An Exploratory Study of How Abundant Display Space May Support Data Analysis

Søren Knudsen, Mikkel Rønne Jakobsen & Kasper Hornbæk
Department of Computer Science, University of Copenhagen, Denmark
sknudsen@diku.dk, mikkelrj@diku.dk, kash@diku.dk

ABSTRACT
Large, high-resolution displays offer new opportunities for visualizing and interacting with data. However, interaction techniques for such displays mostly support window manipulation and pointing, ignoring many activities involved in data analysis. We report on 11 workshops with data analysts from various fields, including artistic photography, phone log analysis, astrophysics, and health care policy. Analysts were asked to walk through recent tasks using actual data on a large whiteboard, imagining it to be a large display. From the resulting comments and a video analysis of behavior in the workshops, we generate ideas for new interaction techniques for large displays. These ideas include supporting sequences of visualizations with backtracking and fluid exploration of alternatives; using distance to the display to change visualizations; and fixing variables and data sets on the display or relative to the user.

Author Keywords
Large high-resolution displays, interaction techniques, user study, workshop, visualization.

ACM Classification Keywords
H.5.2 [Information interfaces and presentation]: User Interfaces—Graphical user interfaces (GUI).

General Terms
Human Factors.

INTRODUCTION
Large, high-resolution displays are becoming ubiquitous, with size and resolution increasing at impressive speeds. Displays now offer sizes well over 100 megapixels [2], resolutions over 100 DPI [24], and more stable and fine-grained support for multi-touch (e.g., Microsoft Surface 2.0). Research has shown that such displays improve performance and user satisfaction [12,33].

An additional hope for large, high-resolution displays is that they support data analysis by giving “space to think” [1]. We use data analysis in a broad sense to denote gathering, organizing, reading, extracting, visualizing, checking, and narrating data; we see it related to sensemaking [26] as well as to the types of activity supported in visual analytics [35]. The contention here is that large, high-resolution displays may fundamentally change how data analysis is done by affording new opportunities for visualizing and interacting with data.

Much research has dealt with how users can interact with large displays, proposing and evaluating techniques for pointing [6], gestures [22,36], text input [29], and using physical movement as a navigation aid [2]. Such techniques are typically generic and support data analysis only indirectly by facilitating input. Less work has been done on supporting complex analysis, though some papers discuss how to support sensemaking [1] and collaboration on large displays [9]. Studies such as [1,37] have helped understand how single or multiple users benefit from large displays in analysis tasks in a particular domain. However, they rarely identify new visualization or interaction techniques for using space to think.

Although recent work has helped understand complex analysis tasks with large displays, we know little about how to support analysis beyond efficient pointing and window manipulation techniques. It is unclear how abundant display space can support data analysis tasks in general. Moreover, we lack visualization and interaction techniques that help users benefit from large displays when analyzing large amounts of data. This raises several questions: How may large displays support what-if analysis? How may abundant display space be used to reason about alternatives? Can we come up with interaction techniques that support analysts in hypotheses testing?

The present paper tries to answer these questions by taking a complementary approach to existing studies [e.g., 1]. We conduct workshops that focus on analysis activities and how they may be supported on large displays. Workshop participants redo analysis tasks from their work using a simulated large display, mocked up by whiteboards and various paper representations of data. As participants redo tasks, we probe them with questions on how to do their analysis given the large display. Workshop participants are sampled from diverse domains and different types of data analysis so as to maximize variation and to attempt generalization. We analyze video recordings of the workshops in detail using a grounded theory approach [31].

Based on this analysis, we report findings across domains and present a catalogue of ideas from the workshops.

Our aim with this work is to generate new directions for researchers and practitioners on how to design for large displays in order to make abundant display space work in analysis tasks. The paper makes three contributions:

- An analysis of 11 workshops spanning domains as different as artistic photography, phone log analysis, and health care policy.
- A set of ideas for making use of large, high-resolution displays for data analysis.
- A workshop method for working with concrete tasks using imaginary technology (in our case, a large, high-resolution display).

RELATED WORK

Much work has investigated the use of large displays both for single-person use [12] and for collaboration [18]. Early examples include iLand [32] and Liveboard [13], which focused on office work and face-to-face meetings. Large displays have been shown to improve users' performance and satisfaction in a variety of tasks [2,4,6,8,12]. Increasing display space helps view multiple windows with less navigation [12], improves task switching [3], enhances awareness of peripheral applications [8,16], gives a better peripheral view [7], and may promote physical navigation [2,41]. Even with the view as a normal-sized display, large displays may increase performance in spatial tasks [33].

The present study focuses on data analysis in a broad sense, taking the phrase to denote gathering, organizing, reading, extracting, visualizing, checking, and narrating data. This sense includes the types of activity supported in visual analytics [35] and listed in taxonomies of information visualization [40]. The focus on data analysis differs from many of the studies mentioned in the previous paragraph. They have solved usability problems in interacting with big screens, problems of reaching over a distance, and so forth, and to a lesser degree concerned analysis tasks.

In contrast, we focus on how an abundance of space by way of large, high-resolution displays may support data analysis. For instance, increasing display space may allow analysts to view more data at a time or to organize data spatially as appropriate for their work. Few empirical studies help understand these benefits for specific types of analysis. Andrews et al. [1] described how intelligence analysts benefit from large displays particularly for sensemaking, which is a common analysis activity [26]. Andrews et al. argued that a large, high-resolution display fundamentally changes analysis tasks compared to smaller display sizes.

Isenberg and colleagues [19] studied how visual representations are used in analysis. They had individuals, pairs, and triples work on data sets from SPSS; tasks comprised open discovery tasks and more focused tasks with one correct answer. From coding of videos they derived a description of the analysis process involved in solving the tasks. The conclusions with respect to interaction and visualization design, however, mostly concern the benefit of process-free tools and the drawbacks of implementing a strict structure in tools for supporting analysis. Robinson [25] report on a similar study of how pairs of experts in geography and infectious diseases synthesize collections of analysis artifacts. Robinson noted that collaboration style and organizational strategy varied between pairs even though pairs had similar backgrounds. Ziemkiewicz et al. [42] presented a case study of the use of immunobiology visualizations. They collected videos and screen captures to analyze how visualizations were used and conducted interviews with four researchers that had used the tools. Thereby Ziemkiewicz et al. identified distinct ways of using the visualization, which varied greatly among individuals.

The above work mainly concerns understanding the use of visualizations. While such work help design for visualizations, few studies have directly attempted to identify and propose new ways of interaction and new visualization techniques that work for large displays. This is the motivation for the present study, where we elicit ideas for supporting data analysis with large displays.

In addition to these considerations about large displays, we also briefly want to discuss work that relate to our choice of method. The literature shows several ways of eliciting design ideas from users when the goal is technology innovation [34,39]. The main goal of the present paper is to use workshops to elicit ideas. We draw on participatory design work on conducting workshops, in particular on the inspiration card workshops [17]. In the workshops we use whiteboards as a proxy for large, high-resolution displays. Several papers on visualization and interaction have concerned whiteboard use [10,38]. For instance, Walny et al. [38] analyzed snapshots of whiteboards, created by 69 researchers. They showed how whiteboards contained complex visualizations, using a variety of types of representations and linking. Their study provides an argument for using whiteboards to simulate large displays; next we describe how we do so in the workshops.

METHOD

The question guiding the study is: How would professionals do data analysis tasks on wall-sized interactive displays? To better understand this, we conducted workshops with 11 groups of 2 to 3 analysts from a variety of domains. We chose to conduct a workshop study because we wanted to observe real, hands-on analysis work, carried out on what participants would think of as a large interactive display. The key part of the workshop is to have participants imagine a whiteboard to be a large, high-resolution display and redo tasks on the imaginary display.

We argue that this approach offers several benefits. First, this approach is more general than individual studies of data analysis. Second, this approach is grounded in concrete data analysis tasks, rather than trying to develop general models of analysis activity and derive design implications from them. Third, this approach may offer a sweet spot between contextual studies and generalizability.

Participants

Eleven groups of professional analysts agreed to participate in the study. The groups were recruited from research and business domains confronted with a need to collect, analyze, understand, and act on large amounts of data. Table 1 provides a summary of the groups; their names replaced by the letters A through K and group size indicated as #. Participants were invited in small groups so as to facilitate discussion and to help each other make the leap of faith in simulating that the whiteboard was a large display.

Our sample comprises four (E, G, H, and I) scientific research groups that analyze large data sets. A main objective of their analysis work is to report results to scientific communities. Three groups (B, D, and F) are part of organizations that analyze business data on customers, production, or accounting; they disseminate their analysis results to internal and external stakeholders. Three groups (A, C, and K) belong to organizations concerned with analyzing data about the general population; they disseminate results publicly. Lastly, one group (J) does artistic photography and shows it in media and art exhibitions. The aim of this variety of domains is to attempt more general conclusions than if we did an in-depth study of one domain. We return to the pros and cons of this variety in the Discussion.

Workshop preparation: Interviews, Tasks, Data

To prepare for each workshop, we interviewed one person from each group of participants. The purpose of the interview was to understand the domain of work and to identify tasks for the subsequent workshop (see Table 1). We asked open-ended questions about the data the groups use and the analysis tasks they perform. We requested that tasks and data to be used in the workshop were based on actual analyzes that the interviewee had recently been doing. Some persons were interviewed two times to clarify the domain and find useful tasks. We also identified data in raw and various processed forms that would be used during the workshop to remind participants of their work and generate ideas. The interviews also helped identify co-workers that would be part of the workshop.

For each interview, we identified up to five analysis tasks that would form the focus in the workshops (see Table 1, second rightmost column). A total of 23 tasks were collected: for two groups, analysis tasks were not fixed before the workshop; while one group had five tasks described. Tasks could for example be: How does use of the website relate to country of visitor (workshop B), how are galaxy image features related to galaxy properties

	Participant characteristics			Materials used in the workshops	
	#	Domain	Type and magnitude of analysis data	Tasks	Representations of data
A	3	Health care policy (Public)	Data on 1m (million) annual admissions to Danish hospitals.	Understand errors in computing costs of hip replacement surgery based on activity information from hospitals.	3 sheets of tabular data and 3 sheets of histograms covering a subgroup of hip replacement surgery.
B	2	Website analysis (Business)	Logs of 2m annual visits to an international corp. website.	Understand how use of the website relates to country of visitor and means of access.	89 printouts of reports from Google Analytics based on website in question.
C	3	Health care policy (Public)	Financial and operations data on 1m annual admissions to Danish hospitals.	Compute costs of births with and without epidural block and understand how changes in configuration of financial accounts influence diagnose group costs.	2 sheets of aggregated costs of patients, grouped by disease category; 14 births split on hospitals and 28 sheets with financial accounts of a specific hospital.
D	2	Phone log analysis (Business)	Logs of 5k (thousand) users' smartphone activity.	Understand how separate subscriber segments use smartphones during a day.	Sketched individual and aggregate data over time for particular segments.
E	2	Astrophysics (Research)	Raw and processed images of 1m galaxies.	Understand relation between image features and properties of galaxies.	Raw and processed images of galaxies in 3 different sizes.
F	3	Logistics (Business)	Positioning information of 10k containers on shipping vessels.	Stow containers into partially loaded vessel at current port minding stability, stresses of vessel and optimal ballast use.	14 sheets of user interface from an actual product used for analyzing loads of containers on shipping vessels.
G	2	Internet game statistics (Research)	Logs of 1m internet game users in-game activity.	How are communicational patterns defined and how do they relate to player age, leveling, and number of players?	20 sheets of: a tabular overview of database tables, a box and whisker plot, 2 scatter plots, and 3 bar charts.
H	2	Information retrieval (Research)	Mapping of 30k rare diseases to 120k medical concepts.	Understand relation between mappings; why these results and why poor/no match.	20 sheets of tabular data describing input and from a semantic mapping tool.
I	3	Information retrieval (Research)	Results of 1k queries to an IR system based on 1m documents.	Gain overview of different IR scores and their relation considering the queries.	3 sheets of tabular data of query results for a rare diseases search engine and aggregates based on 27 IR metrics.
J	2	Artistic photography (Arts)	100k photographs of people in the street.	Sort photographs in categories, construct new categories, select exhibition photographs and design exhibition layout.	100 photograph sheets covering 5 different categories, as well as 5 contact sheets with miniature photos.
K	2	EU air emission statistics (Public)	Statistical reports from multiple public sources.	Find and extract relevant information and analyze sources to understand trends.	8 sheets of paper with data describing air pollution in the EU.

Table 1. Characteristics of participants' domains and data analysis tasks. Numbers of workshop participants are indicated as #.

(workshop E), and how are photographs sorted into meaningful categories (workshop J).

Each interview also resulted in some representation of data to be used during the workshop. We collected 452 sheets of paper containing tabular data, histograms, scatter plots, bar charts, photographs, images, feature images, line charts, geographical maps, and user interface components showing data to use during workshops. Participants brought these sheets of paper to the workshops; in most workshops, additional representations of data were produced during the discussion.

Conducting the workshops

We conducted one workshop for each group of participants; workshops lasted up to two hours (on average 92 minutes). The workshops were held in a meeting room, accommodating up to 20 persons, equipped with a whiteboard of 6 meters by 1.3 meters. We had post-its and whiteboard markers (in 4 different colors), magnets and magic tape available as well as the data printouts that workshop participants brought along.

Each workshop began with introducing participants and facilitators, and explaining the agenda for the workshop. We explained participants the tools that were available.

For each of the tasks identified, we asked the interviewee to walk through the task, the associated data, and the conclusions reached. While doing so they were told to imagine that the whiteboard was a high-resolution display. Next, we encouraged the other participants to discuss how to do the task, how to interpret the data, and to discuss the findings – while reminding the participants that they should use the imaginary display to support their discussion. Figure 1 shows a typical workshop situation: Here, participants discussed how related data could be used in relation to their main data.

When this discussion had lasted about 10 minutes or had dried out, we probed participants with questions in relation to their discussion. The questions come from three sources:

- Information Visualization taxonomies [11,28,40].
- The possibilities enabled by large displays and how participants would use them.
- The tasks brought to the workshop.

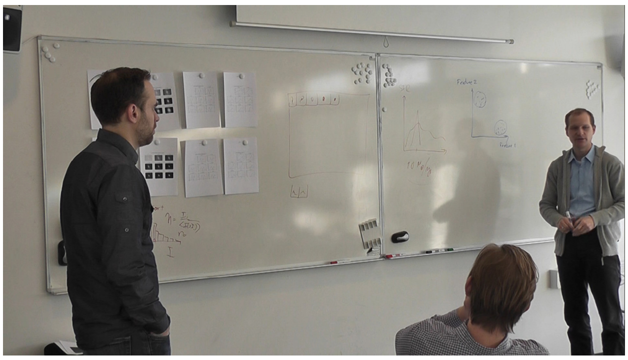

Figure 1. Typical situation in a workshop (workshop F).

When asking questions, we framed or explained them in light of the discussion to ensure participants would understand our questions. For example, we asked "How would you want this shown so as to be able to compare it to the other example?", "Would you prefer to have both a visual representation and a table?", and "How would you use the entire whiteboard to support this task?"

Data Collection and Analysis

Our data comprise notes from the interviews and workshops, data, analysis tasks, and video recordings gathered during the workshops.

We recorded each workshop using two video cameras, each viewing the whiteboard from a different angle. Videos were in 16:9 HD format so as to enable us to observe gestures, pointing, body language, and movement, and were merged into 32:9 video files to be able to easily switch between angles.

Initially, workshops A and B were transcribed and coded by one analyst, both to describe interesting themes to pursue in following workshops and to develop codes. After coding these workshops, we also conducted a collaborative coding session. We looked for themes, topics, and issues related to abundant display space, although other interesting observations were kept as well.

After having conducted all 11 workshops, one analyst coded the remaining workshops. The codes were developed further during this second pass and codes describing activity and general behavior in the workshops were added; we also added codes describing the phase in the workshop (intro, task intro, task discussion, task roundup, workshop summary and pause), interaction on the whiteboard (writing, placing paper, moving paper), gestures (on-screen, in front of screen, in-air), and movement (stepping back, approaching). Following this pass, codes with low coverage were revisited; if we were able to call up instances of these codes from memory, we added them – otherwise, the codes were left out of subsequent analyses.

In the third and final pass, we held short collaborative discussion sessions in which workshop observations were discussed. This resulted in identification of six themes that one analyst related to the coding. For each theme, we identified codes from the second pass that related to these and coded the themes on these; we used axial coding [31] to develop codes further.

RESULTS

The following section first gives an overview of what happened in workshops. Then follows six themes developed during the third analysis pass (see above). The themes concern (1) persistency, (2) showing data side-by-side or one-by-one, (3) space to spread out data, (4) trail of thoughts, (5) movement, and (6) gestures.

Overview of workshop activity

We began each workshop with an introduction (4min). The average time in minutes used for each task broken down in phases was: task setup and introduction (9m), task discussion (30m), and task roundup (3m). Tasks varied much across workshops (see Table 1), but did contain common types of analysis such as comparison between sets of data (10 workshops). Another example was discussing overviewing (all 11 workshops), in some instances in relation to obtaining an overview and in some instances to losing the overview. We ended the workshops in an open dialogue and thanked the analysts for participating (4m).

Participants brought data from their analysis domain to the workshops and used them in various ways on the whiteboard (e.g., attaching them using magnets). In addition, participants drew sketches of user interfaces and different representations of data. Some common types include histograms (used in 6 workshops), tables (9 workshops), and plots (8 workshops); see Figure 2.

Annotations on the whiteboard were coded 20.9 times on average, varying from 0 to 38 instances between workshops, while annotations on sheets of paper were coded 1.3 times per workshop. Placing the paper sheets were coded 6.3 times on average, while moving papers was coded 6.8 times per workshop, varying from 0 to 25 instances between workshops, indicating that in some workshops paper was not used at all.

Participants were actively engaged in discussion during most of the time in the workshops. We saw few pauses in speech lasting more than a couple of seconds. Most participants were gesticulating while speaking. Most gesticulations supported communication between participants and facilitators, yet we coded 172 gestures relating to interaction with the imaginary display.

In all workshops, participants moved along the whiteboard, and closer to or farther from the whiteboard. In 6 workshops (A, C, D, F, G, I), only one participant was active in front of the whiteboard at a time, whereas in the other 5 workshops (B, E, H, J, K) participants shared the whiteboard fluidly. When one participant was active, other participants would sit, but keep engaged in the discussion. We identified 3 typical positions in relation to the display: (a) interacting or looking at the display, (b) interacting with other participants with the back to the display, and (c) away from the display facing it.

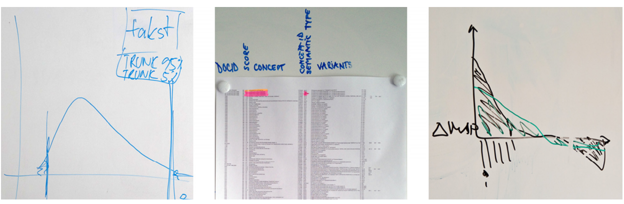

Figure 2. Frequent types of representations used: Histogram (left, workshop A), table (middle, H), and plot (right, I).

Persistency

The most frequent use of abundant space we call persistency: partitioning the display space so that designated areas have a particular purpose in support of analysis throughout a task. Participants' idea behind this usage seems to be that when display space is abundant, one may use more of it to show data for longer periods of time. Persistency was seen in 6 workshops (D, E, F, G, H, J) where participants fixed key variables, data sets, or views to particular areas.

A typical example of persistency was seen in a workshop where participants worked with analyzing how cellphone subscribers use smartphones. In that workshop, an interface was sketched during the workshop (D: 32:30-36:40, see Figure 3 top-left). The top part of the display was reserved for a dimension layer displaying simple data representations (e.g., histograms) of variables preselected among all variables in the system (the examples given were gender, age, smartphone model, questionnaire answers), which could be used to modify data representations in a working area in the central region of the display. Participants also imagined the bottom display area designated for showing a fixed set of groups of data (D: 46:10-46:40).

While most instances of persistency concerned fixed display areas, we saw 2 instances suggesting a need for participants to define persistency relative to their position. In workshop F participants worked with allocating containers onto sections of a ship. They talked about seeing sections of an entire ship in front of them and having related information such as stability metrics and overview of ports placed persistently around this view. Participants went on to imagine the entire ship spread out over the display and having the related information available in their horizontal periphery (F: 36:30-37:00). Having this information fixed in their periphery would enable them to focus on a particular section of the ship while still being able to glance at the important information from time to time. In Figure 3 (top-middle) a participant is gesturing how these views would be positioned.

In the above example, we described variants of persistency pertaining to seeing an overview of the ship and a detailed view of information. We saw instances of persistent overviews in 5 workshops (D, E, F, G, and J) and of persistent detail views in 3 workshops (F, H, and J).

Persistency was talked about or used with raw data, variables, groups of data, and aggregate/calculated information. Recall the example above from workshop D where areas were designated to hold specific variables and groups of data. Likewise, the example above from workshop F involved detailed information. An example of raw data was seen in workshop G (G: 1:21:12-1:21:22), where participants imagined using an area for raw data that could be selected and moved to a more active area for analysis.

Showing data side-by-side or one-by-one

We saw two distinct approaches to how participants worked with multiple representations of data. In one approach, two or more representations of data were used side-by-side. We saw this approach in 10 workshops (all except J). In the other approach, a single representation was changed by interaction, showing data representations one-by-one. We saw this approach in 7 workshops (A, D, E, G, H, I, J).

A typical example of using representations side-by-side was seen in workshop C (39:00-45:20), where participants tasked with understanding cost structures in Danish hospitals analyzed patients with related sub-diagnoses and where they were admitted. Participants used a stacked bar plot showing proportions between individual hospitals. Clicking on a specific bar opened a pie chart next to the other visual representation showing diagnose broken down into procedure codes (see Figure 3, bottom-left). Participants went on to discuss seeing histograms and averages of individual slices of the pie – for example showing distribution over age, admission time, or gender. In this style, representations of data unfold over a series of interaction steps, forming a tree-like path of interactions.

A typical example of using a one-by-one approach was seen in workshop D (46:20-51:35) where participants who worked with smartphone usage logs imagined a middle working area showing a data plot of smartphone usage averaged over a 24-hour period. They wanted to drag variables onto this data plot and thereby let the variables act as filters for the data shown. For example drag the segment *20-29 years* of the variable *age* onto the data plot thereby filtering on this criterion. This is illustrated in Figure 3 (top-left). The boxes in the top of the figure represent variables which can be dragged down onto the graph in the center of the figure to filter the data.

The examples above concern drilling-down in data by filtering on variables. We also saw the approaches of side-by-side and one-by-one used when comparing groups of data. In workshop E (11:25-12:05), for example, participants looked at original grayscale images, processed images, and image feature-plots of two galaxies to compare and understand how visual properties of galaxies were represented in the plots of image features. This configuration of data is shown in Figure 3 (bottom-middle).

The two approaches represent a tradeoff between use of space and interaction. Although space is preferred for many purposes, interaction over time is nevertheless preferred in some situations. For instance, in workshop I (75:15-75:35), participants compared sets of data by flicking back and forth between them. They started by defining what data to compare using checkboxes. Then they talked about viewing data one-by-one: *You could perhaps define two views that are [in] the same space and then say; well can I have one or the other, one or the other* [said while doing a flicking gesture and looking at the data]. They did this to understand

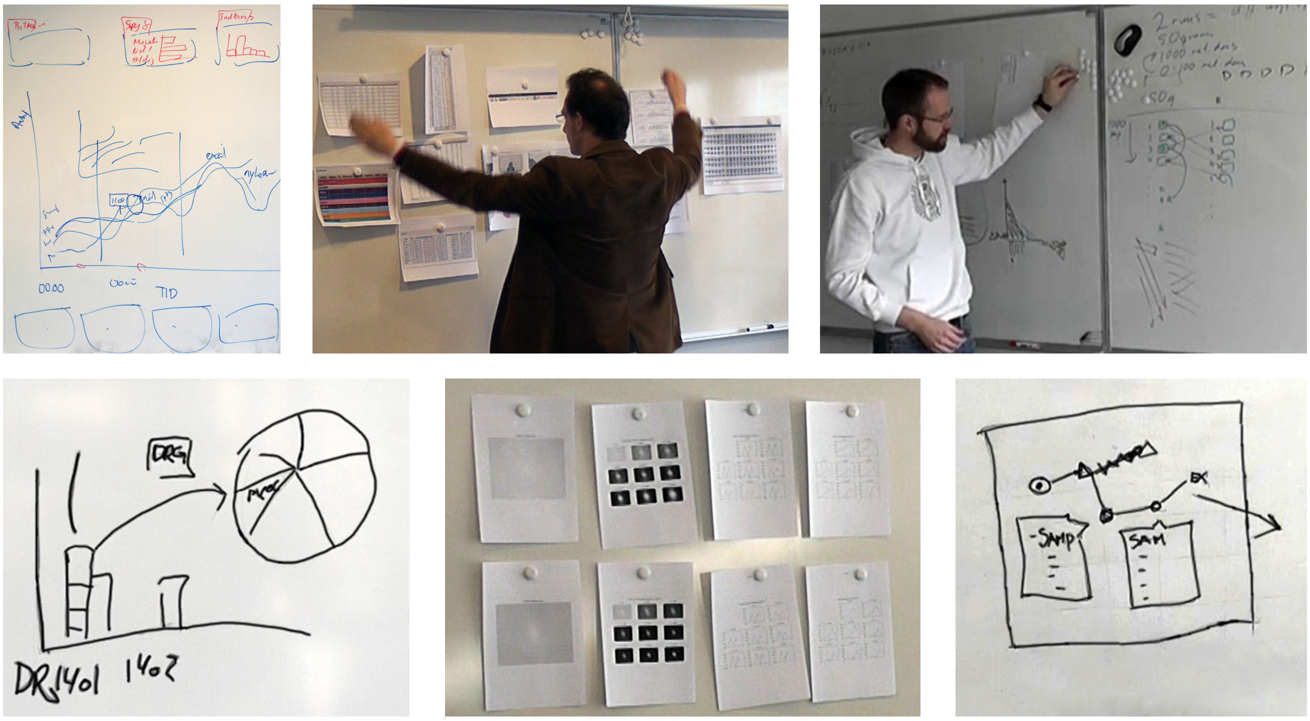

Figure 3. Top-left: Analyzing cellphone subscriber behavior on smartphones. Top-middle: A participant show how information views would be positioned in a user's peripheral view. Top-right: A participant uses a magnet to illustrate a flicking gesture. Bottom-left: Tree of plots. Bottom-middle: gray-scale images, processed images and image feature-plots of two galaxies. Bottom-right: Representation of data processing flow.

the difference between the two views: ... *to see visually, to swap between [the views] and [see] what happens actually* [flicking gesture]. They preferred this rather than having sets of data shown next to each other. This situation is illustrated in Figure 3 (top-right) where a participant uses a magnet to illustrate a flicking gesture.

Space to spread out data

In 3 workshops (D, I, J), participants used space to spread out choices over large areas so that they could select from multiple options shown with rich representations. The abundant display space enabled participants to use several meters of the display for a temporary view to help select from a list of choices or to assign something to an item.

A typical example of this was seen in workshop J (97:18-97:34), where participants were working with categories containing thousands of photos: *Then I am able to take for instance these categories* [pointing gesture towards the area of the categories] *and spread them out over the upper part of the display* [doing a spreading gesture over a large area of the upper display area]. This enabled the participant to assign photographs to the categories.

Another variant of using space to spread out data were seen in workshop I (40:46-41:08), where participants analyzed results from an information retrieval system. Part of this work compared measures of different algorithms. In this situation participants imagined using the overview as an entry point to data: *If we could generate on the fly* [vocal: bouuf, snapping and doing a spreading gesture] *all the measures in one big table* […] *if we rather than having to look at it one by one* [while doing flicking gesture in the air] *could have a starting place with lots of information about, on summary data* […] *and then move to, ok let's go into the details and look at the ranks and what actually happened.*

The use of space to spread out data differs from participants' use of space to view information side-by-side. Using space to spread out data is temporary and typically used when participants need to select or modify data.

Trail of thoughts

With abundant display space, participants commented on the value of being able to see earlier steps of analysis by having these steps represented visually; they also referred back to and used representations of such steps in the workshops. In some workshops, data processing flows were used to represent this idea (A, C, G) and in others snapshots of the display state were shown in small (G, I). We saw examples of such trails of thought in 4 workshops.

An example of using a data processing flow was seen in workshop G (55:10-57:50), where participants drew steps of data processing as vertices and the order of processing as edges. The representation of the data processing flow is shown in Figure 3, bottom-right. Participants explained that it was useful to have an overview of how data were processed and be able to go back and look at earlier steps in the analysis. Results from individual vertices could be represented using histograms or other representations. A related observation was seen in workshop C (39:00-45:20, also described in the section on side-by-side viewing). Here the steps were represented directly by visual representations of results instead of by vertices. When participants wanted to explore a part of the results further, they would press this part, which would make an edge appear that led to a more detailed view of part of the data (see Figure 3, bottom-left).

An example of using snapshots was seen in workshop G (81:57-83:05) where participants discussed how to mark important findings while doing analysis to be able to summarize at the end of an analysis session: *If you could let it make up a summary so you simply could have a description of this [analysis] in time so that you at the end of a meeting quickly could summarize what we have been doing. ... if you simply had the display time your progress along the analysis* [gesturing over the lower part of the screen to indicate a horizontal line of display snapshots] *so that at some point you could say; now we are rewinding to the start of the meeting and then quickly go through the points we have touched upon. ... Then you would be able to do a commented summary [based on this*]. Participants also remarked that marking dead ends in analysis was important.

Movement

The size of the display naturally caused participants to move around in front of the display, and moving closer to or farther from it. Moving away from the display seemed to facilitate obtaining an overview and moving closer seemed to facilitate seeing details. When participants moved in front of the display, they did so to get to data or views of interest, to move out of other participants' view, to gather an overview, or to point to something on the display.

In workshop J (59:40-60:00) for example, participants moved close to the display to look at details in specific photographs and quickly back again to position this detail in their overview: *I can construct an overview of the photographs, I can see what's on the photographs while still being able look at the entire overview.* The sequence of first standing away from the display and thinking, then walking up close to interact with the display and then slowly backing up, as if to make sure things were as expected, was seen in 8 workshops; it was most visible in workshop J. To confirm this observation, we inspected movement patterns in workshop J by sorting still images grabbed with 15 second intervals. Three main categories of positions in relation to the display were observed: interacting or looking at the display (close), with the back turned to the display and interacting with other participants (middle), and away from the display facing it (far). Sorting the grabbed images into these categories showed that participants spent an equal amount of time in all three (close: 34%, middle: 33%, far: 33%).

In some workshops, we observed participants only taking half a step backwards to get distance from the display and to get an overview (e.g., workshop E: 26:40-26:41).

Another variant of movement relates to small movements with both feet on the ground. An example of this was seen in workshop B (23:05-23:07), where participants did a task on one part area of the display that required data placed in another area. To be able to grab the data located far away at the display, one participant leaned backwards, thus getting an improved field of view to the distant display area.

Gestures

We saw 172 gestures with the imaginary display that were significant or interesting enough for coding. We grouped these gestures into three types according to their occurrence in workshops: (a) on-screen (9 workshops, 44 gestures); (b) in front of screen (8 workshops, 43 gestures); and (c) in-air gestures (10 workshops, 85 gestures). Most of these gestures have been described in the literature. For instance, we coded 46 instances of sync- or asynchronous bimanual interactions.

An observation that surprised us was the use of very large gestures (13 gestures in total, 6 workshops). We see the size of these gestures to be related to display space. An example of a large gestures was seen in workshop J (95:50-95:55) where participants talked about changing overall states of the display (see Figure 4): *If there was a permanent image viewing function, which is this one* [pointing to a spot on the display] *having the large view. This is a view which you actually could do like this to* [gesturing with one hand from the left of the display to the right, almost 6 meters] *and draw it all the way over here, because now I just need it to be here*.

Figure 4. Example of a very large gesture.

DESIGN IDEAS

Our results suggest that information visualization systems could be designed with consideration for persistent views, not only as tool palettes and other interface objects, but also to show and interact with data such as raw data, variables, slices of data and general information views.

Views were fixed to top and bottom areas of the display for specific purposes, thus promoting the center area to a working or thinking area. This area was kept for things that were part of a thought process, whereas items supporting constructing and reconfiguring the working area were positioned in harder to reach positions (i.e., in the vertical periphery). Likewise, areas in the horizontal periphery could be used as persistent areas displaying for instance aggregated information. Participants moved back and forth in front of the display. This implies that such an area may need to move with the user. Participants also moved away from a display, for instance to gain an overview of items on the display. In this situation, these peripheral views may be irrelevant and could be hidden to not block important data.

Participants used views of data both side-by-side and one-by-one depending on the situation. This suggests enabling both styles of interaction with data. It also suggests a need to improve our understanding of when it makes sense to use space rather than interaction.

Participants also used one view of data to create new views next to the current view by interacting with parts of data in the view, thereby forming paths of interaction that enabled backtracking. Another method of providing backtracking was to show representations of previous display states, for instance in the bottom display area. This method seemed to be relevant for analysts when constructing a summary of a collaborative analysis session.

Data were temporarily spread out over large areas to enable participants to select from choices. Using space to show choices in rich detail and high resolution seems ideal. When the use is only temporary, these areas may block other data.

Gestures may be relevant to use both on, close to, and from a distance to the display. Large gestures seem to be relevant and perhaps the size of a gesture and the distance to what it refers to may carry meaning in itself.

DISCUSSION

We have presented a cross-domain workshop study of how domain experts would analyze their data with abundant display space. The workshops were analyzed to generate design ideas for interaction and visualization with abundant display space. The most prominent design ideas were:

- Use abundant display space for persistent views of data.
- Use middle center area to support thinking.
- Use vertical periphery to configure middle area.
- Enable both side-by-side and one-by-one views.
- Enable paths of interaction.
- Use abundant display space to support backtracking.
- Use abundant display space to show rich representations of choices.
- Enable use of large gestures.
- Support interaction from a distance.

Relation to Existing Work

In relation to the literature on large high-resolution displays, our design ideas warrant some comments. Earlier work has suggested that large displays promote physical navigation [2]. Certainly, movement in the workshops was necessary as no virtual navigation was possible. However, the workshops suggested that pairs use and switch between parts of the simulated display flexibly. This is similar to findings that high-resolution displays with touch may lead to less territorial behavior (e.g., [20]). The finding that people move not only sideways but also back and forth in relation to the display is related to the recent interest in proxemics for interaction [5].

Although we saw use of abundant display space to support backtracking, probing for styles of interaction related to undo/redo techniques from the desktop such as [27] did not resonate with participants. Implicit use of space to support backtracking seems a sensible way of using abundant display space, and is similar to how [30] represent history. In our workshops, however, history was integrated in the primary view. Participants suggested constructing summaries of analysis by marking important findings and representing these as snapshots of the display, which seems to have a different purpose than both undo/redo techniques and backtracking, and is perhaps similar to what Mahyar and colleagues saw [21].

Our observation of side-by-side and one-by-one views are in line with Gleicher and colleagues' notion of juxtaposition [15]. Here, side-by-side views are similar to juxtaposition in space, whereas one-by-one views reminds of juxtaposition in time and in some instances of blink comparison.

Workshop Methodology

We used cross-domain workshops as a methodology for uncovering new interaction styles and new uses of visualization. In the introduction to this paper, we speculated that cross-domain workshops with simulated large displays might lead to interesting insights. Next, we want to revisit this speculation based on the experience of running the workshops and of analyzing them.

In many and important parts of the workshops we found that participants' imagination was vivid. After a workshop, we showed a participant a large high resolution display. He commented that seeing this display would have made him think differently about the whiteboard during the workshop, which would probably have both positive and negative effects on participants' imagination.

One reason why we were able to derive design ideas from participants seems to be that we used both their behavior and their comments to derive ideas. Another reason seems to be that comparing across domains helps identify common threads of data analysis. The present study has identified some of the same uses of abundant display space across domains as varied as photo management, health care, and container loading.

In other parts of the workshops, it seemed difficult for participants to imagine new technology: it was clear from some participants' dialogue that they thought in terms of the data representations, software, and interaction techniques they know and use today. For instance, in one workshop participants would have alternate terms related to a given sentence by a number describing the part of the sentence to which the term was an alternative. They did not see, however, that with abundant display space the position information could be substituted by placing terms directly around the sentence (similar to a large version of excentric labels [14]). In other workshops, participants would talk about using arrow keys to sift through pictures, or talk about how to access syntax information.

The use of a whiteboard as a large display generally worked well: whiteboards are ubiquitous and can be used right away for drawing and attaching prints. Compared to studies of whiteboard use in visualization [34,38], we saw similar rich and unconstrained use. This suggests that the range of representations and interactions with the whiteboard might be varied enough to inform design. However, we see at least two ways the workshops can be improved.

First, whereas the whiteboard worked well to convey a sense of abundant display space it did not convey any sense of resolution. Most likely, such a sense was developed based on the resolution of the prints that participants carried with them, in addition to the scale at which they drew on the whiteboard. We think that physically large and small prints of data, as well as high and low resolution images may exemplify the role of resolution to participants, making it unnecessary for workshop moderators to explicitly describe or probe for resolution.

Second, we saw a lack of motivation to remove data once it had been placed on the whiteboard. It is unclear how this relates to how people use whiteboards' available space and only erase on an as-needed basis [23]. It might also be related to working out the tasks in the workshop setting (i.e., as a group) which differs from how some of the participants normally work.

Limitations

Our paper has a number of limitations. First, our workshop approach attempts to bridge doing field studies of data analysis to derive implications for design and using models/theories to derive implications. Recommendations derived from this attempt, however, need to be validated using other types of method. In particular, we are interested in trying to implement the interaction techniques developed and test them across domains, following the idea that cross-domain explorations may integrate the concrete (task solution in a single domain) with the general. Second, the number of participants in each workshop was low and thus, we cannot extend our findings to larger group sizes.

REFERENCES

1. Andrews, C., Endert, A., and North, C. Space to Think: Large, High-Resolution Displays for Sensemaking. In *Proc. CHI 2010*, (2010), 55–64.
2. Ball, R., North, C., and Bowman, D.A. Move to Improve: Promoting Physical Navigation to Increase User Performance with Large Displays. In *Proc. CHI 2007*, (2007), 191–200.
3. Ball, R. and North, C. Analysis of User Behavior on High-Resolution Tiled Displays. In *Proc. INTERACT 2005*, (2005), 350–63.

4. Ball, R. and North, C. The Effects of Peripheral Vision and Physical Navigation on Large Scale Visualization. In *Proc. GI 2008*, (2008), 9–16.
5. Ballendat, T., Marquardt, N., and Greenberg, S. Proxemic interaction: Designing for a Proximity and Orientation-Aware Environment. In *Proc. ITS 2010*, (2010), 121–30.
6. Baudisch, P., Cutrell, E., Robbins, D., et al. Drag-and-Pop and Drag-and-Pick: techniques for accessing remote screen content on touch- and pen-operated systems. In *Proc. INTERACT 2003*, (2003), 57–64.
7. Baudisch, P., Good, N., Bellotti, V., and Schraedley, P. Keeping things in context: a comparative evaluation of focus plus context screens, overviews, and zooming. In *Proc. CHI 2002*, (2002), 259–66.
8. Bi, X. and Balakrishnan, R. Comparing usage of a large high-resolution display to single or dual desktop displays for daily work. In *Proc. CHI 2009*, (2009), 1005–14.
9. Birnholtz, J.P., Grossman, T., Mak, C., and Balakrishnan, R. An Exploratory Study of Input Configuration and Group Process in a Negotiation Task Using a Large Display. In *Proc. CHI 2007*, (2007), 91–100.
10. Branham, S., Golovchinsky, G., Carter, S., et al. Let's Go from the Whiteboard: Supporting Transitions in Work through Whiteboard Capture and Reuse. In *Proc. CHI 2010*, (2010), 75–84.
11. Card, S.K., Mackinlay, J.D., and Shneiderman, B. Knowledge Crystallization. In *Information Visualization*. 1999, 10–12.
12. Czerwinski, M., Smith, G., Regan, T., Meyers, B., Starkweather, G., and Way, O.M. Toward Characterizing the Productivity Benefits of Very Large Displays. In *Proc. INTERACT 2003*, (2003), 9–16.
13. Elrod, S., Lee, D., Gold, R., et al. Liveboard: a large interactive display supporting group meetings, presentations, and remote collaboration. In *Proc. CHI 1992*, (1992), 599–07.
14. Fekete, J.-D. and Plaisant, C. Excentric Labeling: Dynamic Neighborhood Labeling for Data Visualization. In *Proc. CHI 1999*, (1999), 512–19.
15. Gleicher, M., Albers, D., Walker, R., Jusufi, I., Hansen, C.D., and Roberts, J.C. Visual comparison for information visualization. *Information Visualization 10*, 4 (2011), 289–09.
16. Grudin, J. Partitioning Digital Worlds: Focal and Peripheral Awareness in Multiple Monitor Use. In *Proc. CHI 2001*, (2001), 458–65.
17. Halskov, K. and Dalsgård, P. Inspiration card workshops. In *Proceedings of DIS 2006*, (2006), 2–11.
18. Huang, E.M., Mynatt, E.D., Russell, D.M., and Sue, A.E. Secrets to success and fatal flaws: the design of large-display groupware. *IEEE Computer Graphics and Applications 26*, 2006, 37–45.
19. Isenberg, P., Tang, A., and Carpendale, S. An exploratory study of visual information analysis. In *Proc. CHI 2008*, (2008), 1217–26.
20. Jakobsen, M.R. and Hornbæk, K. Proximity and Physical Navigation in Collaborative Work With a Multi-Touch Wall-Display. In *EA CHI 2012*, (2012), 2519–24.
21. Mahyar, N., Sarvghad, A., and Tory, M. Note-taking in co-located collaborative visual analytics: Analysis of an observational study. *Information Visualization 11*, 3 (2012), 190–04.
22. Malik, S., Ranjan, A., and Balakrishnan, R. Interacting with large displays from a distance with vision-tracked multi-finger gestural input. In *Proc. UIST 2005*, (2005), 43–52.
23. Mynatt, E.D. The Writing on the Wall. In *Proc. INTERACT 1999*, (1999), 196–04.
24. Nancel, M., Wagner, J., Pietriga, E., Chapuis, O., and Mackay, W. Mid-air Pan-and-Zoom on Wall-sized Displays. In *Proc. CHI 2011*, (2011), 177–86.
25. Robinson, A.C. Collaborative synthesis of visual analytic results. In *Proc. VAST 2008*, (2008), 67–74.
26. Russell, D.M., Stefik, M.J., Pirolli, P., and Card, S.K. The cost structure of sensemaking. In *Proc. CHI 1993*, (1993), 269–76.
27. Seifried, T., Rendl, C., Haller, M., and Scott, S.D. Regional Undo / Redo Techniques for Large Interactive Surfaces. In *Proc. CHI 2012*, (2012), 2855–64.
28. Shneiderman, B. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *Proc. Visual Languages 1996*, (1996), 336–43.
29. Shoemaker, G., Findlater, L., Dawson, J.Q., and Booth, K.S. Mid-Air Text Input Techniques for Very Large Wall Displays. In *Proc. GI 2009*, (2009), 231–38.
30. Shrinivasan, Y.B. and van Wijk, J.J. Supporting the analytical reasoning process in information visualization. In *Proc. CHI 2008*, (2008), 1237–46.
31. Strauss, A. and Corbin, J. *Basics of qualitative research. Techniques and procedures for developing grounded theory*. SAGE Publications, Thousand Oaks, CA, 1998.
32. Streitz, N.A., Geibler, J., Holmer, T., et al. i-Land: an interactive Landscape for Creativity and Innovation. In *Proc. CHI 1999*, (1999), 120–27.
33. Tan, D.S., Gergle, D., Scupelli, P.G., and Pausch, R. With Similar Visual Angles, Larger Displays Improve Spatial Performance. In *Proc. CHI 2003*, (2003), 217–24.
34. Tang, A., Lanir, J., Greenberg, S., and Fels, S. Supporting Transitions in Work: Informing Large Display Application Design by Understanding Whiteboard Use. In *Proceedings of GROUP 2009*, (2009), 149–58.
35. Thomas, J.J. and Cook, K.A. *Illuminating the Path: The Reseach and Development Agenda for Visual Analytics*. 2005.
36. Vogel, D. and Balakrishnan, R. Distant freehand pointing and clicking on very large, high resolution displays. In *Proc. UIST 2005*, (2005), 33–42.
37. Vogt, K., Bradel, L., Andrews, C., North, C., Endert, A., and Hutchings, D. Co-located Collaborative Sensemaking on a Large High-Resolution Display with Multiple Input Devices. In *Proc. INTERACT 2011*, (2011), 589–04.
38. Walny, J., Carpendale, S., Henry Riche, N., Venolia, G., and Fawcett, P. Visual thinking in action: visualizations as used on whiteboards. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (2011), 2508–17.
39. Wobbrock, J.O., Morris, M.R., and Wilson, A.D. User-defined gestures for surface computing. In *Proc. CHI 2009*, (2009), 1083–92.
40. Yi, J.S., Kang, Y.A., Stasko, J., and Jacko, J. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (2007), 1224–31.
41. Yost, B. and North, C. The perceptual scalability of visualization. *IEEE Transactions on Visualization and Computer Graphics 12*, 5 (2006), 837–44.
42. Ziemkiewicz, C., Gomez, S., and Laidlaw, D.H. Analysis Within and Between Graphs: Observed User Strategies in Immunobiology Visualization. In *Proc. CHI 2012*, (2012), 1655–58.