Investigating the Use of a Dynamic Physical Bar Chart for Data Exploration and Presentation

Faisal Taher, Yvonne Jansen, Jonathan Woodruff, John Hardy, Kasper Hornbæk, Jason Alexander

Investigating the Use of a Dynamic Physical Bar Chart for Data Exploration and Presentation

Faisal Taher, Yvonne Jansen, Jonathan Woodruff, John Hardy, Kasper Hornbæk, and Jason Alexander.

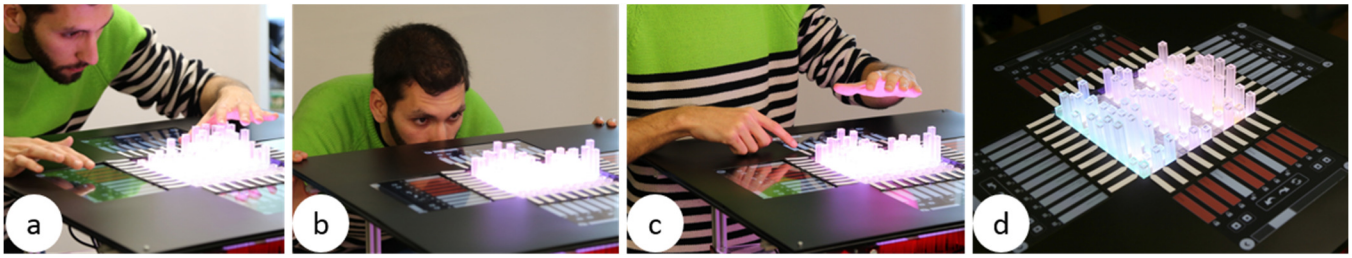

Fig. 1. (a) User pointing at a label and at a data point, (b) user crouching to inspect the data, (c) user hovering their hand above a group of data points, (d) data organized into three sections and separated by hidden data points.

Abstract—Physical data representations, or data physicalizations, are a promising new medium to represent and communicate data. Previous work mostly studied passive physicalizations which require humans to perform all interactions manually. Dynamic shape-changing displays address this limitation and facilitate data exploration tasks such as sorting, navigating in data sets which exceed the fixed size of a given physical display, or preparing "views" to communicate insights about data. However, it is currently unclear how people approach and interact with such data representations. We ran an exploratory study to investigate how non-experts made use of a dynamic physical bar chart for an open-ended data exploration and presentation task. We asked 16 participants to explore a data set on European values and to prepare a short presentation of their insights using a physical display. We analyze: (1) users' body movements to understand how they approach and react to the physicalization, (2) their hand-gestures to understand how they interact with physical data, (3) system interactions to understand which subsets of the data they explored and which features they used in the process, and (4) strategies used to explore the data and present observations. We discuss the implications of our findings for the use of dynamic data physicalizations and avenues for future work.

Index Terms—Shape-changing displays, physicalization, physical visualization, bar charts, user behaviour, data presentation.

1 Introduction

Physical data visualizations, or data physicalizations, are "artifacts whose geometry or material properties encode data" designed to better support "cognition, communication, learning, problem solving, and decision making" [18]. Recent work provides evidence of their benefits including their utility as education tools [32], as mediators to engage people in data exploration ([24], [31]), or to increase the understanding of statistical data [11]. In popular media, Hans Rosling uses physicalizations to communicate data on global health [7]. While the majority of previous work in this area has focused on passive, fabricated physicalizations, shape-changing technology (e.g., inFORM [9], EMERGE [29]) promises to increase the interactivity of physicalizations to eventually reach a similar level of control as that possible with on-screen visualizations. For example, Microsoft's Physical Charts [26] demonstrates dynamic bars and pie charts, and Taher et al. [29] derived initial interaction preferences for a physically dynamic bar chart.

- *Faisal Taher is with Lancaster University. E-mail: f.taher@lancaster.ac.uk.*
- *Yvonne Jansen is with University of Copenhagen. E-mail: yvja@di.ku.dk.*
- *Jonathan Woodruff is with Lancaster University. E-mail: j.woodruff@lancaster.ac.uk.*
- *John Hardy is with Lancaster University. E-mail: john@heinventions.com.*
- *Kasper Hornbæk is with University of Copenhagen. E-mail: kash@di.ku.dk.*
- *Jason Alexander is with Lancaster University. E-mail: j.alexander@lancaster.ac.uk.*

Despite the promise of dynamic physicalizations, our knowledge of how users interact with these systems is limited. Specifically, there are no empirical studies exploring user behaviour around physicalizations or how physicalizations are used to present and describe the data to others. We are left with important questions: How do users move around and interact with such a display? How do they orient themselves in a dataset larger than the physical display? What strategies are used to describe their observations? This paper aims to answer these questions through empirical observations of how people behave, interact with, and present observations using a dynamic physical data system.

We present a user study with 16 participants using EMERGE [29], a physically dynamic bar chart with a grid of 100 self-actuating bars. During the study, participants were asked to explore an unseen dataset to discover themes, and present these themes and their relationships using the dynamic physicalization. We recorded and analyzed video and birds-eye-view Kinect data to examine participants' hand, arm, and body movements (e.g., see Fig. 1). System interaction logs were examined to analyze how people navigated through a dataset exceeding the size of the physical viewport and which functionalities were used. Further, we discuss the strategies that participants used to explore and present observations using EMERGE. We use these insights to characterize behaviours and interactions, and to identify avenues for future work.

This paper therefore contributes characterizations of: (1) users' body movements to understand how they approach and react to the physicalization, (2) their hand-gestures to understand how they interact with data, (3) system interactions to understand which

subsets of the data they explored and which features they used in the process, and (4) data exploration and presentation strategies.

2 Background and Study Rationale

Physical data representations have existed for as long as civilization. The increasing need for powerful data manipulation mechanisms in the 20th century brought about an almost complete switch to virtual on-screen data visualizations. Only since the inception of digital fabrication technologies, physical data representations began to reappear and to become an object of study in HCI and visualization. Research has explored specific properties of physicalizations such as engaging diverse audiences with data ([24], [26], [16]), motivating behaviour changes ([28], [20]), or exploring multisensorial data communication [13]. Other work focusses on specific user groups, such as molecular biologists [10], or education [32]. Specific aspects of physicalizations have also been studied, such as how representation modality affects the user-experience of data artifacts [14], how physical visualizations compare to their on-screen counterparts for information retrieval tasks [17], and how people use physical tokens to author and make sense of physical data [15].

However, most of the aforementioned work is based on passive physical visualizations: those that are either static or require a human to perform tasks such as filtering or re-organization manually. Applying shape-changing technologies [26] to data physicalization promises to merge the benefits of physical data representations with the power of computation. This physical dynamicity creates additional cost and complexity, meaning that few works have explored this area (however, first attempts at facilitating access to the required technologies are being made [12]). Currently, only a few dynamic shape displays for data physicalization exist. Among these are the Relief [23] and inFORM [9] which have demonstrated data physicalizations for demo installations and in videos, and EMERGE [29], which was explicitly constructed for data physicalization.

Taher et al. explored different types of interactions with the EMERGE system [29]. Overall, they found that people preferred direct interaction with the physical bars to indirect techniques that relied on touch buttons around the physical bars. However, for some repetitive tasks such as scrolling through a larger dataset, an indirect virtual scroll bar was preferred. While providing some initial insights into which type of interaction people prefer with a dynamic physical display, this first study did not look into *how* people interact with the display. For passive physicalizations, Huron et al. [15] and Jansen et al. [17] report people making extensive use of their hands to support their thinking processes. With a dynamic physical display, physical "thinking actions", like resting a finger on a bar or a label while exploring its vicinity, might invoke a system function.

It is currently unclear how people would interact with a dynamic data physicalization. Does the added interactivity encourage them to explore the data? In an open-ended task setting, do they make use of the different interaction techniques? How do people make use of the physical properties when asked to talk about their findings?

3 Observational Study

We chose to study these questions by designing an exploratory study that combines qualitative and quantitative analysis. To elicit a wide range of different behaviours, we included two phases of interest, following a training phase: an initial data exploration phase where people were asked to explore a dataset according to their personal interests, and then a presentation phase where we asked them to use the physicalization to illustrate what they had found.

3.1 Apparatus

3.1.1 Physically Dynamic Bar Chart

To aid our investigations, we used EMERGE, a physically dynamic bar chart [29]. The system (see Fig. 2) consists of a 10×10 grid of plastic bars that are linked to 100 motorized potentiometer sliders capable of 100mm travel. Each bar can be illuminated by an RGB LED. The top of EMERGE consists of four touchscreen panels to display labels as well as additional controls for organizing data. Opposite sides of EMERGE show the same information, allowing users to control the data from any side. Our software setup uses a client-server architecture with web-socket communication between JavaScript clients on the touchscreens, and a C# client running on EMERGE. The interactions with the bar chart were supported by the four touchscreen panels as well as by the physical data points (bars).

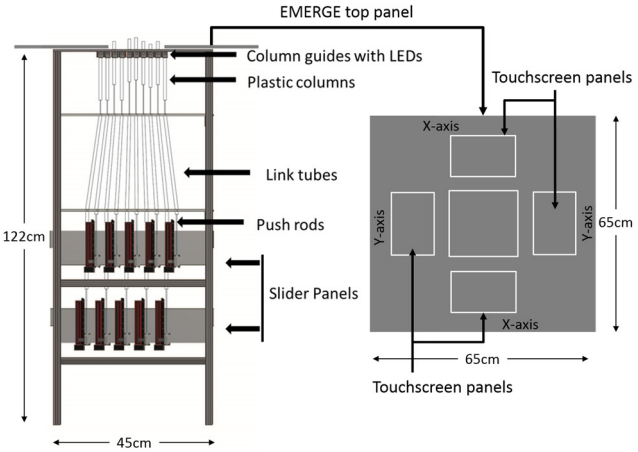

Fig. 2. Component overview and physical dimensions of EMERGE.

We implemented six system interactions based on recommendations from previous work [29]:

- **Highlighting:** users can highlight and emphasize individual data points by pulling a bar, which dims the unselected data.
- **Swapping:** to re-order rows along an axis, users can drag a label on the touch screen panels and drop it on top of the target label, which then swaps the two rows.
- **Scrolling:** users can navigate through a larger dataset (i.e., datasets larger than 10 × 10 items), by (a) dragging the scrollbar slider on the touchscreen panels, or (b) by pressing arrow icons on the touchscreen panels that scrolls through single rows.
- **Locking:** rows along each axis can be locked in place by pressing the lock icon on the touch panels. This keeps that particular row of data points in place as users are scrolling, and restricts interactions such as hiding data points.
- **Hiding:** to temporarily remove irrelevant data, users (a) press individual data points to hide them, or (b) press two data points around the edges to single out two rows for comparison, which then hides all of the other data points (except for those which are *locked*)
- **Snapshot**: users can save and return to particular "views" (i.e., the current 10×10 set of data points) where they may have highlighted and re-organized several data points, by pressing the snapshot icon on the touchscreen.

Additional functions include *undo* and *redo*, which enables users to go back and forth in the interaction history and *reset,* which clears all changes and returns to the initial view.

3.1.2 Data Capture

To study user interactions with EMERGE, we set up two video cameras that captured user actions from the top, and from the side. User actions consist of system interactions, hand gestures, and body movements. The top-view camera was placed directly above the system to capture any actions obscured from the side-view camera by the user. The side-view camera was placed two meters away from the participants to capture depth-of-field perspective of their actions.

The movement of participants around EMERGE was additionally captured using a Microsoft Kinect. We were interested in whether participants were static or mobile while interacting with the data, and so we tracked their head and body movements to generate movement heatmaps. In addition, interactions with the four touchscreen panels, as well as with the bars, were directly logged by the system to quantitatively analyze participant interactions.

3.2 Participants

The observational user study was carried out with 16 participants (9 female) with a mean age of 30 years (four participants were above 39 years). None of the participants had previous experience with dynamic physicalizations or with interactive on-screen visualization tools. Four participants created static bar charts on a monthly basis or more, and eight participants gave presentations with bar charts a few times a year. One participant gave weekly presentations with bar charts.

3.3 Setup and Procedure

During the study, participants were welcomed individually and asked to fill out a demographics questionnaire. Following this, they were introduced to the study setup, which included a demonstration and training phase with the dynamic physical chart. They then went on to an open-ended exploration of an unknown dataset (explained below) with the goal to prepare a short presentation of their insights. At the end of the study, participants were interviewed and asked to fill out a questionnaire. The study phases are described next. The datasets used in the study were encoded by using unique colours to differentiate between rows, and by using the height of the bars to represent the values. While the exploration and presentation phases were open-ended, we included a set of different tasks in the initial training phase as detailed below to cover different types of visualization tasks [1]. The purpose of these training tasks was to give participants the opportunity to familiarize themselves with the system in the context of different types of questions typically asked during data exploration.

3.3.1 Demonstration and Training Phase

The demonstration and training phase was designed to introduce the functionalities of EMERGE to the participants, and to allow them to practice the interactions. This phase lasted for 30 minutes. A UK rainfall dataset was used for this phase[1] that showed average rainfall (encoded in the height of the bars) for 11 regions over 103 years. The labels along the x-axis showed years, and the labels on the y-axis showed locations within the UK. Each interaction and its functionality was described (e.g., highlighting a data point can be useful to emphasize an interesting observation). Participants were also asked to practice these interactions for 5 minutes.

To allow participants to build some proficiency with the EMERGE system, participants then carried out a training exercise for 20 minutes where the experimenter verbally asked participants to perform a variety of different tasks. The dataset for this training was a survey from 1974 where 52 college students had provided ratings for the appropriateness of 15 actions in 15 situations [25]. For example, ratings would reflect how appropriate (height of the bar) college students felt it was to sleep (shown on the x-axis) in class (shown on the y-axis). Participants were asked to do the following:

1. Focus on either actions or situations, then select and group interesting categories together.
2. Scroll to find more interesting categories and group them next to the locked ones. Lock the new ones and take a snapshot.
3. Keep the data points that have been selected and hide the rest. Then take another snapshot.
4. Compare with situations or actions (if participants picked actions, compare with situations, and vice versa) and highlight any unusual or interesting data points. Then take a snapshot.
5. Briefly explain the highlighted data points. Then cycle through the three saved snapshots and explain what the snapshots are communicating.
6. Repeat the above by focusing on the other category (i.e., situations if participants initially selected actions).

3.3.2 Exploration Phase

Following the training phase, participants were presented with a new dataset, i.e. a subset of World Value Survey data from 2006, which consists of ratings from inhabitants of 46 European countries on 31 topics, i.e., religion, social issues, politics, military, healthcare, and economy [8]. Topic labels were shown along the x-axis, and country labels were shown along the y-axis. Participants were encouraged to explore the dataset for at least 10 minutes (no time limit was indicated nor enforced) and to identify themes, which they would be asked to informally present as the last part of the study.

3.3.3 Presentation Phase

Participants were asked to present their observations informally to the experimenter in their preferred way after the exploration phase. The experimenter remained in the same location to prevent bias stemming from participants relating to different locations.

3.3.4 Post Study

After the presentation, a short semi-structured interview was carried out for approximately five minutes to receive feedback on how participants felt about the 10×10 grid size, using EMERGE as a presentation tool, the ease of discovering themes, and how they felt carrying out the system interactions (e.g., swapping rows, hiding data). This was followed by a short questionnaire where participants provided Likert ratings for the above questions.

4 Research Questions & Data Analysis

Our data analysis is driven by three exploratory research questions:

(1) How do people make use of a dynamic data physicalization to explore data and to communicate their insights? Do they simply apply the strategy learned during the initial training session or do they develop their own?

(2) How do they move and behave around the physicalization? Do they move around or stay in place on one side? Do they interact predominantly with the physical bars or the labels on the side-screens?

(3) How do system limitations such as the fixed number of concurrently visible data points affect people's data exploration behaviour? Do they explore all or most of the dataset or do they focus on specific subsets?

We collected three types of data from participants' actions to answer these questions: body movements, hand gestures, and system interactions. Body movements and hand gestures were gathered through video-coding, whereas system interactions were logged by the EMERGE system.

4.1 Video-Coding

Prior to analysis, all videos were separated into the two study phases. We developed a codebook through open coding of a random subsample of videos (from 3 participants) from both phases by all co-authors. This led to five key hierarchical categories (see below).

Hand and arm gestures that do not trigger the device, e.g., pointing at bars, pointing at labels, palm hovering of the bars and labels, hesitations.

Hand and arm gestures that trigger the device, e.g., gestures such as pressing and pulling the bars with one hand or two hands.

[1] http://www.metoffice.gov.uk/climate/uk/summaries/datasets (last accessed 11/03/2016)

Body/head movements that change a user's view onto the graph bars, e.g., walking around the device, leaning on top of the device, head tilts, leaning back, sidestepping, and crouching.

Failed or impossible interactions, e.g., actions that are not implemented, such as pressing a bar, which is already hidden.

Unusual actions, e.g., surprised reactions or movements.

The complete set of videos was then divided among three coders. Interrater reliability was coded from 10% of the videos (from 6 different participants to the initial codebook creation) for both phases. We used Lasecki et al.'s approach [22] to calculate reliability using Cohen's Kappa [5]: category events were placed into 1-sec bins to determine if each coder contributed a 'yes' or a 'no' to the events in that bin. A mean kappa for each phase was then calculated. This gave agreement rates of 0.42 for the exploration phase and 0.56 for the presentation phase. Using Landis and Koch's classification [21] the coders had moderate agreement in both phases.

4.2 Log Data

The Kinect tracking data was only used to analyze large-scale body movements of participants, such as at which sides of the chart they stood during the study session. The subsets of the system interaction logs were analyzed to learn which physical bars participants interacted with and how participants navigated through the dataset.

5 Results

We first report high-level findings across the study phases before providing details for the exploration and presentation phases.

5.1 General Findings

Both phases were open-ended and driven by the interests and motivation of participants. We enforced no time limits and thus observed a wide variation of time spent exploring the data (see bar lengths in Fig. 3–top), averaging at around 16 minutes, as well as presenting findings, averaging at about 5 minutes (median ~3.5 min).

Figure 4–top gives an overview of participant behaviour over time, showing the prevalence of different types of movements, gestures, and system actions over the course of their study session. We found that body movements were frequent and equally present across both phases. Body movements include movement of the head or upper body (shown in purple) or around the display (shown in green) to change one's perspective onto the physical 3D chart. While this is to be expected during the exploration phase, we expected to observe this less during the presentation phase. However, since the presentations were informal and not carefully prepared and practiced, these movements likely indicate participants reaffirming themselves of their previous observations during the exploration phase.

In terms of hand gestures, we found a pronounced difference between the two phases. Hand gestures comprise gestures that did not result in executing a system function[2]. While we expected a high frequency of such gestures during the presentation phase, e.g., to direct the attention of the listener, gestures performed during the exploration phase cannot serve the purpose of communication as participants were working alone. These gestures varied considerably between participants with some performing only few to none (e.g., P1, P3, P15) while others made extensive use of their hands (e.g., P7, P8, P14). We discuss this finding later on in more detail.

For system interactions, that is, navigation through the dataset and organization interactions such as sorting, filtering, and highlighting, we observed a large difference between the two phases as well as different styles between participants. Notably, few participants (P1, P5, P10, P15) made use of the interactive features during the presentation phase apart from the snapshot recall function. Some participants made no use of snapshots during the presentation despite having prepared some during their exploration (P3, P7, P13), some chose not to prepare any snapshots (P5, P6), while the majority structured their presentations around two to four snapshots.

We deliberately chose a dataset for this study which exceeded the physical display size considerably: the EMERGE display can show 10×10 data points while our dataset contains 31×46 data points. Fig. 4 shows which parts of the entire dataset each participant explored with topics plotted on the x-axis and countries plotted on the y-axis. The dark rectangle in the bottom left of each participants' area in the figure indicates the initial viewport at the start of their exploration or after a reset. Colour indicates time-on-display using a logarithmic colour scale. Blue-shaded areas indicate which parts of the dataset were on display during the exploration phase and the overlaid red-shaded areas indicate time-on-display during the presentation phase while white areas indicate that these were not viewed at all by a participant (the logarithmic scale ensures that even a short viewing time of 3 sec is visible as a light blue).

Overall, we find that data exploration strategies varied considerably between participants and between phases. Notably, no participant explored the entire dataset. Of those participants who spend more time on data exploration, some viewed a large subset of the data (e.g., P10, P14) whereas others did not (e.g., P3). Whether this was intentional or due to missing overview features of the system is unclear. The observed behaviour could indicate that participants had trouble keeping an overview. Alternatively, the limited exploration time and the set goal to present findings about the data could mean that certain aspects of the dataset quickly spiked their interest and led them to focus on those areas. For the presentation phase, all participants used small subsets of the complete dataset. While some talked with subsets that were already adjacently placed in the original layout (e.g., P3, P6), others rearranged and compiled more elaborate views (e.g., P1, P10, P12).

5.2 Exploration Phase

Here we analyze the data in more detail for the exploration phase.

5.2.1 System Interaction Behaviours

We observed distinct behaviours surrounding interactions such as scrolling, swapping, locking, and snapshots. Participants often scrolled one row at a time – both in quick succession and by pausing between each scroll to inspect the data.

Organization. Data was organized in multiple ways and to differing extents across participants. Reorganization of the data by participants is visible in Fig. 3 in the form of striation: large differences in terms of time-on-display between adjacent rows and columns can only occur if participants reordered them. Thus, large uniform blocks indicate little reordering while striation with individual darker rows or columns indicate that these were 'locked' within the viewport while the participant continued to scroll. While for some participants (e.g., P1, P5, P13) striation only occurs along one dimension, others reorganized data along both dimensions (e.g., P9, P10). Organization occurred in two ways: first by relocating rows (by swapping) from one location to another (e.g., countries on one end of the dataset to another end) which involved a large amount of scrolling and swapping, and second by scrolling and locking rows of interest. The second approach is generally quicker for grouping rows, and participants often used this approach. Rows were also organized in three different configurations to compare and contrast: single groups where all rows of interest were placed next to each other, two groups separated by irrelevant rows, and unstructured groupings where rows of interest were in random locations. Grouped rows were nearly always locked. We also observed that most participants (9) had hidden irrelevant data to either emphasize or create a barrier between grouped rows (e.g., see Fig. 1d).

[2] Note that hand gestures in Fig. 3–top only include gestures which did not result in system interactions whereas the detailed breakdown of interactions in the charts on the bottom include both (press/pull results in a system interaction whereas point does not).

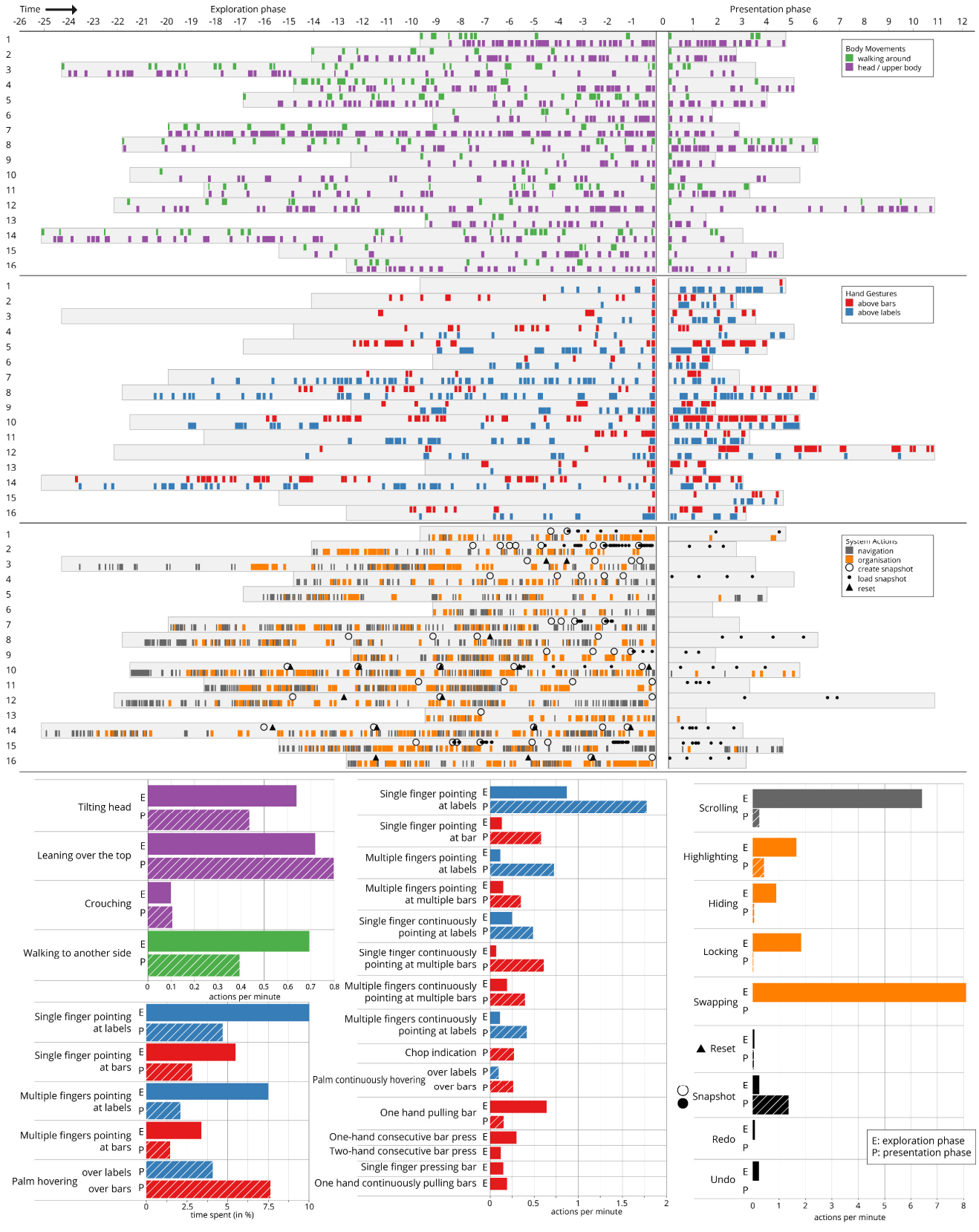

Fig. 3. Top: cinematic log visualization for all participants and both phases. Bottom: detailed breakdown of the different interactions, movements, and gestures averaged over all participants.

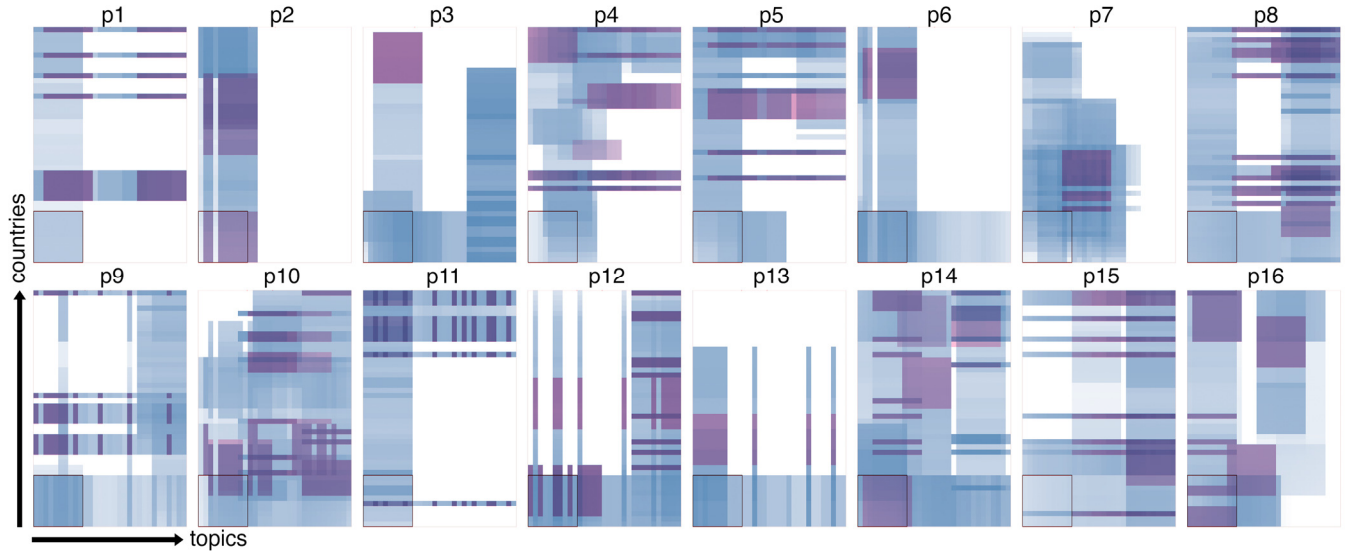

Fig. 4 - Heatmaps illustrating participants' navigation behaviour through the entire dataset. Blue indicates parts viewed during the exploration phase, red indicates which parts were used during the presentation phase

Notably, one participant had highlighted two entire rows rather than hiding the irrelevant ones.

Provenance through snapshots. Two participants saved snapshots that did not pertain to specific observations; they preserved their exploration history. For example, P11 initially grouped a set of countries, took a snapshot, grouped a set of topics, took another snapshot, hid irrelevant countries, took a third snapshot, then hid irrelevant topics, and took a final snapshot. Similarly, P13 grouped a set of topics, hid irrelevant rows, took a snapshot, and continued exploring the data. In both these cases, the snapshot function was used to keep a history of the different stages of their exploration.

5.2.2 Movement Behaviour

Participants moved in four different ways around the physicalization: walking between different sides of EMERGE, tilting their head during interaction, leaning over the top of EMERGE, and crouching down (e.g., see Fig. 1b). We observed that head tilts were typically subtle movements, whereas leaning over the top and crouching were either obvious or subtle. Fig. 3–bottom shows the number of movements per minute for each of the four movement behaviours. Leaning, walking and head-tilting were the most common during exploration and most frequently occurred after scrolling interactions.

The relationship between scrolling and movements such as walking, head tilts, leaning and crouching is likely to be caused by participants inspecting the data from various angles to better understand relationships between the data points. The 3D nature of the data representation creates different views from different angles. Thus, multiple perspectives can help to confirm relationships. In particular, we observed that P14 repeatedly combined scrolling interactions with crouching and leaning 13 times within a 15-second time window. P14 would scroll (which would change the 'shape' of the bar matrix as new data is scrolled into view), stop and inspect by crouching and leaning, and repeat this sequence. Similarly, P13 frequently scrolled and walked between different sides of EMERGE to inspect the data, and P1 frequently tilted their head while scrolling. We also observed that walking between sides and leaning were carried out to interact with different touchscreen axes.

All participants moved between at least two sides of EMERGE and four participants moved between three sides. Fig. 5-left shows a heatmap generated from the top view of the Kinect depth camera, which tracked the movement of all participants during the exploration phase. Participants mainly moved between the north and west side of EMERGE (represented by the white square). This choice is likely due to the experimenter's location, which was a few feet away from the south of EMERGE.

Fig. 5-right shows which of the physical bars in the 10x10 grid participants predomfinantly interacted with. The heatmap indicates that participants mainly interacted with bars along the edges (the comparison form of hiding, such as concurrently pressing bars to hide) is not included as it is only triggered by the bars around the edges of the grid, and would therefore skew the data), with fewer interactions scattered through the rest of the grid. We observed that bar interactions in the center regions were less frequent as they were harder to reach amongst surrounding bars (especially if participants wanted to reach a low bar that was surrounded by higher ones).

5.2.3 Gesturing, Pointing, and Physical Interactions

A range of pointing and touching gestures were observed during the exploration phase (see Fig. 3–bottom). Common gestures included both discrete and continuous pointing at labels with one finger, pulling the bars (highlighting interactions), and one-handed consecutive bar press (hiding). Similarly, participants also spent a considerable amount of time using continuous pointing gestures, e.g. participants spent 10% of their exploration time pointing at labels.

We observed that pointing gestures during exploration were likely related to aiding the participants' thinking process, for example, row labels or data points of particular interest. Participants typically remained static during pointing gestures. In particular, P14

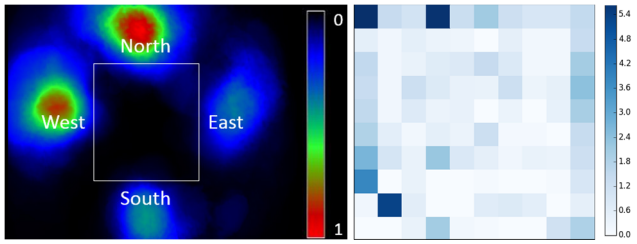

Fig. 5. Left: Heatmap of the movement of all participants during exploration (Range: 1 – more time spent in a region, 0 – less time spent). Right: heatmap of regions of the 10×10 grid showing the percentage of where participants pressed or pulled.

used a large number of discrete single-finger pointing gestures at the row labels and the bars (e.g., see Fig. 1a). For example, they pointed at a label, and then followed with their finger along (over the top) a single row of bars on three separate occasions. P13 also repeatedly displayed this sequence of gestures.

Physical interactions directly with the bars only occurred when the bars were not moving (i.e., after they had scrolled, or some rows had been re-ordered). We believe the bar movements indicated to participants that an interaction was still in progress. Bars were mainly highlighted individually, but we observed a few cases where participants confidently highlighted several data points in quick succession. In particular, P1 exhibited this behaviour while highlighting data points, and used both hands consecutively to "pull out" interesting observations. Participants normally paused before and after these bursts of highlighting. P1 continuously highlighted data points that were spread across the data view, whereas P12 would highlight an entire row. For example, P12 created a snapshot with two entire rows highlighted for comparison. Here we observe two different strategies: P1 highlighted specific data points of interest, whereas P12 highlighted to compare trends for two countries in the dataset.

Press gestures on bars would cause the individual bar to hide. Participants mainly used a single-finger on one hand to trigger consecutive bar presses. To trigger row comparison, participants usually paused for a few seconds with their fingers over the two bars before pressing, which likely indicates that they wanted to confirm their selection (of the rows to keep), and to also correctly carry out the interaction (both bars need to be pressed at the same time). We observed instances of participants "celebrating" after carrying out consecutive bar presses. For example, P14 waved both arms on two occasions, P13 would smile or show content facial expressions, and P13 verbally expressed content on two occasions. We attribute these behaviours to this simultaneous press being the most difficult interaction and to the resulting effect of many rows hiding away.

5.3 Presentation Phase

All participants were able to develop insights from the data and describe what they found interesting or surprising. While participants were mostly left to speak about their observations, the experimenter occasionally asked questions to clarify a point or make a comment about an observation. All participants generally began by stating which topics or countries they had chosen to focus on, and three participants also described the interactions they had carried out (e.g., locking a group of rows and taking a snapshot). Next, we describe in more detail the styles and strategies that participants used, the themes participants focused on, and participants' movements, gestural behaviours and difficulties.

5.3.1 Presentation Style and Strategies

We observed four key presentation styles: (1) participants cycled through snapshots and described interesting insights within these views, (2) participants described their observations based on a single view, (3) participants interacted with EMERGE to describe observations (e.g., by scrolling to different parts of the data), and (4) participants did not interact with EMERGE and described general insights they had formed during exploration.

Snapshot-Centric Presentation. Nine of sixteen participants described sets of observations by using the snapshot feature to cycle through different views. Participants would show a snapshot with a particular set of organized data points and talk about what is being shown. The next snapshot would typically switch to a different location within the dataset and include a different set of rows that have been grouped together. For instance, P16 started the presentation with all rows except two hidden, pertaining to two topics (importance of good pay in work, and importance of work) and described two examples of countries that were unlike the others on these topics. P16 then pulled a bar (to highlight it) while talking to emphasize the country that placed least importance in work. P16 proceeded to show the next snapshot where again two rows remained unhidden (pertaining to two countries: Russia and Romania) and discussed observations on religion. P16 then showed the third snapshot where four rows had been grouped next to each other and remained unhidden, which showed four major European countries and described that their confidence in government was relatively low. Finally, the fourth snapshot focused on two countries (UK and Switzerland) that remained unhidden and they were compared in terms of how the public perceived their government.

Single-View Presentation. Three participants (P6, P11, and P13) presented their observations based on one single view. All three participants had prepared the views by grouping rows of interest, highlighting data points of interest and by hiding irrelevant data points. Notably, although P11 described insights using a single view, she initially cycled through the snapshots and described the interactions she had carried out in order to reach the final snapshot.

Interactive Presentation. Two participants had grouped a set of rows along one axis and scrolled to different locations along the adjacent axis to describe their observations. P15, for instance, had grouped a set of countries and saved snapshots in different locations along the list of topics, but chose to scroll to the topics of interest. P15 frequently scrolled back and forth depending on the topic being described. Similarly, P5 had grouped a set of countries and sequentially scrolled through the topics and compared them across the grouped countries. P1, P10, and P15 also occasionally pulled a bar to highlight an interesting observation while talking.

Non-Interactive Presentation without Data. Two participants (P3 and P7) did not prepare any views (i.e., no highlighted or hidden data) nor interact with EMERGE during their presentation. Both participants simply described general impressions they had formed during exploration. For example, P3 described that they had found Cyprus to be generally positive about various topics, and that in general immigration as a threat was not high in any of the countries. Both participants made generalized statements similar to "some countries felt a certain way about a certain topic" without supporting their statements with relevant data.

5.3.2 Themes of Presentations

Participants generally selected topics concerning social issues, religion, and politics, and countries that were perceived as major European powers, or as particularly liberal or conservative. Five participants chose to focus on topics (i.e., by grouping them) during their presentation and then cycle through the countries to compare how the topics were perceived in various countries. Similarly, five participants focused on countries, and six participants focused on both countries and topics. Six participants also grouped sets of countries and topics into comparable categories.

Five participants selected topics and countries based on personal experiences and interests, such as countries that participants had lived in, topics that related to their place of work, and personal beliefs. For example, P11 spent time in various Slavic countries and therefore chose to compare them against topics that they experienced during her stay, and P9 selected topics related to politics as they worked in a foreign exchange office. Other participants were more opportunistic when selecting topics and countries. For example, P1 stated that they selected a set of countries for no particular reason.

Most participants (9) provided descriptive observations about the data and seven participants provided some reflection on their observations. Descriptive accounts were related to what participants found surprising and descriptions of topics and countries in relation to each other. For example, P1 grouped five countries (UK, Spain, Norway, Russia, and France) and described that they all had high national pride, and that participation in elections contributed to that. They then moved on to immigration as a threat to society and stated that this was particularly high in the UK and Russia. In contrast, P5

had chosen to compare socio-economic topics with what they regarded as nations that are more powerful and smaller nations, and described that none of the countries feel a duty towards society to have children, which is why she felt that they lacked human resources, and that this was very different to Asian countries.

5.3.3 Movement Behaviour

We observed that participants mainly leaned over the top of EMERGE (0.8 leans per minute) during the presentation phase (Fig. 3-bottom), to inspect the data which they were presenting. Other movements such as walking around the display, head tilts, and crouching were less common. These movements are likely associated with participants reminding themselves of the relationships between the data (e.g., participants would pause while speaking, lean over to look at data points and resume speaking) as well as read the axis labels (i.e., walking around the display to read the countries and topics they had grouped together). For example, P12 initially forgot why two of the four snapshots had been saved, which led to repeated crouching, leaning, and walking behaviours to inspect the data and recall their observations before resuming their presentation. P10 and P11 frequently walked between different sides whose axis labels pertained to topics or the countries that were being discussed (e.g., when talking about topics P11 walked to the side that listed topics).

5.3.4 Gesturing, Pointing, and Touching

Five participants did not interact with EMERGE and mainly used gestures to communicate their findings. Participants mainly pointed (discretely and continuously) at either the bars or the labels using both one finger, and multiple fingers. For instance, a single-finger label point gesture was carried out at 1.8 times per minute (Fig. 3-bottom). The longest continuous gesture that was observed included hovering the palms over the bars (7.7% of presentation time).

Gestures during the presentations were typically used to indicate both single and groups of data points. Notably, P6 made several palm indications on top of the bars to emphasize the dips and arches that had been discovered (in the shape of the bar grid) that showed differences in values between countries. P10 used palm gestures to describe that a group of data points are higher compared to others (which would be followed by a palm hover and palm raise, e.g., see Fig. 1c). P10 grouped topics based on two themes (politics and religion) with three rows on one end, and three rows on the other end (with hidden data points in the middle to create a valley). P10 was comparing and contrasting these themes with different countries and would use both arms and hover palms above each side of the bars.

5.3.5 Difficulties

Four participants had trouble recalling the observations they had saved in the snapshots. This was portrayed through hesitation, verbally questioning why they had saved a particular snapshot, and inspecting the grouped rows or highlighted data points. Recall issues can be attributed to grouping different sets of topics and countries in each snapshot (thus increasing complexity). It is also possible that this is correlated to not hiding irrelevant data and thus having too many data points in one view. For instance, four of the seven participants who grouped different sets of rows in both axes had hidden irrelevant data and had no issues recalling their observations. However, further study is required to examine whether hiding unrelated data lowers cognitive load and allows participants to more easily recall the context of previously created snapshots.

5.4 User Feedback

The post-study interview and questionnaire showed that participants found interactions, such as pulling a bar to highlight it, relatively intuitive (M=3.94, SD=0.85) as shown in Fig. 6. Six out of sixteen participants found it difficult organizing rows (due to the large dataset) and would prefer an easier way of achieving this. In contrast, 15 participants found the 10×10 configuration of bars sufficient, as a larger grid would increase data density and become overwhelming. P14 stated *"actually, what I was constantly trying to do was drill down. I can't process that much in one go"*.

All participants stated that dynamic physicalizations such as EMERGE would be effective for discussions in smaller groups (such as seminars) rather than to a large audience (e.g., in a lecture theatre, where two participants suggested using a camera). Use in smaller groups would also allow the members of a group to be more "hands-on" and interact together with the data. For example, one said: *"I've given presentations as an English as a Foreign Language teacher and I think for lots of learners something like this it's moving, it's changing shape, it's changing colour it's interesting and exciting and makes things more easily memorable."*

Fifteen out of sixteen participants felt that the data was easy to interpret (M=4.5, SD=0.8). Six participants stated that it is more accessible as *"any layman can understand, because if you start putting models and figures it starts becoming difficult, you have to cater to the audience"*. One participant, however, stated that they would not feel comfortable presenting without exact numbers.

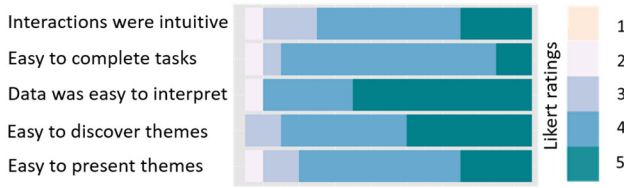

Fig. 6. Likert scale ratings from 1 – Strongly Disagree to 5 – Strongly Agree on the post-study questionnaire.

5.5 Results Summary

In summary, all 16 participants were able to successfully develop and present insights by exploring an unseen dataset. We observed that participants frequently moved around EMERGE during the exploration phase in order to inspect the data from different angles (e.g. leaning over the top and walking between different sides). We surprisingly found some participants made extensive use of hand gestures while inspecting the data (e.g. pointing along a row of bars); these likely aided their thinking process. As participants were more engaged in discussion with the experimenter during the presentation phase, movement around EMERGE was less frequent, but gestures were more common (e.g. to direct the experimenter's attention). Movements that did occur, such as leaning over the top, were likely indicative of reaffirming an observation. We found that although system interactions related to organization were most common during exploration (scrolling, locking, swapping), physical interaction with the bars (e.g. pulling to highlight) were also relatively frequent. It was also clear that none of the participants explored the entire dataset (with 2 participants coming close), which is possibly related to viewport size (10×10) or the ease of navigating through the data. Different types of data organization techniques were observed during exploration, such as grouping blocks of data in different ways (swapping vs. locking) and comparing grouped blocks or showing prominence by hiding surrounding irrelevant data. Presentations were structured according to saved snapshots, a single view of data points currently visible, interactive presentations (e.g. they scrolled to different parts of the data), and non-interactive (i.e. they spoke about general impressions).

6 Discussion

Our findings provide promising evidence that physicalizations encourage people to engage in data exploration, to support them in data-based presentations, and have the potential to support thinking about data in ways that are different from non-physical

visualizations. Here, we reflect on the key findings and identify avenues for further work.

6.1 'Thinking Actions'

Previous work on static data physicalizations by Jansen et al. found that people made extensive use of their hands to break down information retrieval and processing tasks into simpler physical actions [17]. We found, unsurprisingly, a high frequency of hand gestures during the presentation phase. More surprising was a considerably high frequency of hand gestures during the exploration phase (see Fig. 3). In contrast to Jansen et al. our participants were engaged in an open-ended task ("find something interesting in this data"), thus the gestures we observed served a large variety of purposes, and we cannot attribute them easily to specific task goals. Nonetheless, our observations do confirm that people frequently use their hands while engaging with data. An interesting question now is whether this is specific to *physical* data representations or whether people perform some form of hand gestures also with non-physical visualizations. It is so far unclear how people interact in the physical world when analyzing on-screen visualizations. Do they point and gesture at screens in a comparable fashion? Some prior work suggests that such behaviour can be observed with mouse pointers [3]. Does it serve similar purposes? What are the conditions under which such behaviour occurs?

6.2 Perception of 3D Visualizations

Besides hand gestures, we observed a high frequency of body movements to interrogate the visualization, both in the exploration and presentation phases. This might be interpreted to mean that body movements, such as crouching and tilting one's head, contribute to exploring data. Another interpretation is that body movements turn harder perceptual tasks into simpler ones. Prior work found that people are able to make reasonably accurate estimates (+/- 7%) of size differences between 3D bars without much movement [19]. However, comparisons between isolated bars with clearly visible baselines are easier than with bars that are partially occluded by surrounding ones. It remains an open question as to whether body movements are correlated with better insights from the data, if they allow participants to read data more accurately, or if they compensate for occlusion. Nevertheless, such movements highlight the importance of designing physicalizations that allow inspection from multiple perspectives and placing them in locations that supports this behaviour (e.g. placing the system against a wall would restrict movement). Future work will need to investigate this question, e.g., by using experimental datasets designed to cause differing amounts of occlusion.

6.3 Interacting with Physical Data

Dynamic physicalizations provide the ability to go beyond the touch-screen and introduce tangible controls. While many of the system interactions were supported by the touch screen, we were interested in how participants responded to physical controls (e.g., pressing to hide rows). We observed that physical interaction with the bars were moderately frequent and that all participants were generally confident while doing so. This suggests that they were happy in temporarily adjusting the bar values to trigger a function (e.g., P1 would highlight an entire row by pulling the bars in quick succession), thus creating an opportunity to further explore physical controls for manipulating data. In particular, we found that the bars around the edges were pressed and pulled the most, and can therefore be useful candidates for control mechanisms. We only observed hesitation before participants carried out concurrent bar presses to hide data. This is likely related to participants ensuring their selection and to prevent hiding data that they are interested in.

6.4 Scalability

Data physicalizations are, at least with current available technology, fixed in size – they have a fixed amount of concurrently visible data. The size of the dataset used in our study purposefully exceeded the viewport size to test how this would affect people's exploration behaviour. Our findings show that participants found this challenging (7 of 16). For instance, P2 stated that there was too much data to take in and decided to focus on a subset. Fig. 4 shows that only a few participants looked at most of the dataset. At the same time, 15 of 16 participants stated that the size of the viewport (10×10) is sufficient and that a larger grid would become overwhelming and unmanageable. Although it is contradictory that no participant explored the entire dataset, and that they felt the grid size was adequate; perhaps the open-ended nature of the task, the unfamiliarity of the physicalization, or the system's ability to facilitate navigation of the dataset, caused participants to explore limited parts of the data. For instance, six participants wanted more efficient ways of grouping and navigating through the data (e.g. P6 suggested scrolling methods similar to an address book on a Smartphone) and quicker ways of dragging row labels from one end of the dataset to the other. This highlights the importance of providing controls that allow faster ways of data navigation and organization. Furthermore, no context visualization for the data surrounding the viewport was available apart from the scrollbars on the side-screens. It would be useful to explore on-screen visualization techniques, such as the focus and context techniques [6] and whether they can be applied in to data physicalizations.

6.5 Limitations

Our investigations are based on a configuration with specific hardware and interaction capabilities. Many different designs are possible and it is unclear in how far factors such as the physical height of the overall device, the height of the individual bars, the distance between bars, or the speed with which they move might affect how people approach, perceive, and interact with a physicalization. Nonetheless, we believe that our observations provide useful insights into how people engage with physically dynamic data, which can be used as a starting point in further developments and investigations of dynamic physicalizations.

Participants were also limited to a lab-based setting and, although we included a demonstration and training phase to build proficiency, participants were neither expert system users nor visualization experts. However, given the timeframe in which participants carried out their tasks, their ability to understand the data and to extract and present these themes provide a promising outlook.

Our study was not a comparative experiment, and we therefore cannot draw any conclusions about how visualizations and physicalizations compare. The exploratory nature of the study, however, is a necessary step in order to provide a foundation for such comparative studies. The fact that all participants were able to engage with the data and talk about their observations by using an unfamiliar physicalization prototype is encouraging for further work.

7 Conclusion

The objective of this paper was to observe and understand users' movements, gestures, system interactions, and usage strategies with a dynamic physicalization. This was achieved through a user study where participants explored an unknown dataset and then presented interesting observations. We provide an in-depth report and characterize the above factors in order to set a starting point for the development of dynamic physicalizations into a usable and ultimately, more useful method of data presentation.

Acknowledgments

This work was supported by the FET Open Scheme's GHOST project (grant #309191) and the EPSRC's MORPHED project (grant #EP/M016528/1).

REFERENCES

[1] R. Aigner, D. Wigdor, H. Benko, M. Haller, D. Lindbauer, A. Ion, Z. Shengdong, J. T. Kwan, and J. T. K. V. Koh. Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for HCI. Microsoft Research TechReport MSR-TR-2012-111, 2012

[2] M. Brehmer and T. Munzner. A Multi-Level Typology of Abstract Visualization Tasks. IEEE Transactions on Visualization and Computer Graphics, 19(12), pages 2376-2385, 2013.

[3] S. Breslav, A Khan and K. Hornbæk. Mimic: visual analytics of online micro-interactions. In Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces (pages. 245-252). ACM, 2014.

[4] C. Brown and A. Hurst. VizTouch: Automatically Generated Tactile Visualizations of Coordinate Spaces. In Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction (TEI '12), Stephen N. Spencer (Ed.). ACM, New York, NY, USA, pages 131-138, 2012.

[5] J. Cohen. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. Psychological bulletin, 70(4), 213, 1968.

[6] A. Cockburn, A. Karlson and BB. Bederson. A review of overview+ detail, zooming, and focus+ context interfaces. ACM Computing Surveys (CSUR), 41(1), page 2, 2009.

[7] P. Dragiecevic and Y. Jansen. "List of physical visualizations", http://dataphys.org/list/hans-rosling-adopts-physical-visualizations/, 2013.

[8] EVS. 2011. European Values Study 2008: Integrated Dataset (EVS 2008). GESIS Data Archive, Cologne. ZA4800 Data file version 3.0.0.

[9] S. Follmer, D. Leithinger, A. Olwal, A. Hogge and H. Ishii. inFORM: Dynamic Phsyical Affordances and Constraints through Shape and Object Actuation. In Proceedings of the 26th annual ACM symposium on User interface software and technology (UIST '13). ACM, New York, NY, USA, pages 417-426, 2013.

[10] A. Gillet, M. Sanner, D. Stoffler and A. Olso. Tangible Interfaces for Structural Molecular Biology. Structure 13(3), pages 483-491, 2005.

[11] I. Gwilt, A. Yoxall and K. Sano. Enhancing the Understanding of Statistical Data Through the Creation of Physical Objects. In DS 73-1 Proceedings of the 2nd International Conference on Design Creativity, Volume 1, pages 117-126, 2012.

[12] J. Hardy, C. Weichel, F. Taher, J. Vidler and J. Alexander. ShapeClip: Towards Rapid Prototyping with Shape-Changing Displays for Designers. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15). ACM, New York, NY, USA, pages 19-28, 2015.

[13] T. Hogan and E. Hornecker. In touch with space: embodying live data for tangible interaction. In Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction (TEI '13). ACM, New York, NY, USA, pages. 275-278, 2013.

[14] T. Hogan and E. Hornecker. How Does Representation Modality Affect User-Experience of Data Artifacts? Haptic and Audio Interaction Design, Springer, pages 141-151, 2012.

[15] S. Huron, Y. Jansen and S. Carpendale. Constructing Visual Representations: Investigating the Use of Tangible Tokens. IEEE Transactions on Visualization and Computer Graphics, 20(12), pages 2102-2111, 2014.

[16] S. Houben, C. Golsteijn, S. Gallacher, R. Johnson, S. Bakker, N. Marquardt, L. Capra and Y. Rogers. Physikit: Data Engagement Through Physical Ambient Visualizations in the Home. CHI 2016 - Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2016.

[17] Y. Jansen, P. Dragicevic and J.-D. Fekete. Evaluating the Efficiency of Physical Visualizations. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13). ACM, New York, NY, USA, pages 2593-2602, 2013.

[18] Y. Jansen, P. Dragicevic, P. Isenberg, J. Alexander, A. Karnik, J. Kildal, S. Subramanian and K. Hornbæk. Opportunities and Challenges for Data Physicalization. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15). ACM, New York, NY, USA, pages 3227-3236, 2015.

[19] Y. Jansen, and K. Hornbæk. A psychophysical investigation of size as a physical variable. IEEE Transactions on Visualization and Computer Graphics 22, no. 1, pages 479-488, 2016.

[20] R. A. Khot, J. Lee, D. Aggarwal, L. Hjorth and F. Mueller. TastyBeats: Designing Palatable Representations of Physical Activity. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15). ACM, New York, NY, USA, pages 2933-2942, 2015.

[21] J.R. Landis and G.G. Koch. The measurement of observer agreement for categorical data. Biometrics 33, pages 159-74, 1977.

[22] W. S. Lasecki, M. Gordon, D. Koutra, M. F. Jung, S. P. Dow, and J. P. Bigham. Glance: rapidly coding behavioral video with the crowd. In Proceedings of the 27th annual ACM symposium on User interface software and technology, pages 551-562, 2014.

[23] D. Leithinger and H. Ishii. Relief: a scalable actuated shape display. In Proceedings of the fourth conference on Tangible, Embedded, and Embodied Interaction, pages 221–222, 2010.

[24] B. Nissen and J. Bower. Data-Things: Digital Fabrication Situated within Participatory Data Translation Activities. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pages 2467-2476, 2015.

[25] R. H. Price and D. L. Bouffar. Behavioral Appropriateness and Situational Constraint as Dimensions of Social Behavior. Journal of Personality and Social Psychology, 30(4), pages 579–586, 1974.

[26] M.K. Rasmussen, E.W. Pedersen, M.G. Petersen and K. Hornbæk. Shape-changing interfaces: a review of the design space and open research questions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 735–744, 2012.

[27] T. Regan, D. Sweeney, J. Helmes, V. Vlachokyriakos, S. Lindley and A. Taylor. Designing Engaging Data in Communities. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15). ACM, New York, NY, USA, pages 271-274, 2015.

[28] S. Stusak, A. Tabard, F. Sauka, R. A. Khot and A. Butz. Activity Sculptures: Exploring the Impact of Physical Visualizations on Running Activity. IEEE Transactions on Visualization and Computer Graphics, 20(12), pages 2201-2210, 2014.

[29] F. Taher, J. Hardy, A. Karnik, C. Weichel, Y. Jansen, K. Hornbæk and J. Alexander. Exploring Interactions with Physically Dynamic Bar Charts. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15). ACM, New York, NY, USA, pages 3237-3246, 2015.

[30] A. Tang, M. Tory, B. Po, P. Neumann and S. Carpendale. Collaborative Coupling over Tabletop Displays. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06), Rebecca Grinter, Thomas Rodden, Paul Aoki, Ed Cutrell, Robin Jeffries, and Gary Olson (Eds.). ACM, New York, NY, USA, pages 1181-1190, 2006.

[31] A. S. Taylor, S. Lindley, T. Regan, D. Sweeney, V. Vlachokyriakos, L. Grainger and J. Lingel. Data-in-Place: Thinking through the Relations Between Data and Community. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15). ACM, New York, NY, USA, pages 2863-2872, 2015.

[32] A. Vande Moere and S. Patel. The Physical Visualization of Information: Designing Data Sculptures in an Educational Context. Visual Information Communication, Springer, pages 1-23, 2010.

[33] J. Zacks, E. Levy, B. Tversky, and D. J. Schiano. Reading bar graphs: Effects of extraneous depth cues and graphical context. Journal of Experimental Psychology: Applied, 4(2), page 119, 1998.