 is better as it provides a display where everything is “in focus”. P3 on the other hand, stated that S3 is too small and in S7 there is too much information in periphery. Although larger displays seemed to be favoured slightly (~45% preferred S7, ~33% preferred S5, which is the largest preference group), each option had some popularity among a certain group of subjects.

Looking at the VA experience level among my participants, I was able to identify 4 groups:

- Level 0 (no VA experience): Both P1 and P8 preferred S7.
- Level 1: Both P2 and P6 preferred S3.
- Level 2: P5 preferred S5, P0 and P4 preferred S7.
- Level 3 (very experienced with VA): P3 and P7 preferred S5.

Overall, more experienced (level 2 and 3) users seem to prefer S5 (or S7, but less so), while the least experienced users (level 0) preferred S7. Users with intermediate experience (level 1 and 2) do not have a clear preference.

Similarly, for the navigational preferences answers also did not show a clear pattern. While some subjects liked to turn around with the chair and rotated their body and head to view the information, others preferred vertical scrolling through the mouse wheel. Another group of subjects said that they preferred *a mix of both* alternatives. As no clear pattern of preference could be identified navigation preferences might be attributable to individual differences.

I also observed mixed results for the initially displayed charts that served as recommendations or alternatives to consider. Some users (P1, P4) did not like them, with P1 asking directly: “*Can we clear [the] space?”* In contrast, P2, P3 and P7 used those charts effectively and they helped them to finish tasks in a shorter time. Here are some related quotes:

> *P2: I feel the system was effective at [giving me] a bird's eye view of the trends and correlations.*

> *P3: 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 only look the graphs I needed.*

> *P7: The initial [charts] that are generated automatically made the process and exploration faster. The filters and sorting were quite useful to find a specific detail about the data.*

Here it is interesting to note that if a participant likes using the recommended charts, display size seems to gain importance. Both P0 and P4 reported that the smaller screens restricted them, as they adapted a strategy of reducing the number of dimensions to tackle the issue of limited space, which affected the time to complete tasks as well as the variety and depth of their analysis negatively.

Participants also provided additional comments about chart sizes and display space. P5 commented: "*I usually moved the charts I am interested in to the middle screen and enlarged them to see better.*" This comment suggests that the user felt the chart was too small to work with. I chose the initial chart size to be small enough to accommodate a large number of charts in the overview mode, with the intent to provide a reasonable overview. With this number of charts, users could see many charts, and easily perceive their titles and any trends visible in the data. However, details, such as the values on the axes, were harder to see. When a user felt that they needed to see more information on a given chart, they enlarged that chart and started working on it. On the other hand, moving charts to the middle cannot be related to visibility, since the display system was curved and the distances of each display unit to the user was the same. As far as I can tell, the reason why some participants adapted the behaviour of moving content to the center, is that they wanted to reduce their physical eye and/or head movements.

Finally, I observed some task performance inhibitors during Study II, namely, the bezels and limits mouse interaction. P2 was adversely affected by both factors. Bezels seemed to be misleading this participant as P2 read the data wrong and eventually found an incorrect answer for the asked question. P8 also seemed to have problems with using a mouse on such a large display and experienced issues such as losing track of the cursor, wrong clicks, and motions that were too fast. That caused P8 to lose time, which resulted in time pressure and a reduction in time task performance.

5.6. Limitations

In this section I review several limitations of my work and discuss their potential impact.

5.6.1. System-Based Limitations

In some cases, bezels limited task success in this study, e.g., because participants misread data presented across a bezel, or had to spend extra time to avoid having charts going across bezels. Here is what some participants reported:

> *P3: Sometimes the bezels were distracting and getting in the way of reading the data effectively.*
>
> *P2: The bezels cut off data.*
>
> *P5: The seams between screens interrupted the view of the charts. I had to move some of them to see clearly.*

Potential future work that could address some of the disadvantages of bezels is discussed in Chapter 7.

5.6.2. Methodology-Based Limitations

The limited number of participants and constraints on the time available per participant is also a limitation of the kind of study I performed. Fundamentally, clearer results might be found if there had been more users taking part in the study. More time for using the system to do analytical tasks, in multiple sittings if possible, would also likely lead to more insights into behaviours. However, real-world time and resource constraints apply for my work as well.

Decomposing the above-mentioned limitations, the first limitation is the number of users taking part in the study. Ideally, there should be more than nine participants for a better study. With this number of participants, only three experienced people each of

G1, G2 and G3 for testing RQ3. A study with a multiple of nine, say 27, participants would likely yield more concrete results, as each condition would have been tested with multiple people. Recruiting a large enough number of subjects with VA experience is one of the main constraints that made this too challenging for the work reported here.

To address the other limitation, it would be desirable to conduct more in-depth, longitudinal studies with the participants. Some limitations of my current study were:

- Time was limited to approximately 90 minutes. For that reason, only reasonably simple analysis tasks could be performed.

- Even though there were practice sessions, it is likely that at least some learning effects are present in the data. One potential reason is that participants experienced only a single session and some participants seem to have forgot some of the features right after learning and had to be reminded again about them during the main study tasks. From my observations and records of having to assist during the study, I believe that two users were particularly affected (P3, P8). Another user, P7, commented as follows: "*I got used to the system in the second analysis. I think I was two times faster after the first analysis.*"

- Time constraints prevented me from testing all conditions within-subjects, since such a methodology would require over 5 hours per participants.

- One participant, P7, got noticeably fatigued during the user study. This may well be due to external influences.

Chapter 6.

Summary and Discussion

In summary, I conducted two VA studies on V4-SPACE using DynSpace+ within the scope of this thesis. The first study was a classification and spatial organization task using a VA tool on LHRDs, whereas the second study investigated effects of changes in display size and resolution on VA task success. My goal was to address three RQs. Study I was designed to tackle RQ1, while Study II helped me to answer my other two research questions. RQs are stated below for reminding purpose:

- 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? What are the factors that users consider when clustering; topical relations, visual similarity, or common dimensions of charts? 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?

- RQ2: Do larger displays help the user to do better in visual analytics tasks? How does task success change as screen size is increased? Task success is here defined as shorter task completion time and better accuracy in visual analytics tasks, such as answering fact-based questions or finding insights in relatively complex datasets among categorical, numerical or time data using scatterplots, bar charts and histograms.

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

The results of the first study are mostly based on (qualitative) observations, as all users were given the same task and there was only a single condition in the experiment. In the second study, I explored the effect of display size and resolution in a mixed design, where display size was investigated within subjects (RQ2) while resolution was explored between subjects (RQ3), respectively. I recorded both quantitative and qualitative measures to analyze task performance and participant preferences.

Through the results of the first study, I was able to identify that users follow substantially different classification and spatial organization strategies, even though they all started with the same state. I observed clear distinctions in terms of clustering strategies, space usage, and preferred navigation techniques. Clustering strategies were further analyzed in section 4.5.1.

In the 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. Based on my observations during the study, the difference could have been even sharper, because I think that even when users are faced with the disadvantages of a given display size or resolution conditions, they seemed to be able to compensate for those disadvantages by changing their analysis strategy. Therefore, I believe that qualitative observations and user self-reports also play a role in identifying the impact of display size and resolution on VA tasks with large displays.

In the following sections, I summarize and discuss the methodology, system design and results.

6.1. Discussion of Methodology

The research done in this thesis explored the impact of a very-large ultra-high-resolution display system on user performance with visual analytics tasks. As Swaminathan and Sato [104] observed, “when a display exceeds a certain size, it becomes qualitatively different.”

Through both qualitative and quantitative measures prior research had observed benefits in terms of user behavior and task performance on large, high-resolution display systems. The methodologies I used in my research builds on this body of work.

In Ball and North's qualitative study [10], the following methods were used: Direct observations as individuals interacted with the display, formal interviews involving a predetermined list of questions, and informal interviews to let the user talk more freely about their experience. Quantitative methods, on the other hand, usually require exposing participants to multiple conditions and measuring the performance. I use both approaches in the two studies that were conducted with the scope of this thesis. Between the two studies, I recorded direct observations of the user interacting with the display, tracked the progress during ongoing interaction, obtained written responses to qualitative questions, asked (predefined) performance-determining questions to measure effectiveness, and used hybrid oral interviews that consisted of a formal interview involving pre-determined questions as well as additional informal questions, which targeted issues that participants had reported during the tasks.

For the two studies that were part of the research, I used various methods to obtain the results. Thus, there are both qualitative and quantitative outcomes, both direct and indirect feedback, as well as verbal, visual or textual responses. Thus, as also discussed in previous work [92], my methodology is labor intensive. Although it seems unlikely that one can completely eliminate the intensity of the labor, while at the same time still using a detailed and thorough methodology, I still attempted to reduce the labor requirements with some design decisions, such as qualitatively observing usage patterns and asking participants for self-reports, as well as keeping configuration logs and screen-capturing the system during the sessions, which provides evidence for further quantitative measurements and analyses.

When seeking study participants for VA studies, sufficiently trained participants are not necessarily easy to find, especially when one is looking for a high number of participants for multiple studies that potentially require substantial amounts of time. Unlike the undergraduate students that I could recruit reasonably easily for participation in my first study, VA knowledgeable participants for the second study, which was heavier in terms of analytic requirements of the tasks, were harder to find and had many more scheduling constraints. However, even though I recruited more knowledgeable participants, Saraiya et al. [92] argue that that this recruitment strategy does not necessarily improve performance in VA studies:

> *One might conjecture that users with more domain experience or software development experience would gain more insight from the data visualizations. Yet, we found that the insight domain value and total number of insights did not appear to depend on participant background. Averages were similar and no significant difference between user categories was detected. Rather, we found that these factors were more dependent on the user motivation. [92]*

Participant motivation is an important factor to consider when designing user studies [79]. Some previous work reported that users working on VA tasks did not analyze the data with much care, looked only for overall effects, but did not delve into great detail [92]. Consequently, I think it would be appropriate to expect a breadth-over-depth approach from average participants, and experimental tasks should be created accordingly. Also, since the motivation factor affects how much engagement a user has with the tasks, motivation should be kept high as much as possible. For this reason, in both of my studies I watched for a certain “voluntariness” factor when recruiting participants. With the expectation of increasing motivation, I used rewards, through course credits for the undergraduate students in the first study and through cash remuneration in the second study.

One of the biggest differences between experimental studies and real-world analysis conditions is the amount of time an analyst could spend on a problem. Analysts are able to work for tens or hundreds of hours on a given dataset or potentially even months. However, experiments typically take at most only a few hours. Longitudinal experiments might help here, but there are number of trade-offs to consider, such as the number of participants to track, controlling the tasks, collecting data, and study timelines.

In my studies, users interacted with a new system they had never used before. They worked with a VA tool that they were completely unfamiliar with. They worked with a dataset that they were not necessarily personally interested in. All those factors required that there was sufficient training and that the time available to perform tasks we long enough, so that the results would not be affected too strongly by inexperience. For the training, the display system was introduced before the tasks, and participants were encouraged to use, train, and adapt to factors such as their position relative to the displays, adjusting seats, mouse controls, cursor tracking, the bezels and looking at the individual screens of the system. Then the VA tool was introduced to the participants and they received tool-related training. For this part of the training, I explained the tool to them, how it works, what was expected from them and also gave them some useful usage tips based on my observations in pilot studies prior to the main studies reported here.

Finally, the participants were introduced to the datasets and they were encouraged to ask questions, make comments, or think aloud during the entire study, also so that I could record feedback and/or provide immediate help for the system, if necessary. In summary, I still attempted to maximize the time that I could spend with each participant, especially for providing training and minimizing learning effect that could potentially impact my measurements.

6.2. Discussion of System and Software Design

6.2.1. Bezels

One notable observation from both studies is that the participants complained about bezels going through a chart. In fact, in some cases bezels intruded upon the analysis and negatively affected the obtained results. That shows the fact that “bezel adaptation” is highly important in VA on LHRDs.

6.2.2. Auxiliary Monitor

In my studies, I kept the analysis space (LHRD) and evaluation space (auxiliary display) separate. While users worked on LHRDs, study information displayed on the small monitor in front of them and they answered the main task questions on this monitor, too. I believe that using such an auxiliary display was a good design decision as displaying the questions on the main screen during analysis would either reduce the available screen space or would require many tab switches between the analysis window/tab and a question window/tab.

An alternative to an auxiliary monitor is paper-based instructions and forms. However, considering my observations of the users during informal, pilot experiments in the earlier stages of the studies, I believe working continuously on the system is a better experimental design, since it does not require the user to switch between keyboard - mouse and pen-and-paper during the task. In support of this statement is the fact that finding the mouse pointer on a LHRD can be costly in terms of time and abandoning the mouse could easily increase the likelihood of losing the mouse pointer whenever the user transitions back from the paper interface to the system.

Another alternative would be to put the reference material on the main display; however, I did not want to interrupt constant visibility of the analysis space by displaying the reference material on top of the VA tool.

6.2.3. Display Settings

Prior to my pilot studies, I did not realize how the brightness of the displays could become a significant issue on such large displays. However, in the pilots I observed that poorly designed user interfaces, e.g., interfaces that use very large white areas, might cause the user to be exposed to too much light, and even heat(!), coming from a nearly 8-meter-wide display system, creating glare and causing discomfort.

Therefore, display settings such as brightness, color, contrast should be calibrated, i.e., set to reasonable levels, for LHRDs to avoid potential discomfort created by heat that can be generated by large display surfaces. Additionally, from the software side, color schemes should be chosen wisely, as their choice can seriously reduce the amount of glare, e.g., compared to the glare caused by a very large display with a bright white background.

6.3. Discussion of Overall Results

My first study showed that LHRD users adapted a variety of classification and spatial organization strategies, and there is no overall, generalizable result other than the observation that clustering approaches can be very person specific. My second study indicated a lack of significant quantitative improvement between tested display resolution conditions, which was investigated by RQ3, however it showed that display size has significant impact on task success, as a response to RQ2. Outcomes of both studies are still very important, as they challenge some of previously identified results for the potential benefits of VA on LHRDs and also point out that there is substantial variation in terms of user behaviours during analysis tasks.

In the conclusion section of Andrew and North's paper [4], they explain their results as follows: "*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*." Some of my findings have a similar character, as I discovered how the needs and preferences of users vary.

My overall conclusion from the studies are:

- Although still being the most popular choice, physical navigation is not uniformly preferred by all users of large displays working on VA tasks. In fact, I observed that the users prefer a happy medium of those.

- As in the literature, large displays are preferred over conventional monitors. However, for different sizes of large displays, bigger is not always better. S5 appeared to be best in my studies and I believe that it is deeply connected to FOV limitations and/or preferences of the users.

- I did not observe a significant performance difference between display resolution conditions, however I do not conclude that resolution has absolutely no effect on VA tasks. Since VA processes involve substantial amount of analytical reasoning, i.e., "thinking", which constitutes some of the time spent on a task, potentially inferior display size or resolution conditions might not impact task performance that badly, if task performance is measured as a whole. That is, even if a certain condition takes extra time in VA processes, the added time caused by interaction operations could be covered by getting analytical reasoning faster in order to meet time restrictions, as I observed during the experiments. Writing ten insights or answering four questions take much less than 50% of the allowed time if those are done in a serialized way without spending any time on thinking. According to my observation of the analysis sessions, the portion of the analysis "cycle" requiring most time seemed to be thinking and figuring out what strategy to follow. Once user decides on that, execution of those thoughts took less time.

- Similarly, if a given display size or resolution has a negative effect on task performance or accuracy, users adopt strategies to reduce those disadvantages through taking extra interaction steps, which effectively translates any negative impact into time loss. For example, if resolution impacts readability, we do not observe a sharp decline in task accuracy. Instead, participants resized the charts for better readability, which overcame any potential readability issues through the price of increased interaction time. However, since interaction time is small relative to the time spent on analytical reasoning during the task, any losses become practically insignificant.

- Bezels can be a major problem for VA task performance, as they can even lead to errors in information interpretation.

- I observed that users fall into different categories through their preferences and strategies. Thus, certain design factors might impact only certain types of users. For example, if a user always takes a chart, places it in the middle of the displays, and then does the analysis always one-chart-at-a-time, changes in the display size will be practically irrelevant for that user, as the user essentially ignores information in the periphery.

Though LHRDs have the potential to offer increased productivity over conventional monitors, they are fundamentally different environments, presenting both unique usability challenges and opportunities, which need to be better understood. I partially addressed this need for additional studies to analyze the impact of LHRD size and resolution settings on content spatialization strategies and VA task performance through the work presented in this thesis. I believe that each finding and observation improves our understanding of the role of LHRDs in VA.

Chapter 7.

Future Work

In a potential future expanded study, it would help to have more participants available, which would help to classify participants more clearly into groups according to their use of the system, and then be able to arrive at stronger conclusions for the behaviour pattern(s) associated with each group.

In the first study, I observed an unexpected behaviour by two of the participants in their use of the LHRD, where they used only a very narrow space and performed their tasks through heavy use of vertical scrolling, i.e., in a way that I think they were accustomed to by everyday devices such as smartphones. It would be very interesting to have further studies on this topic and this is an interesting area for future investigation.

In VA tasks on LHRDs, I noticed that the amount of content displayed may become too much and seemingly simple operations can create noticeable overhead that slows the analysis process. In my study, participants needed to move approximately 100 charts on the analysis space, which started to become overwhelming. Thus, users asked if they could move multiple charts at once, as explained in section 4.6. Therefore, interaction with larger amounts of visual components should become more efficient, and operations that speed repeated tasks should be supported.

One observation from the second study (section 5.5.2) is that display resolution seemed to affect the analysis process qualitatively, even though we did not observe quantitative changes in task success. With 720p (lower than retina display) no user worked on charts as they were in their initial size. In contrast, with Retina resolution or above, 50% of the users (3 of 6) identified some insights without resizing any of the initially provided charts. This means that the qualitative impacts of display resolution on VA task process should be explored in future work.

Generalizing from the previous issue, further studies could benefit from improvements to the methodology, and further build on the observations, experience, and results I obtained. These studies should also address at least some of the challenges encountered in this thesis work.

At the current point in time, it is unlikely that bezels will be completely eliminated in a large (non-projector-based) high-resolution display system since the display system I used was almost 8 meters wide and single unit 8-meter wide monitors do not exist. However, a few improvements might be possible in the future:

- Smaller bezels
- Graphic card level bezel correction techniques (e.g. parts of image that span displays are hidden behind the bezel)
- When improving the software, the chart layout methods could and should be improved to consider the bezels.

Finally, considering the experience reported in my thesis, I believe that any evaluation regarding the size or resolution of LHRDs should be based on qualitative user feedback as well as quantitative measures. Due to several uncontrollable factors like different levels of learning ability, fatigue, various distractions, or different degrees of motivation for sense-making processes, it is very hard to control all the factors and ensure same environment to observe purely quantitative performance during VA tasks. Moreover, users seem to be able to reasonably perform well even if they experience some forms of mild discomfort, which then does not adequately reflect the impact of a given experiment conditions. Considering this, I believe that involving subjective measures in the evaluation of LHRDs might serve as a useful way to further explore characteristics of LHRDs.

References

[1] Alfano, Patricia L., and George F. Michel. "Restricting the field of view: Perceptual and performance effects." *Perceptual and motor skills* 70, no. 1 (1990): 35-45.

[2] 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.

[3] Andrews, Christopher, Alex Endert, Beth Yost, and Chris North. "Information visualization on large, high-resolution displays: Issues, challenges, and opportunities." *Information Visualization*10, no. 4 (2011): 341-355.

[4] 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.

[5] Anscombe, Francis J. "Graphs in statistical analysis." The American Statistician 27, no. 1 (1973): 17-21.

[6] Arias-Hernandez, Richard, Linda T. Kaastra, Tera Marie Green, and Brian Fisher. "Pair analytics: Capturing reasoning processes in collaborative visual analytics." In System Sciences (HICSS), 2011 44th Hawaii International Conference on, pp. 1-10. IEEE, 2011.

[7] Arias-Hernandez, Richard, Linda T. Kaastra, and Brian Fisher. "Joint action theory and pair analytics: In-vivo studies of cognition and social interaction in collaborative visual analytics." In Proceedings of the Cognitive Science Society, vol. 33, no. 33. 2011.

[8] Arlen, Gary (January 6, 2013). "Ultra High-Def TV: Super-Sizing An Immersive Experience". Consumer Electronics Association. Retrieved January 17, 2013.

[9] Arthur, Kevin Wayne, and Frederick P. Brooks Jr. "Effects of field of view on performance with head-mounted displays." PhD diss., University of North Carolina at Chapel Hill, 2000.

[10] Ball, Robert, and Chris North. "Analysis of user behavior on high-resolution tiled displays." *Human-Computer Interaction-INTERACT 2005* (2005): 350-363.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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).

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

[17] 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.

[18] 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.

[19] Bishop, Gary, and Greg Welch. "Working in the office of 'Real Soon Now'." *IEEE Computer Graphics and Applications* 20, no. 4 (2000): 76-78.

[20] Blyth, Colin R. "On Simpson's paradox and the sure-thing principle." *Journal of the American Statistical Association* 67, no. 338 (1972): 364-366.

[21] 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.

[22] Card, Stuart K., Jock D. Mackinlay, and Ben Shneiderman, eds. *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999.

[23] Card, Stuart K., Peter Pirolli, Mija Van Der Wege, Julie B. Morrison, Robert W. Reeder, Pamela K. Schraedley, and Jenea Boshart. "Information scent as a driver of Web behavior graphs: results of a protocol analysis method for Web usability." In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 498-505. ACM, 2001.

[24] Carrasco, Marisa, Denise L. Evert, Irene Chang, and Svetlana M. Katz. "The eccentricity effect: Target eccentricity affects performance on conjunction searches." *Attention, Perception, & Psychophysics* 57, no. 8 (1995): 1241-1261.

[25] Casiez, Géry, and Nicolas Roussel. "No more bricolage!: methods and tools to characterize, replicate and compare pointing transfer functions." In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pp. 603-614. ACM, 2011.

[26] Cater, Kirsten, Alan Chalmers, and Greg Ward. "Detail to attention: exploiting visual tasks for selective rendering." In ACM International Conference Proceeding Series, vol. 44, pp. 270-280. 2003.

[27] Chacksfield, Marc. "UHDTV to be name for both 4K and 8K television standard?" http://www.techradar.com/news/television/uhdtv-to-be-name-for-both-4k-and-8k-television-standard-1082328 (accessed October 19, 2017)

[28] Chang, Remco, Caroline Ziemkiewicz, Tera Marie Green, and William Ribarsky. "Defining insight for visual analytics." *IEEE Computer Graphics and Applications* 29, no. 2 (2009): 14-17.

[29] Cockburn, Andy, Amy K. Karlson, and Benjamin B. Bederson. "A review of overview+ detail, zooming, and focus+ context interfaces." *ACM Comput. Surv.* 41, no. 1 (2008): 2-1.

[30] Cook, Kristin A., and James J. Thomas. "Illuminating the path: The research and development agenda for visual analytics." (2005).

[31] Czerwinski, Mary, Desney S. Tan, and George G. Robertson. "Women take a wider view." In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 195-202. ACM, 2002.

[32] 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.

[33] Dachille, Frank, and Arie Kaufman. "High-degree temporal antialiasing." In *Computer Animation 2000. Proceedings*, pp. 49-54. IEEE, 2000.

[34] Davis, Fred D. "Perceived usefulness, perceived ease of use, and user acceptance of information technology." MIS quarterly (1989): 319-340.

[35] Dill, John, Lyn Bartram, Albert Ho, and Frank Henigman. "A continuously variable zoom for navigating large hierarchical networks." In Systems, Man, and Cybernetics, 1994. Humans, Information and Technology., 1994 IEEE International Conference on, vol. 1, pp. 386-390. IEEE, 1994.

[36] Dill, John C., David J. Kasik, and David J. Darvill. "Case study: successful deployment of industry-university collaborative visual analytics research." In System Sciences (HICSS), 2013 46th Hawaii International Conference on, pp. 1505-1511. IEEE, 2013.

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

[38] Draper, Mark H., Erik S. Viirre, Thomas A. Furness, and Valerie J. Gawron. "Effects of image scale and system time delay on simulator sickness within head-coupled virtual environments." *Human factors* 43, no. 1 (2001): 129-146.

[39] Drawbaugh, Ben. "Ultra HD TVs stole the show at CES 2013, but they're just part of the puzzle" https://www.engadget.com/2013/01/11/ultra-hd-tvs-stole-the-show-at-ces-2013/ (accessed October 19, 2017)

[40] Duh, HB-L., J. W. Lin, Robert V. Kenyon, Donald E. Parker, and Thomas A. Furness. "Effects of field of view on balance in an immersive environment." In *Virtual Reality, 2001. Proceedings. IEEE*, pp. 235-240. IEEE, 2001.

[41] Dulberg, Martin S., Robert St Amant, and Luke S. Zettlemoyer. "An Imprecise Mouse Gesture for the Fast Activation of Controls." In *Interact*, vol. 99, p. 375. 1999.

[42] 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.

[43] 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.

[44] 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.

[45] 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.

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

[47] Folk, Charles L., Roger W. Remington, and James C. Johnston. "Involuntary covert orienting is contingent on attentional control settings." Journal of Experimental Psychology Human Perception and Performance 18 (1992): 1030-1030.

[48] Folk, Charles L., and Roger Remington. "Selectivity in distraction by irrelevant featural singletons: evidence for two forms of attentional capture." Journal of Experimental Psychology: Human perception and performance 24, no. 3 (1998): 847.

[49] Furnas, George W., and Benjamin B. Bederson. "Space-scale diagrams: Understanding multiscale interfaces." In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 234-241. ACM Press/Addison-Wesley Publishing Co., 1995.

[50] González, Victor, and Alfred Kobsa. "A workplace study of the adoption of information visualization systems." In *Proc. I-KNOW*, vol. 3, pp. 92-102. 2003.

[51] Grudin, Jonathan. "Partitioning digital worlds: focal and peripheral awareness in multiple monitor use." In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 458-465. ACM, 2001.

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

[53] Huber, Peter J. "Huge data sets." In *Compstat*, pp. 3-13. Physica, Heidelberg, 1994.

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

[55] Johnston, Casey. "Whatever happened to 4K? The rise of “Ultra HD” TV" https://arstechnica.com/gadgets/2013/01/whatever-happened-to-4k-the-rise-of-ultra-hd-tv/ (accessed October 19, 2017)

[56] 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.

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

[58] 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.

[59] Lamprecht, Andreas, Annette Hautli, Christian Rohrdantz, and Tina Bögel. "A Visual Analytics System for Cluster Exploration." In *ACL*, pp. 109-114. 2013.

[60] Larkin, Jill H., and Herbert A. Simon. "Why a diagram is (sometimes) worth ten thousand words." *Cognitive science* 11, no. 1 (1987): 65-100.

[61] Lehrer, Jonah. "The eureka hunt." *The New Yorker* 28 (2008): 40-45.

[62] Lischke, Lars, Sven Mayer, Katrin Wolf, Niels Henze, Harald Reiterer, and Albrecht Schmidt. "Screen arrangements and interaction areas for large display work places." In *Proceedings of the 5th ACM International Symposium on Pervasive Displays*, pp. 228-234. ACM, 2016.

[63] Liu, Can, Olivier Chapuis, Michel Beaudouin-Lafon, Eric Lecolinet, and Wendy E. Mackay. "Effects of display size and navigation type on a classification task." In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pp. 4147-4156. ACM, 2014.

[64] Liu, Can, Olivier Chapuis, Michel Beaudouin-Lafon, and Eric Lecolinet. "Shared interaction on a wall-sized display in a data manipulation task." In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 2075-2086. ACM, 2016.

[65] Mai, Xiao-Qin, Jing Luo, Jian-Hui Wu, and Yue-Jia Luo. "“Aha!” effects in a guessing riddle task: An event-related potential study." *Human brain mapping* 22, no. 4 (2004): 261-270.

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

[67] 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.

[68] Matejka, Justin, and George Fitzmaurice. "Same stats, different graphs: Generating datasets with varied appearance and identical statistics through simulated annealing." In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1290-1294. ACM, 2017.

[69] Mayer, Richard E. "The promise of multimedia learning: using the same instructional design methods across different media." *Learning and instruction* 13, no. 2 (2003): 125-139.

[70] McCreary, Faith A., and Robert C. Williges. "Effects of age and field-of-viewon spatial learning in an immersive virtual environment." In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 42, no. 21, pp. 1491-1495. Sage CA: Los Angeles, CA: SAGE Publications, 1998.

[71] Moyle, Michael, and Andy Cockburn. "Analysing mouse and pen flick gestures." In *Proceedings of the SIGCHI-NZ Symposium on Computer-Human Interaction*, pp. 19-24. ACM, 2002.

[72] Navrátil, Paul A., Brandt Westing, Gregory P. Johnson, Ashwini Athalye, Jose Carreno, and Freddy Rojas. "A practical guide to large tiled displays." In *International Symposium on Visual Computing*, pp. 970-981. Springer, Berlin, Heidelberg, 2009.

[73] North, Chris, Remco Chang, Alex Endert, Wenwen Dou, Richard May, Bill Pike, and Glenn Fink. "Analytic provenance: process+ interaction+ insight." In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pp. 33-36. ACM, 2011.

[74] 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.

[75] 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.

[76] Papadopoulos, Charilaos, Kaloian Petkov, Arie E. Kaufman, and Klaus Mueller. "The Reality Deck--an Immersive Gigapixel Display." *IEEE computer graphics and applications* 35, no. 1 (2015): 33-45.

[77] Pashler, Harold, James C. Johnston, and Eric Ruthruff. "Attention and performance." Annual review of psychology 52, no. 1 (2001): 629-651.

[78] Pirolli, Peter, and Stuart Card. "The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis." In *Proceedings of international conference on intelligence analysis*, vol. 5, pp. 2-4. 2005.

[79] Plaisant, Catherine, Georges Grinstein, Jean Scholtz, Mark Whiting, Theresa O'Connell, Sharon Laskowski, Lynn Chien et al. "Evaluating visual analytics at the 2007 VAST symposium contest." IEEE Computer Graphics and Applications 28, no. 2 (2008).

[80] 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.

[81] 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.

[82] Rajabiyazdi, Fateme, Jagoda Walny, Carrie Mah, John Brosz, and Sheelagh Carpendale. "Understanding researchers' use of a large, high-resolution display across disciplines." In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*, pp. 107-116. ACM, 2015.

[83] Rieman, John. "A field study of exploratory learning strategies." *ACM Transactions on Computer-Human Interaction (TOCHI)* 3, no. 3 (1996): 189-218.

[84] 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.

[85] 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.

[86] Rooney, Chris, and Roy A. Ruddle. "HiReD: a high-resolution multi-window visualisation environment for cluster-driven displays." In *Proceedings of the 7th ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, pp. 2-11. ACM, 2015.

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

[88] 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.

[89] Sandin, Daniel J., Todd Margolis, Jinghua Ge, Javier Girado, Tom Peterka, and Thomas A. DeFanti. "The Varrier TM autostereoscopic virtual reality display." In *ACM Transactions on Graphics (TOG)*, vol. 24, no. 3, pp. 894-903. ACM, 2005.

[90] Sandstrom, Timothy A., Chris Henze, and Creon Levit. "The hyperwall." In Coordinated and Multiple Views in Exploratory Visualization, 2003. Proceedings. International Conference on, pp. 124-133. IEEE, 2003.

[91] 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.

[92] 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.

[93] Sawant, Amit P., and Christopher G. Healey. "A survey of display device properties and visual acuity for visualization." *Raleigh: Department of Computer Science, North Carolina State University, Raleigh, North Carolina* (2005).

[94] Schaffer, Doug, Zhengping Zuo, Saul Greenberg, Lyn Bartram, John Dill, Shelli Dubs, and Mark Roseman. "Navigating hierarchically clustered networks through fisheye and full-zoom methods." ACM Transactions on Computer-Human Interaction (TOCHI) 3, no. 2 (1996): 162-188.

[95] 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.

[96] 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.

[97] 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.

[98] Shupp, Lauren Marcy. "The Effects of Curving Large, High-Resolution Displays on User Performance." (2006).

[99] 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.

[100] Sibert, Linda E., and Robert JK Jacob. "Evaluation of eye gaze interaction." In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pp. 281-288. ACM, 2000.

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

[102] Smith, Geneva, Robert J. Teather, Jordan Lass, and Jacques Carette. "Effects of interior bezel size and configuration on gaming performance with large tiled displays." In Games Entertainment Media Conference (GEM), 2015 IEEE, pp. 1-8. IEEE, 2015.

[103] Stroud, Forrest. Retina Display https://www.webopedia.com/TERM/R/retina_display.html (accessed October 19, 2017)

[104] Swaminathan, Kishore, and Steve Sato. "Interaction design for large displays." interactions 4, no. 1 (1997): 15-24.

[105] Tan, Desney S., Darren Gergle, Peter Scupelli, and Randy Pausch. "With similar visual angles, larger displays improve spatial performance." In Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 217-224. ACM, 2003.

[106] Tan, Desney S., Darren Gergle, Peter Scupelli, and Randy Pausch. "Physically large displays improve performance on spatial tasks." *ACM Transactions on Computer-Human Interaction (TOCHI)* 13, no. 1 (2006): 71-99.

[107] Thomas, James J., and Kristin A. Cook. "A visual analytics agenda." *IEEE computer graphics and applications* 26, no. 1 (2006): 10-13.

[108] Venkatesh, Viswanath, and Fred D. Davis. "A theoretical extension of the technology acceptance model: Four longitudinal field studies." *Management science* 46, no. 2 (2000): 186-204.

[109] Wallace, James R., Daniel Vogel, and Edward Lank. "Effect of bezel presence and width on visual search." In *Proceedings of The International Symposium on Pervasive Displays*, p. 118. ACM, 2014.

[110] Wallace, James R., Daniel Vogel, and Edward Lank. "The effect of interior bezel presence and width on magnitude judgement." In *Proceedings of Graphics Interface 2014*, pp. 175-182. Canadian Information Processing Society, 2014.

[111] Wang, Yi, Kunmi Otitoju, Tong Liu, Sijung Kim, and Doug A. Bowman. "Evaluating the effects of real world distraction on user performance in virtual environments." In *Proceedings of the ACM symposium on Virtual reality software and technology*, pp. 19-26. ACM, 2006.

[112] Witmer, Bob C., John H. Bailey, and Bruce W. Knerr. *Training Dismounted Soldiers in Virtual Environments: Route Learning and Transfer*. No. ARI-TR-1022. ARMY RESEARCH INST FOR THE BEHAVIORAL AND SOCIAL SCIENCES ORLANDO FL ORLANDO FIELD UNIT*, 1995.

[113] 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.

[114] Zhai, Shumin, Carlos Morimoto, and Steven Ihde. "Manual and gaze input cascaded (MAGIC) pointing." In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pp. 246-253. ACM, 1999.

[115] ACCELERATE YOUR CREATIVITY NVIDIA QUADRO K620 (2015) NVIDIA

[116] Is This Retina? https://isthisretina.com (accessed October 19, 2017)

[117] NASA. http://photojournal.jpl.nasa.gov/catalog/pia04182, 2005. (accessed October 19, 2017)

[118] NASA. http://www.spitzer.caltech.edu/media/releases/ssc2008- 11/ssc2008-11a.shtml, 2008. (accessed October 19, 2017)

[119] NASA. http://earthobservatory.nasa.gov/features/bluemarble, 2009. (accessed October 19, 2017)

[120] REAL INTERACTIVE EXPRESSION NVIDIA QUADRO M5000. (2015). NVIDIA.

[121] The Electron Microscopy Outreach Program. http://em-outreach.ucsd.edu/, 2008. (accessed October 19, 2017)

[122] "4K AND 8K UHDTV DEFINED" https://tech.ebu.ch/news/4k-and-8k-uhdtv-defined-16may12 (accessed October 19, 2017)

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

[124] “Gridstack.js” https://github.com/gridstack/gridstack.js. (accessed November 30, 2017)

[125] “Largest Cursor Ever! Cursor Details” http://www.rw-designer.com/cursor-detail/60277 (accessed October 19, 2017)

[126] “Razer DeathAdder Chroma - The World’s Best Gaming Mouse” https://www2.razerzone.com/ap-en/gaming-mice/razer-deathadder-chroma (accessed October 19, 2017)

[127] “Superstore Sales” http://bit.ly/2hNVA2g. (accessed November 20, 2017)

[128] "Ultra High Definition Television: Threshold of a new age"." *ITU. May* 24 (2012).

[129] "Ultra-high-definition television" https://en.wikipedia.org/wiki/Ultra-high-definition_television (accessed October 19, 2017)

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

[131] “Video Game Sales” http://bit.ly/2mIIOno. (accessed November 20, 2017)

[132] “Visual Analytics on Large Displays: A Study with DynSpace”, http://bit.ly/2nCGGOC. (accessed December 6, 2017)

Appendix A.

Definitions, Concepts, Terms

1. Display size: This is the physical size of the displays. If display size is defined with two numbers, they express the width and height of the displays, as measured in meters or inches. If only one number is used, it refers to the diagonal size of the display.

2. Display resolution: In this work, resolution is the physical pixel configuration in a display. For example, if a display is said to have 3840 x 2160 resolution, the display contains 2160 rows of 3840 pixels each.

3. Pixel count: It is the total number of pixels in the display, found by multiplying horizontal and vertical number of pixels. If a display has 8 million pixels in, it will be referred to as an 8-megapixel display.

4. DPI/PPI: Dots-per-inch or pixels-per-inch. I use the terms PPI and DPI interchangeably, even though they do refer to different (although related) concepts. PPI, the correct term, means the number of pixels per inch and gives the number of pixels in a horizontal (or vertical) inch on the display.

5. Pixel density: Refers to the concentration of pixels on a display. In this work, PPI is used as a measure of pixel density.

6. Distance to displays: Display specifications alone are not enough to identify perceived image quality. The user's position and the distance of the displays are also important to consider. In my experiments, users had fixed positions, at a known distance, in front of the display system.

7. Pixels-Per-Degree (PPD): PPD considers both the pixel density and the distance from which the device is viewed for expressing a viewing quality measure with a specific unit. Apple marketed the *Retina Display* term, basing it on the angular resolution (PPD), which ensures that pixels are not visible at a given viewing distance, creating an impression of sharp, print-like text. PPD depends on both PPI and distance to the display.

8. Angular visual field: *Display size* and *distance to display*s are both distance measurements. Using those two quantities, a triangle can be drawn from the user's eyes to the center and the edges of the display, from which the angular visual field can be calculated.

9. Visual acuity: A measure of the clarity of an individual's vision. At 20 feet, a "normal" human eye can distinguish lines that are approximately 1.75 mm apart. Thus, 20/20 vision, a term used to express normal visual acuity (the clarity or sharpness of vision) means the individual can distinguish at 20' what a normal human can distinguish at 20'; 20/40 means the individual can distinguish at 20' what normal vision can distinguish at 40'. This definition is usually used to describe foveal vision.

10. Field of view (FOV): FOV is the extent of the observable world that is seen at any given moment, while the head does not move but eyes can, and I assume in this work that peripheral as well as foveal vision is included in the FOV.

11. Physical field of view (PFOV): PFOV is the angle subtended from the eye to the left and right edges of the display screen.

12. Software field of view (SFOV): SFOV is defined as the angle subtended from the virtual camera to the left and right sides of the viewing frustum [31].

Appendix B.

Re-Configurability of V4SPACE

Even though I configured 4-SPACE to provide the best VA experience considering the design criteria for large displays in visual analytics, the system is also easily reconfigurable for a variety of use cases, design paradigms and user interactions. The bullet points below provide additional information on this re-configurability:

- Due to the ease of motion that the casters provide, the spatial arrangement of the displays in the grid can be easily rearranged, as needed. The system can even be “split” to create multiple smaller displays.

- The screens can rotate 360 degrees about a horizontal axis on the stands, i.e., be oriented horizontally, vertically or in any orientation while facing the user. A horizontal orientation, for example, provides over 13 meters (43 feet) of horizontal space, which could be used to create a closed circular display space for systems.

- The height of the displays relative to the floor can be changed by adjusting the height of the rotating mount.

- It is also possible to vertically tilt the screens. With this, the center of the display surface remains in position, but the angle can be adjusted to ensure the user views the displays as in a direction perpendicular to the surface. This helps to adjust if the user is not at the same elevation as the center of the screens, or to reduce glare effects.

- Optional touchscreens attached to the display surface support multi-touch-based interaction. Therefore, multiple users can interact with the display system simultaneously. One option with touch interaction is to create multiple virtual work spaces, which can be arranged dynamically.

- Optional 3D glasses support stereoscopic image display (currently only if the displays are oriented horizontally).

- As the number of screens and their orientation can be changed, and multiple types of interaction are supported, the potential use cases for V4-Space vary greatly.

Appendix C.

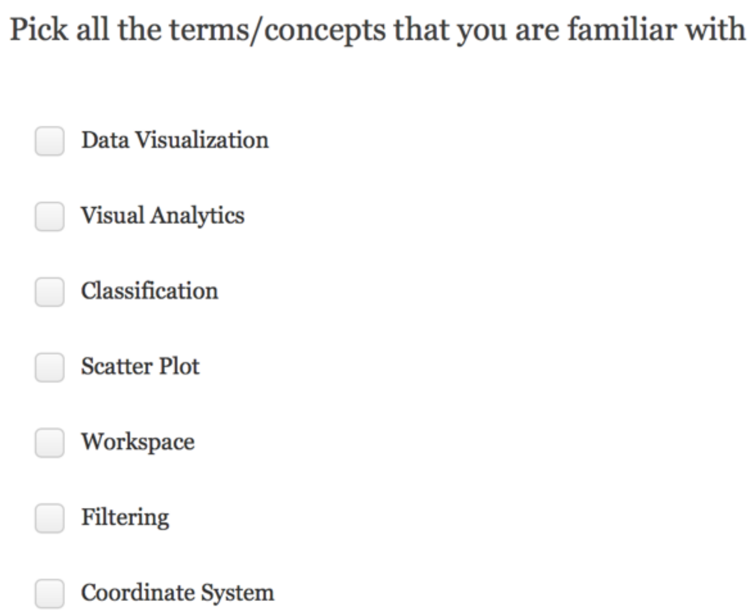
Study I Participation Motivation

Table A1. Participants and the motivation factor mentioned when asked why they had taken part in study I.

Motivation\Participant	P0	P1	P2	P3	P4	P5	P6	P7	P8
Participation reward (course credit, cash, etc.)		X	X	X					
Interest in the experiment tools (large displays)		X		X	X		X		
Interest in activity (Analytics, Visualization)				X		X	X		
Interest in dealing with data				X			X		
Other (please specify)	Helping a friend	-	-	Understand the research methodology and approach used.				As a favor	Volunteered out of the goodness of my heart :)

Appendix D.

Study I Concept Familiarity Survey


Pick all the terms/concepts that you are familiar with

- Data Visualization
- Visual Analytics
- Classification
- Scatter Plot
- Workspace
- Filtering
- Coordinate System

Figure A1. Participants were asked if they were familiar with given terms before the beginning of the study. They could pick multiple terms.

Table A2. The concepts that the participants of first study were familiar with.

Concept\Participant	P0	P1	P2	P3	P4	P5	P6	P7	P8
Data Visualization			X	X		X	X	X	
Visual Analytics	X		X	X		X	X		
Classification	X			X			X		X
Scatter Plot	X			X				X	X
Workspace	X	X	X	X	X	X	X		X
Filtering			X				X		X
Coordinate System		X		X					

Appendix E.

Study I Written Responses for Task (A)

Task (A) of Study I asks participants to arrange the charts into groups based on similarity and to explain the motivation for the way they grouped the charts. Participant responses were:

P0: I chose Y […] instead of X, because visually Y stands out.

P1: I [arranged] these charts as six groups, all dots of charts in group A (first line and third line) are near [the vertical axis], and in group B (second line), the dots in charts are near [the vertical axis] and also more crowded than group A. Most dots in Group C (4th, 5th line) are in the middle of chart. In the charts of Group D (6th,7th), most of dots are in the [right] of the chart. Group E have clear tendency and Group F are at the bottom. Because these charts already include age or land, I need to find a new way to group them.

P2: Some data of table is very concentrated. People may find rules from the table. But some data of table is very dispersed. I can't find data rules in those table. Some data just concentrate in the corner of the table. Some data points shapes are like lines in the table.

P3: I tried to organize the charts using their axis names. Some of them were about average, some about population and some about females (women). I did not care about the data inside them. I just grouped them by the English words I would see, mostly horizontal but sometimes vertical ones.

P4: I grouped them in that way because there was a lot of data and I consider this way will be easier to read it.

P5: I grouped things with something different versus something in common.

P6: I used horizontal axis texts to group charts. And in each group if dots are grouped close, I stretched [the] graphs' X or Y axis, so I can see better and makes them visible. And I created virtual grid in my mind which divides [the] screen [into] three parts horizontal. And I tried to list group elements in some random groups in different order. For example, all group elements are one line vertical or horizontal. The purpose of this is to break the order so I can focus. I tried to avoid screen frames, they make it hard to read but that was not a huge problem. One last thing I did was to try to collect groups to the right, because [the] left bar distracts me.

P7: I grouped the charts by the way the scatterplots looked in terms of correlation. The first group are all the plots where the points cluster by the y axis. The second group is where the points cluster around the origin. The third group is where all the points gathers around the lower half of the y axis. The fourth group is where all the points seem to have no correlation. The fifth group is where the points have a strong, linear correlation.

P8: When starting to organize the graphs, I first looked at the topics that were being compared. For example, above-65 age vs. women percentage. I would pick out the independent value (women percentage) and picked other graphs that fell within the first independent value group I chose. I chose retail as the first group to categorize as it stood out the most to me.

When I realized that some graphs don't fit in the "retail" group that I have originally categorized as, I moved the graphs around. Some groups (ex. poverty) were hard for me to categorize as I was torn between categorizing it with other groups.

Appendix F.

Study I Written Responses for Task (B)

After Task A in Study I, users were asked to use the highlighting facility to see if the highlighted data dimensions aligned with their groupings. The following are their responses.

P0: Data aligns with tab bars about 40%.

P1: I think those highlights will be very useful for grouping if I use [a] normal way to group [those] charts, but I do not use it, because my grouping way do not need that.

P2: When I see those table by each category. Some of them look like completely different. Some of them that we can figure out they are same category just by visually.

P3: I grouped average_income in one area quite well. I grouped population_count, population_change_perc and population_per_land well too. I grouped females_perc and firms_by_women_perc in one place. I grouped manufacturing fine. I also grouped merchandise_sales ok. But I can see above65age_perc, bachelor_degree_perc, below_poverty_perc, employment_change_perc, ESLs_perc, foreign_borns_perc, land_area, retails_sales and under18age_perc are gathered all around the large screen. Maybe because I did not manage to use them for my grouping. I used my grouping by dimensions that I can easily understand and [have] more meaning to me.

P4: The highlighted charts were matching per classification set.

P5: I did something vs common factor; so when you highlight something, that group second elements going to be highlighted.

P6: I already figured out at the beginning that property, so I grouped according to that.

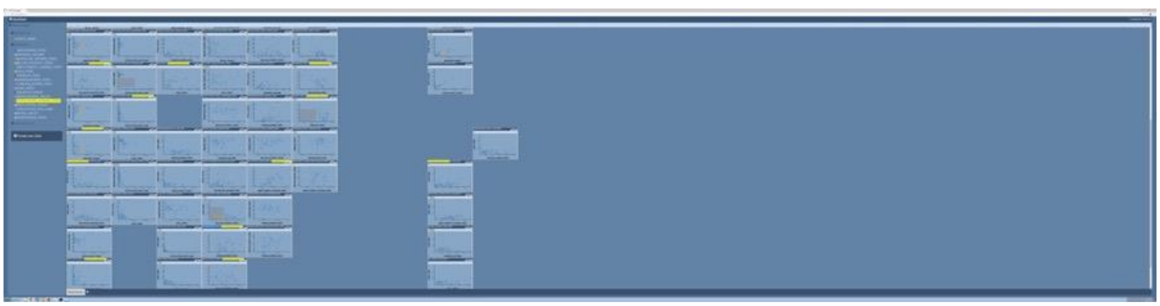
P7: Most of the dimensions when highlighted showed little correspondence to my groupings. The ones that stood out were Land area which showed many highlights in group 3 and Population per land which showed many highlights in group 2. I also noticed that all three scatterplots in group 5 had the dimension or Retail Sales, however the dimension of retail sales also showed up in other groups. Going through the dimensions and comparing them to my groupings did not reveal very much compelling information.

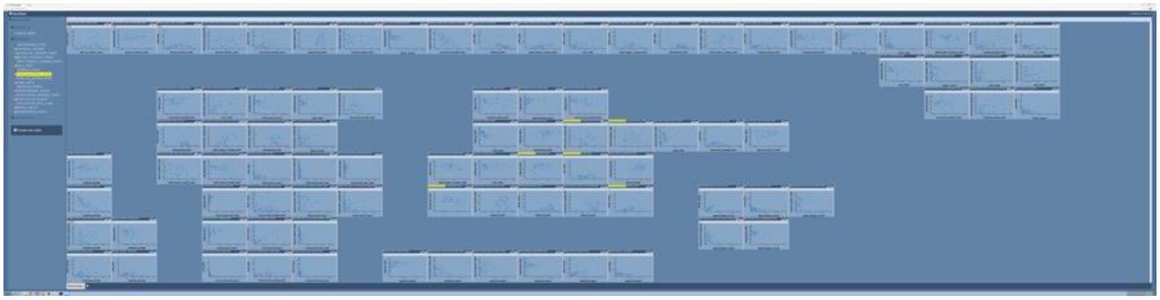
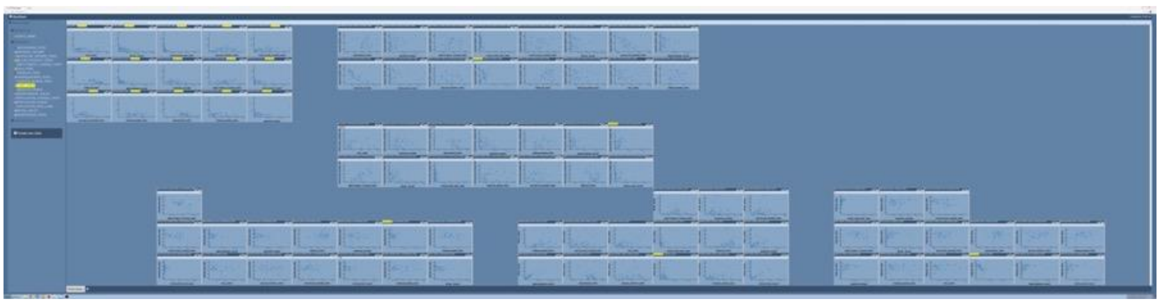
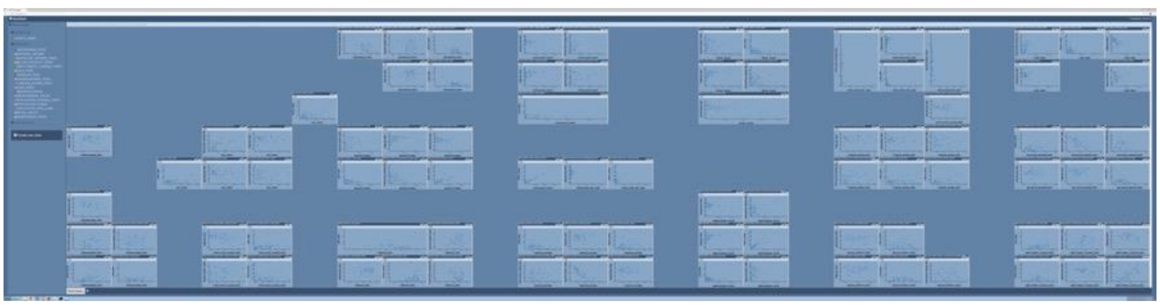
P8: From my categorizations of the graphs, some groups (ex. female, poverty) are widely spread apart, whereas some groups are categorized all into one group (ex. Bachelor’s degree, income). I realized that the areas highlighted are regardless of independent value or dependent value, therefore even if the categorization looks very scattered, the comparing values may not be scattered.

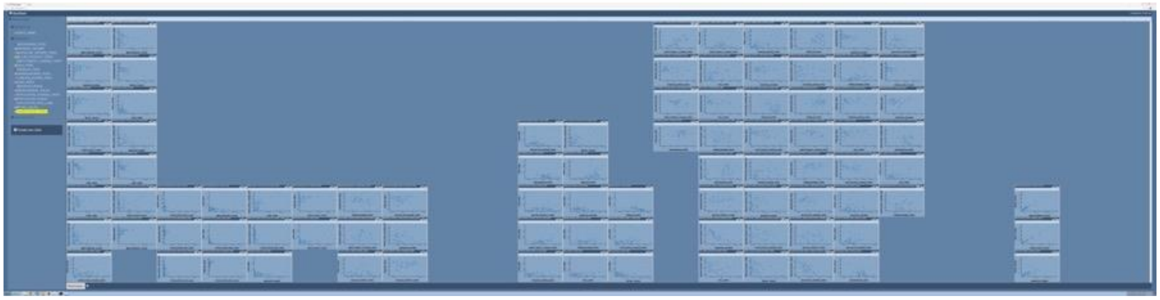
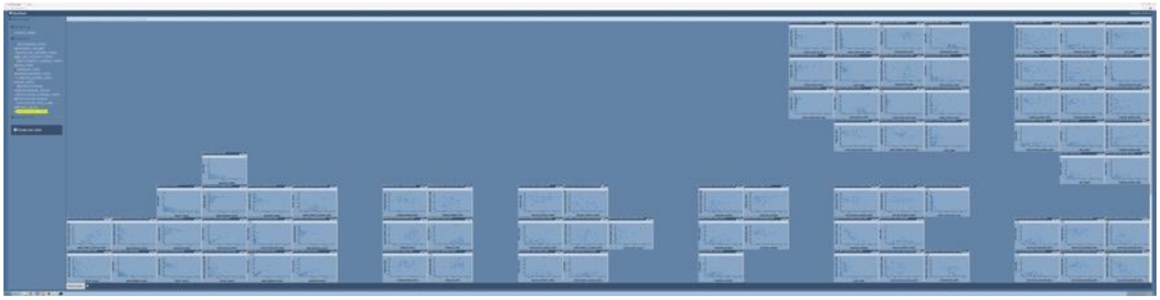
Appendix G.

Final Organization Layouts of Study I

The screenshots below show the final organization of charts for Task A of Study I.

P0	
P1	
P2	

P3	
P4	
P5	
P6	

P7	
P8	

Appendix H.

Study II Familiarity Survey

Table A3. Suryey of knowledge users have of VA concepts, whether they know any VA tools, and if they could name any VA software with they were experienced.

	How much knowledge do you have in Visual Analytics concepts such as bar chart, scatter plot, filter, dashboard, etc?	Do you know or use any Visual Analytics tools, such as Tableau? If yes, please write down the name of it / them.	
P0	I know all of them	Yes	Tableau
P1	I know some of them	No	-
P2	I know some of them	Yes	Tableau
P3	I know all of them and have practical experience	Yes	Tableau, matplotlib, R, spss
P4	I know some of them	Yes	JMP, Excel, SPSS, Minitab
P5	I know all of them	Yes	Tableau
P6	I know some of them	Yes	Tableau
P7	I know all of them and have practical experience	Yes	-
P8	I know some of them	No	-

Appendix I.

Study II Practice Session

Tutorial - How to explore your data

Thank you for completing the pre-study survey. Now we would like to give you a short tutorial about using the visual analytics tool. After that, you will be given a few tasks to practice what you learned in the tutorial.

Practice

	Yes	No
Can you create a new chart ?		
Can you resize a chart / move the chart around?		
Can you see how filter, sort, bookmark and close buttons work?		
Can you mark data dimensions to work with?		
Can you hover on a scatterplot and view details of points?		
Can you create a local filter by dragging and selecting points in a scatter plot ?		
Can you create a local filter by dragging a data dimension onto a chart ?		
Can you drag a filter from one chart to another ?		
Can you create a global filter ?		
Can you edit a filter ?		
Can you remove a filter ?		
Can you create a histogram?		
Can you add a bookmark?		

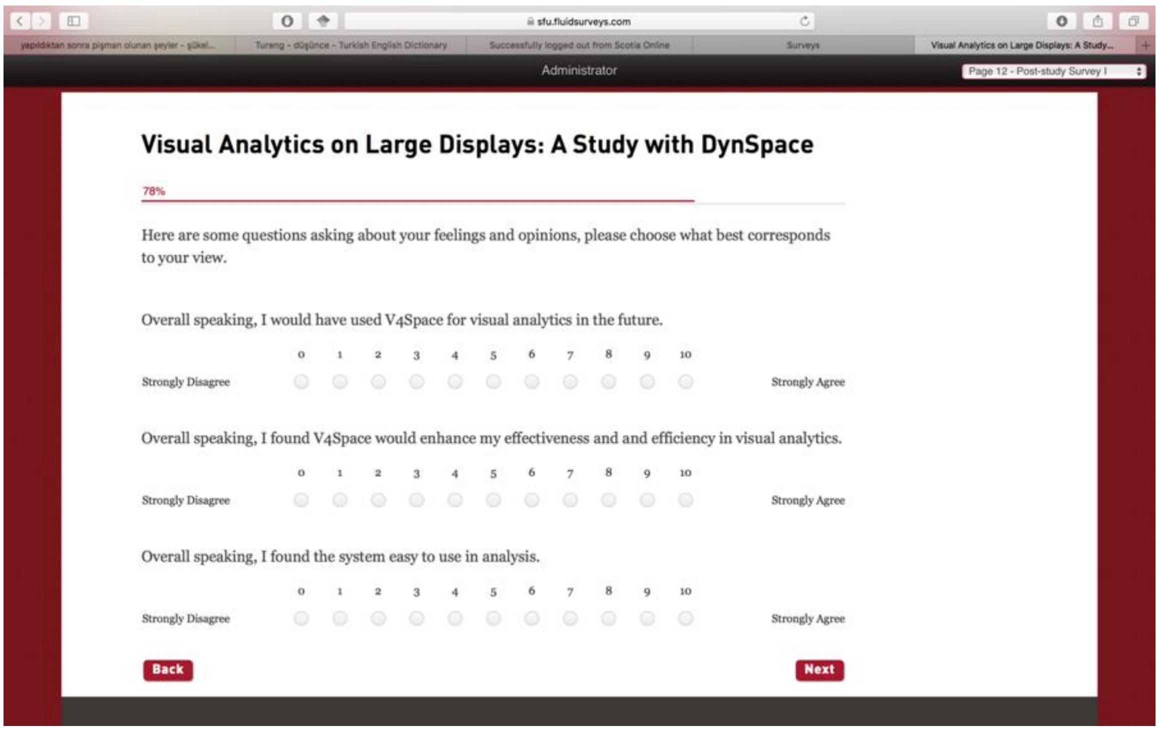
Figure A2. Practice session provided to users for Study II.

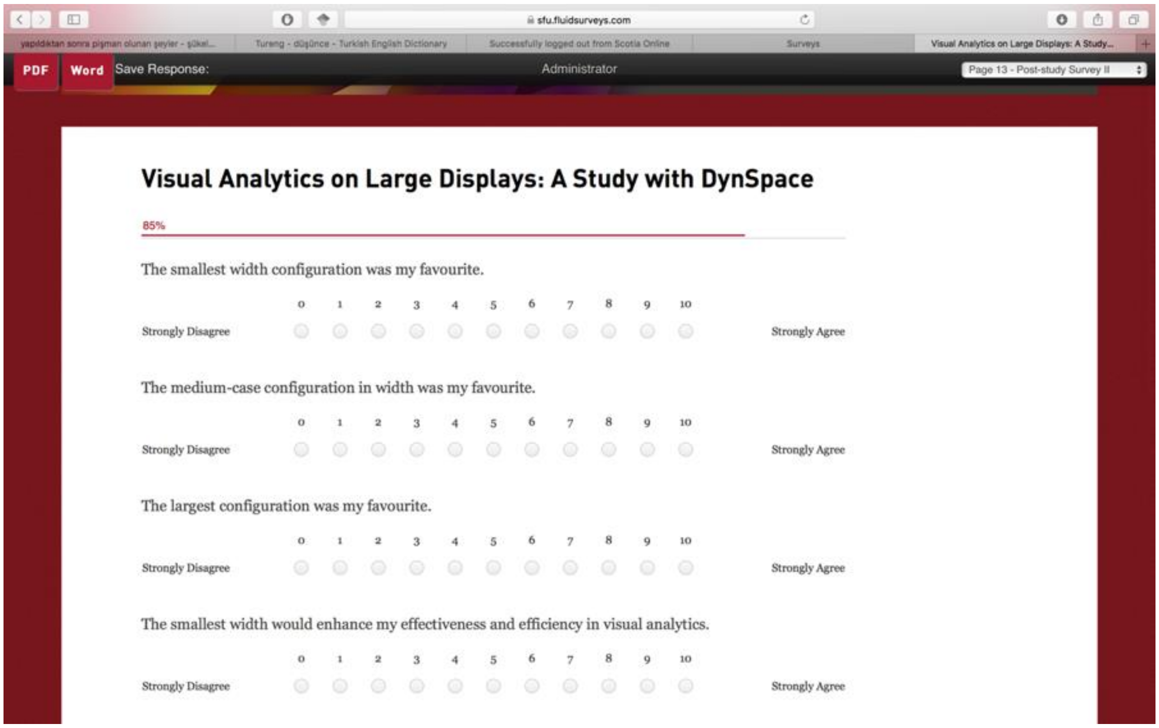
Appendix J.

Study II Post Study Survey and Interview

The first 3 screenshots below show the post study survey, asking Study II users about their experience and asking them to compare three cases of experiment: 3, 5 and 7 screen settings as the display surface.

The next two screenshots show the open ended post-study interview questions.





PDF Word Save Response: Administrator Page 13 - Post-study Survey II

Visual Analytics on Large Displays: A Study with DynSpace

85%

The smallest width configuration was my favourite.

Strongly Disagree 0 1 2 3 4 5 6 7 8 9 10 Strongly Agree

The medium-case configuration in width was my favourite.

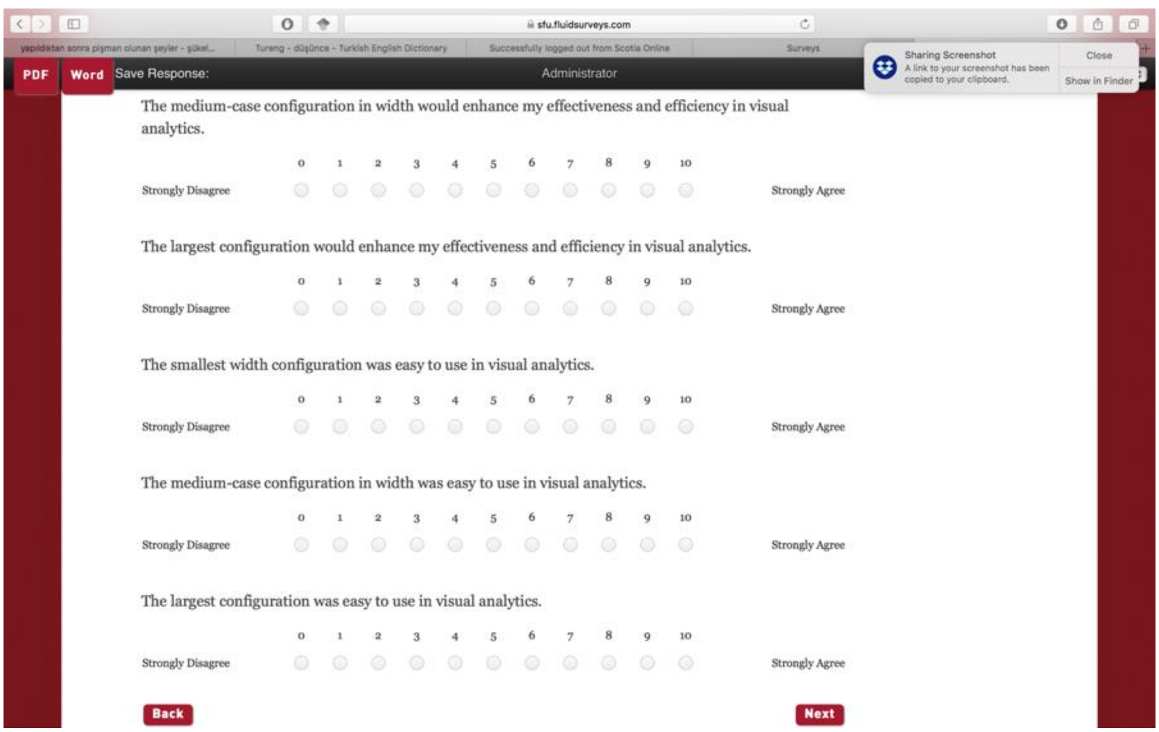
Strongly Disagree 0 1 2 3 4 5 6 7 8 9 10 Strongly Agree

The largest configuration was my favourite.

Strongly Disagree 0 1 2 3 4 5 6 7 8 9 10 Strongly Agree

The smallest width would enhance my effectiveness and efficiency in visual analytics.

Strongly Disagree 0 1 2 3 4 5 6 7 8 9 10 Strongly Agree



PDF Word Save Response: Administrator

Sharing Screenshot
A link to your screenshot has been copied to your clipboard.
Close
Show in Finder

The medium-case configuration in width would enhance my effectiveness and efficiency in visual analytics.

Strongly Disagree 0 1 2 3 4 5 6 7 8 9 10 Strongly Agree

The largest configuration would enhance my effectiveness and efficiency in visual analytics.

Strongly Disagree 0 1 2 3 4 5 6 7 8 9 10 Strongly Agree

The smallest width configuration was easy to use in visual analytics.

Strongly Disagree 0 1 2 3 4 5 6 7 8 9 10 Strongly Agree

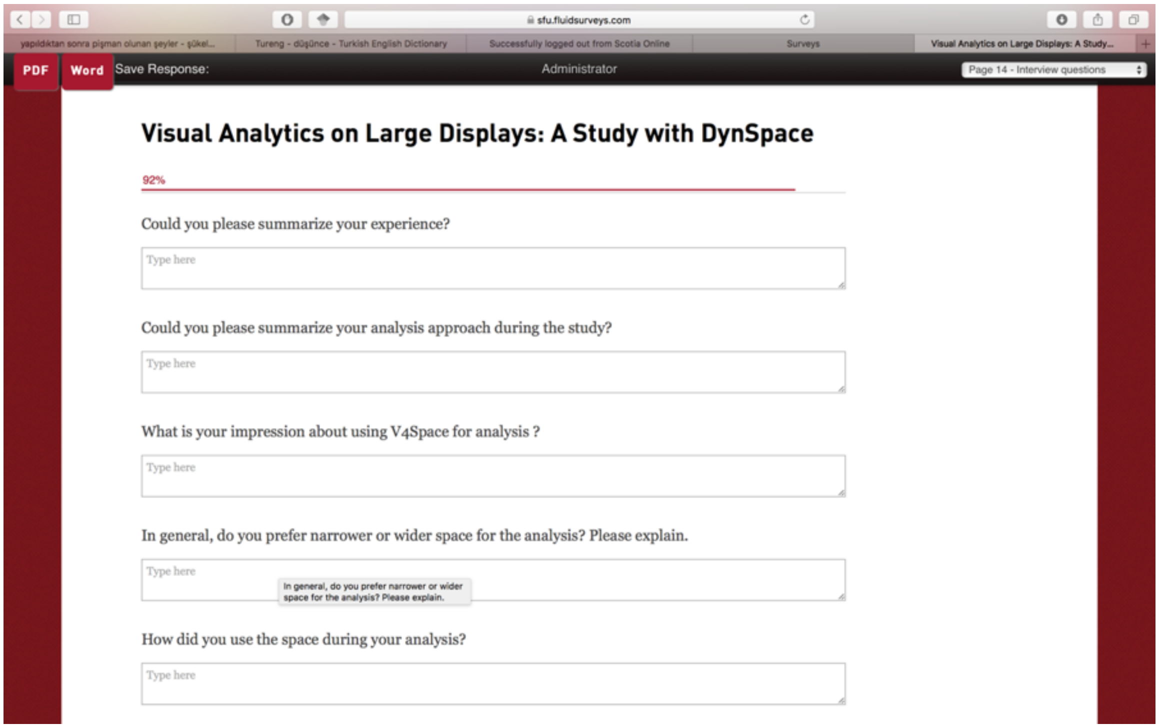
The medium-case configuration in width was easy to use in visual analytics.

Strongly Disagree 0 1 2 3 4 5 6 7 8 9 10 Strongly Agree

The largest configuration was easy to use in visual analytics.

Strongly Disagree 0 1 2 3 4 5 6 7 8 9 10 Strongly Agree

Back Next


PDF Word Save Response: Administrator Page 14 - Interview questions

Visual Analytics on Large Displays: A Study with DynSpace

92%

Could you please summarize your experience?

Type here

Could you please summarize your analysis approach during the study?

Type here

What is your impression about using V4Space for analysis ?

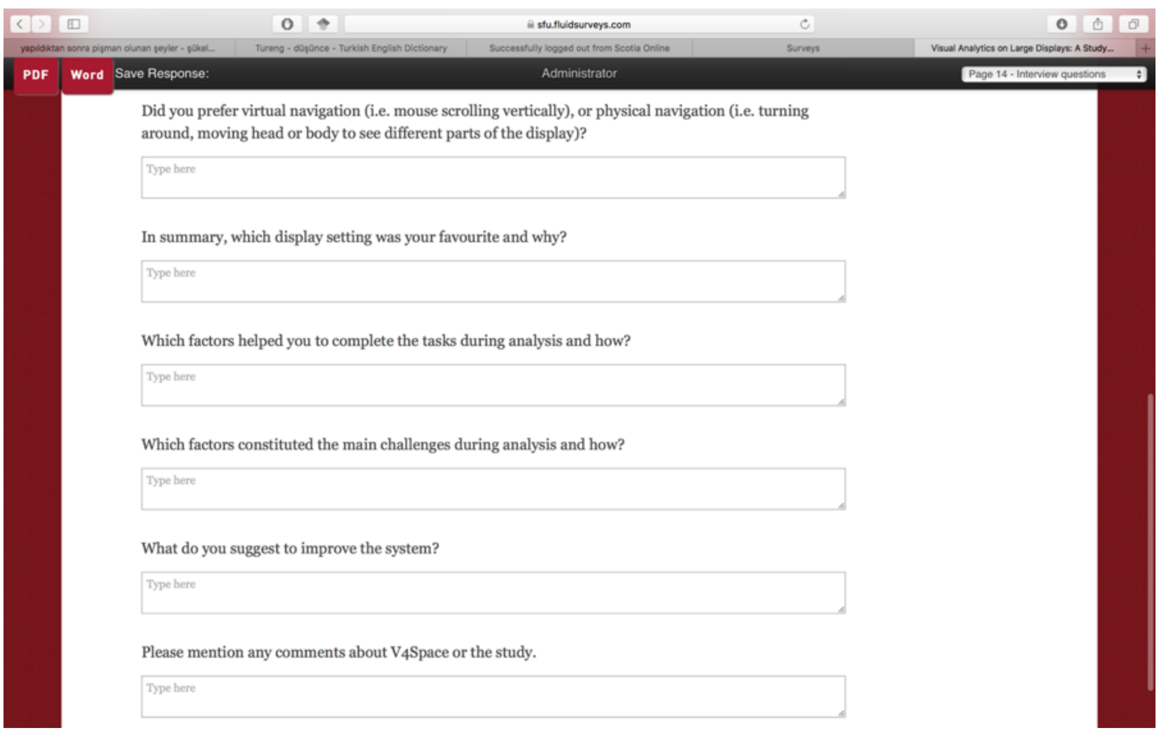
Type here

In general, do you prefer narrower or wider space for the analysis? Please explain.

Type here

How did you use the space during your analysis?

Type here


PDF Word Save Response: Administrator Page 14 - Interview questions

Did you prefer virtual navigation (i.e. mouse scrolling vertically), or physical navigation (i.e. turning around, moving head or body to see different parts of the display)?

Type here

In summary, which display setting was your favourite and why?

Type here

Which factors helped you to complete the tasks during analysis and how?

Type here

Which factors constituted the main challenges during analysis and how?

Type here

What do you suggest to improve the system?

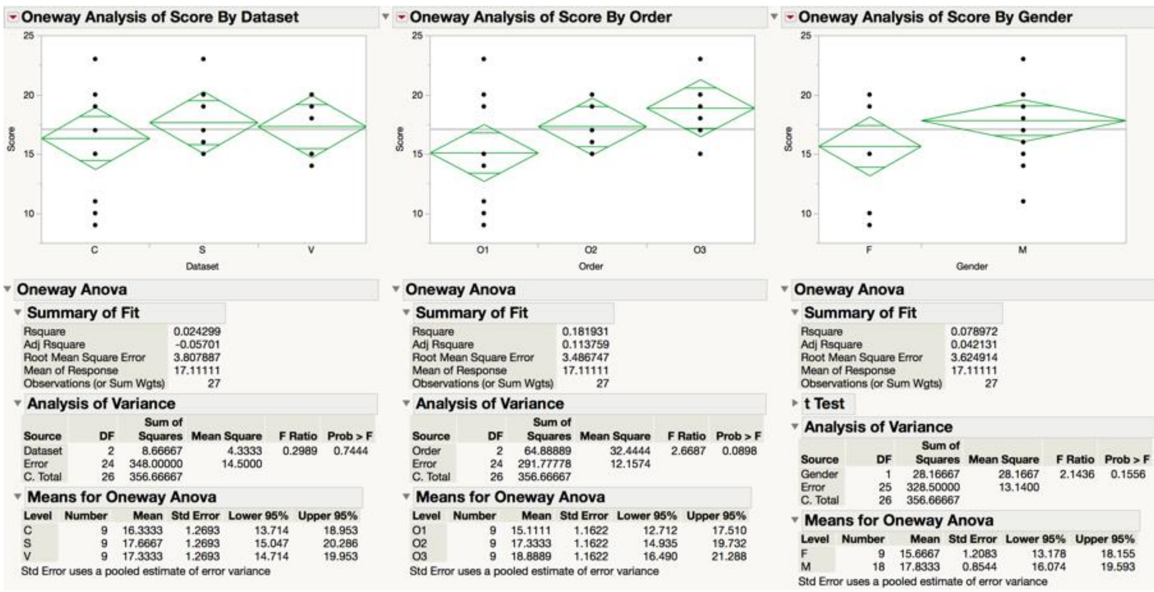
Type here

Please mention any comments about V4Space or the study.

Type here

Appendix K.

Study II Quantitative Analysis of Further Factors



Oneway Anova (Dataset)

Summary of Fit

Rsquare	0.024299
Adj Rsquare	-0.05701
Root Mean Square Error	3.807887
Mean of Response	17.11111
Observations (or Sum Wgts)	27

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Dataset	2	8.66667	4.3333	0.2989	0.7444
Error	24	348.00000	14.5000		
C. Total	26	356.66667			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
C	9	16.3333	1.2693	13.714	18.953
S	9	17.6667	1.2693	15.047	20.286
V	9	17.3333	1.2693	14.714	19.953

Std Error uses a pooled estimate of error variance

Oneway Anova (Order)

Summary of Fit

Rsquare	0.181931
Adj Rsquare	0.113759
Root Mean Square Error	3.486747
Mean of Response	17.11111
Observations (or Sum Wgts)	27

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Order	2	64.88889	32.4444	2.6687	0.0898
Error	24	291.77778	12.1574		
C. Total	26	356.66667			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
O1	9	15.1111	1.1622	12.712	17.510
O2	9	17.3333	1.1622	14.935	19.732
O3	9	18.8889	1.1622	16.490	21.288

Std Error uses a pooled estimate of error variance

Oneway Anova (Gender)

Summary of Fit

Rsquare	0.078972
Adj Rsquare	0.042131
Root Mean Square Error	3.624914
Mean of Response	17.11111
Observations (or Sum Wgts)	27

t Test

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Gender	1	28.16667	28.1667	2.1436	0.1556
Error	25	328.50000	13.1400		
C. Total	26	356.66667			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
F	9	15.6667	1.2083	13.178	18.155
M	18	17.8333	0.8544	16.074	19.593

Std Error uses a pooled estimate of error variance

Figure A3. **Analysis of dataset, order, and gender's impact on quantitative task measure (overall score).**

The results for each condition are fairly similar to each other, in terms of characteristics and complexity, as can be seen from the results shown on the left. In the middle, it is seen that subjects perform better in their later tasks, as they learned and gained experience VA during the experiment. Therefore, I can verify that counter-balancing the order helped me obtain more reliable results. Finally, on the right, 6 male participants performed slightly better than the 3 females, however this is not a statistically significant difference in terms of scores.

Appendix L.

Study II Task Success by Total Number of Pixels