Large Displays and Tablets: Data Exploration and its Effects on Data Collection

Katerina Gorkovenko[1,2], **Lars Lischke**[2], **and Paweł W. Woźniak**[2,3]
[1]DJCAD, University of Dundee, Dundee, United Kingdom, k.gorkovenko@dundee.ac.uk
[2]University of Stuttgart, Stuttgart, Germany, lars.lischke@vis.uni-stuttgart.de
[3]Utrecht University, Utrecht, the Netherlands, p.w.wozniak@uu.nl

ABSTRACT
Data is pivotal to open government initiatives, where citizens are often expected to be informed and actively participate. Yet, it can be difficult for people to understand the meaning of data. Presenting data to the public in an appropriate way may also increase citizen's willingness to participate in data collection. Here we present a study which explores how large screens can support socially relevant data exploration. In a between subject laboratory experiment, we analysed how pairs of participants explored data visualisations on a high-resolution display (LHRD) and a tablet. Our results indicate that LHRDs are less cognitively demanding, while tablets offer more shared control of the interface. Data exploration had limited effect on increasing comfort with sharing personal data but helped increase perceptions of trustworthiness within the data collection process. We observed that appropriately visualised data on either platform has significant potential to increase the public's understanding of large data sets.

Author Keywords
LHRD; tablet; data visualisations; collaboration; data exploration; participation; data collection; Spotfire.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
In the age of big data, governments and healthcare institutions are adopting Open Data models, which give others access to the data they have gathered. This transparent approach to gathering and distributing data allows them to maintain perceptions of trustworthiness and in turn increase public participation [5, 23, 36]. Incentivising the public to take part in the collection of openly available health data can have a positive impact on the research of the causes of complex diseases, such as cancer [36]. Although, vast amounts of data are available online, it is not yet clear how this data can be presented to the public in a meaningful way, and how communities as well as individuals can benefit from the data. State of the art visualisation software can help the exploration of data but may not necessarily lead to precise understanding of its meaning. Democratising data and giving access to data can be empowering, but it is essential that people can understand the data and assess it for their own goals. This is especially important when it comes to health related scientific data, were the vision is that broad access to the information can have lasting implications for people. Providing access to information about risks and risk factors, such as the spread of flu in winter, or the impact of health risks due to behaviour, such as smoking, based on comprehensive data may positively facilitate social awareness.

NordiCHI'18, September 29–October 3, 2018, Oslo, Norway

ACM ISBN 978-1-4503-6437-9/18/09...$15.00
https://doi.org/10.1145/3240167.3240192

This research project is part of a wider open government initiative, which aims to create computational models to help predict and reduce the spread of diseases and understand how contiguous phenomena are linked to mobility and human behaviour. A large segment of the project is investigating ways to visualise and explain data to the public in order to facilitate social response and participation. Current technology offers a range of solutions, which can help spread information and facilitate understanding. Data visualisation software is becoming increasingly more accessible and interactive. Furthermore, the technology that can be used to present visualised data now ranges from a smartphone to large displays many meters across. As the prices of technology are continuously decreasing and the uptake of large screens in offices and homes (e.g. for TV) increases, it is apparent that LHRDs will become commonplace in the near future. It has not yet been explored how technology with varying interactivity and screen size affect the ability of the public to explore big data and collaborate in understanding the meaning of large data sets. We were further interested in the effect that personal experience with data exploration has on the participant's willingness to contribute their own data in the future. The study was designed to address the following questions:

1. How is collaboration and sense making affected by the exploration of data on a LHRDs vs a tablet?

2. In what ways does data exploration impact the willingness of participants to take part in data collection?

Through a between subject lab study we investigated how collaborative data exploration of world health data visualisations on a tablet vs a LHRD affected the cognitive demand, perceived usability and collaboration of completing a set of tasks. We further explored to what extent interacting with the visualisations influenced the participant's comfort with sharing their personal health data. We primarily contribute to the study of large displays and their potential to deliver complex information in a less cognitively demanding way. We further contribute insights into the perceived benefits of using data visualisations as part of the data collection process.

BACKGROUND

Here, we present the motivation behind our work and past research that inspired us to conduct our inquiry.

Open Government Data

The phrase open government is used to describe initiatives where data collected by government institutions is made available online in order to foster openness, transparency, participation and accountability within citizens [5, 23, 27]. Such openly available data can help people make more informed decisions [28]. HCI literature has identified that open government data can also encourage citizens to actively participate [4]. It is not yet clear exactly how open data facilitates and incentivises participation.

The view that government transparency can lead to a more engaged society has led to the development of digital tools to empower the public. In October of 2013 Brazil experienced its first parliamentary hackathon where developers were invited to make use of the data provided by the government in order to increase transparency and engagement [11]. The event resulted in a mobile application, which allowed its users to monitor all proposed laws, commissions, constitutional amendments and parliamentarian's voting histories [11]. Although such initiatives have been widely successful we are yet to establish what aspects of them engage the public. An Austrian app for political participation was downloaded 780 times, but when investigating the public's incentives for using it, it was discovered that game aspects within the application did not significantly influence participation [37]. There is still room within HCI to investigate what interaction techniques incentivise and engage the public.

Health Data

Although past open government initiatives have largely focused on information such as parliamentary voting histories [3, 11], health data has been identified as a key area for future tool development [21]. Generating large open health datasets, such as UK Biobank have the potential to empower research into the causes of complex diseases [36]. Although sharing health data may be useful for a variety of stakeholders, such as researchers, governments, and health-care providers, there are intricate regulations and perceptions, which shape the ways the public can share health data. The European Data Protection Directive has leveraged data privacy of individuals, which allows them to enforce access restrictions on the data they share, this type of regulation creates a strong demand for balance between security and utility of data [13]. Sharing health data is especially important between patients and healthcare providers, where such data can dramatically increase the number and quality of insights that healthcare providers are able to deliver [18]. Meanwhile, current electronic health records fail to support documentation and collaboration between healthcare-professionals [24]. Within clinical data sharing, individuals often report that they want higher levels of transparency and control over the way their data is being used [1]. Meanwhile reassuring them that their health data is used in research that would potentially make a difference for others can have positive effects on the quality of the data itself [1]. Engaging people in data sharing.

There are increasing numbers of initiatives, which encourage data collection directly from citizen's. Public spaces have often been identified as appropriate locations to engage directly with the public. Whether through the use of projections or large displays, research within data collections in public spaces often utilises large visuals in order to gain the public's attention [20, 17]. In a project, which aimed to engage a small community with data collection through the use of distributed voting and visualisations in several shops within a small community, it was observed that the large information visualisations combined with simple input technology can encourage more wide-ranging engagement. This indicates that large displays containing data visualisation have potential to engage the public in data sharing.

Within the sharing of health data, research has identified that patients experience issues when collecting, storing, accessing and sharing data with their clinicians [32]. Research into the development of tools, which enable the sharing of data between patients and health-care professionals, indicate that patients often feel overwhelmed by the data made available to them [29]. This indicates that future data sharing tools need to not only address the factors that incentivise data sharing, but how data should be shared back to patients in a meaningful way.

Interactive Collaborative Data Exploration

With the rise of initiatives encouraging citizens to contribute to data collection, data exploration becomes more important. Weise et al. [41] argues that non-experts have to be able to manage and understand ubiquitous sensing technology, which may be part of future participation initiatives. In line with this, Churchill [10] calls for novel techniques to explore large data sets interactively. Such data exploration can be facilitated through advanced interactive visualisations, which help users explore and derive meaning from data [30]. Mayer et al. [25] remarked that users move differently in space depending on the goal of their interactive task. For mobile data exploration scenarios, research emphasises the benefits of using multiple tablets [41]. Beyond mobile technology large display space has been shown to have a positive influence on data exploration [2]. Furthermore, Andrews and North [3] argued that the spatial representation of information supports the exploration process.

Within the field of technology for learning studies show that tabletop interfaces may be advantageous for the teaching of complex concepts [34, 35]. Schneider et al. discovered that a tabletop interface fosters collaborative learning [34], while Shaer et al. found interactive tabletop learning applications are less cognitively demanding, than traditional screen and mouse setups [35]. Furthermore, Shaer et al. discovered different collaboration dynamics that could increase learning in pairs when both participants are actively engaged in the activity, such as when they take turns exploring, or when one acts as a navigator instructing his partner [35]. These works were centred around the development of novel interfaces and compared the use of new, tailor-made software to a baseline system. In contrast, our work uses the same software in both conditions, thus aiming to unpack the qualities stemming specifically from the interaction modality.

Data exploration on LHRDs and tablets

Large high-resolution displays (LHRDs) have potential to help empower and support people in the process of exploring vast amounts of visual data. In a paper-based study using a white board as a display, it was discovered that large spaces can support collaborative data analysis tasks [19]. Yost and North [45] presented participants with a map populated by graphs, they discovered that even when there was a large number of visualisations the participants were still able to make meaningful connections between the data without being overwhelmed. Furthermore, the extra display space afforded by LHRDs can help participants within map navigation tasks to be faster and to interact in a physical rather than a virtual manner with the data [6]. Interactive technology provides effective support for users reflecting upon data in the process of sensemaking. Goyal et al. [14] showed how implicit sharing can help users in sensemaking over a distance using a desktop interface. RAMPARTS [42] illustrated how mobile devices allowed for more effective sensemaking than a tabletop interface and Thaddeus [43] showed that effective interaction with information visualisation on tablets was possible. On the other hand, Mayer et al. [26] observed that mobile devices can also disrupt discussions. This implies that tablets need to be studied further in the context of group data exploration. Thus, we were inspired to use a tablet in our investigation. Further, Wallace et al. [40] showed that extensive, reconfigurable display space may benefit collaboration in sensemaking. This motivated us to investigate if those properties could be translated to the vast screen space of LHRDs.

METHOD

The study investigates the effect of an LHRD and tablet interface on collaborative data exploration. It further explores, the ways in which data exploration on these interfaces aids the participant's comfort and willingness to contribute their personal health data. The study adopts a mixed method, between subject approach and took placed in a laboratory lab setting. Fifty-six participants were recruited through university emailing lists. They were grouped into pairs and asked to explore a set of health data through Spotfire, a state of the art visualisation and analysis tool, on either a tablet or an LHRD. The grouped participants were video and audio recorded during the data exploration activities.

The tablet condition used the Spotfire IOS app, and the LHRD condition used the Spotfire web tool. The pages were populated by three types of visualisations: maps, tree-maps and scatterplots (See Figure 1). Each dot within the scatterplots or square within the tree-maps was representative of a different country from the dataset. The pairs received a short training task in order to learn how to interact with the interface. Both interfaces contained the same six data exploration pages. Both the tablet and LHRD interface had the same main interaction techniques. These included a lasso tool to select parts of the visualisations and a filtering tool to narrow down the visible content based on selected parameters. The LHRD condition utilised a mouse and keyboard, while the tablet had a touchscreen interface. Sections of the data could be selected, which highlighted the relevant countries in all other visualisations. Before seeing the data, each participant was asked to fill out a questionnaire containing personal health data, which related to the visualisations they were going to explore. They were asked how comfortable they felt with sharing their health data with us once before and once after the collaborative tasks.

The participants were then asked to explore the data inthe six task pages and verbally communicate what they understood from the data. All of the visualised health data was collected from the gapminder.org. The tasks involved open ended explorative questions, which required the participants to interact with the visualisations, for example "What possible explanation is there for the relationship between cancer cases and government health expenditure?". Each task had its own page populated with relevant visualisation ranging from two (see Figure 1 F) to six (see Figure 1 C).

The participants could explore the data and discuss it as long as they wished in order to reach an agreed solution to each task. The data visualisation pages focused on different themes, which aimed to spark discussion and exploration about potential relationships between government health expenditure, cancer, income and health risks. They explored:

1. Government health expenditure per country vs health risks, such as smoking (Figure 1 A).
2. Government health expenditure vs female colon and rectum cancer and female breast cancer (Figure 1 B).
3. Health risks vs different types of cancer (Figure 1 C).
4. Government health expenditure vs income per person (Figure 1 D).
5. Female cervical cancer vs income per person (Figure 1 E).
6. Female cervical cancer vs male colon cancer vs income (Figure 1 F).

When all six tasks were finished the participants were asked to rate the usability of the interface using a SUS questionnaire [7, 9], rate the workload required to complete the tasks using the raw NASA TLX questionnaire [16] and a collaboration questionnaire created by Goyal et al. [14, 15]. We used a between subject approach in order to accurately compare NASA TLX, collaboration questionnaire, SUS, and TCT scores.

Following the questionnaires, the participants each took part in a 10-minute semi-structured interview, which explored their collaboration, understanding of the data, and thoughts on sharing their own health data. The interviews were conducted

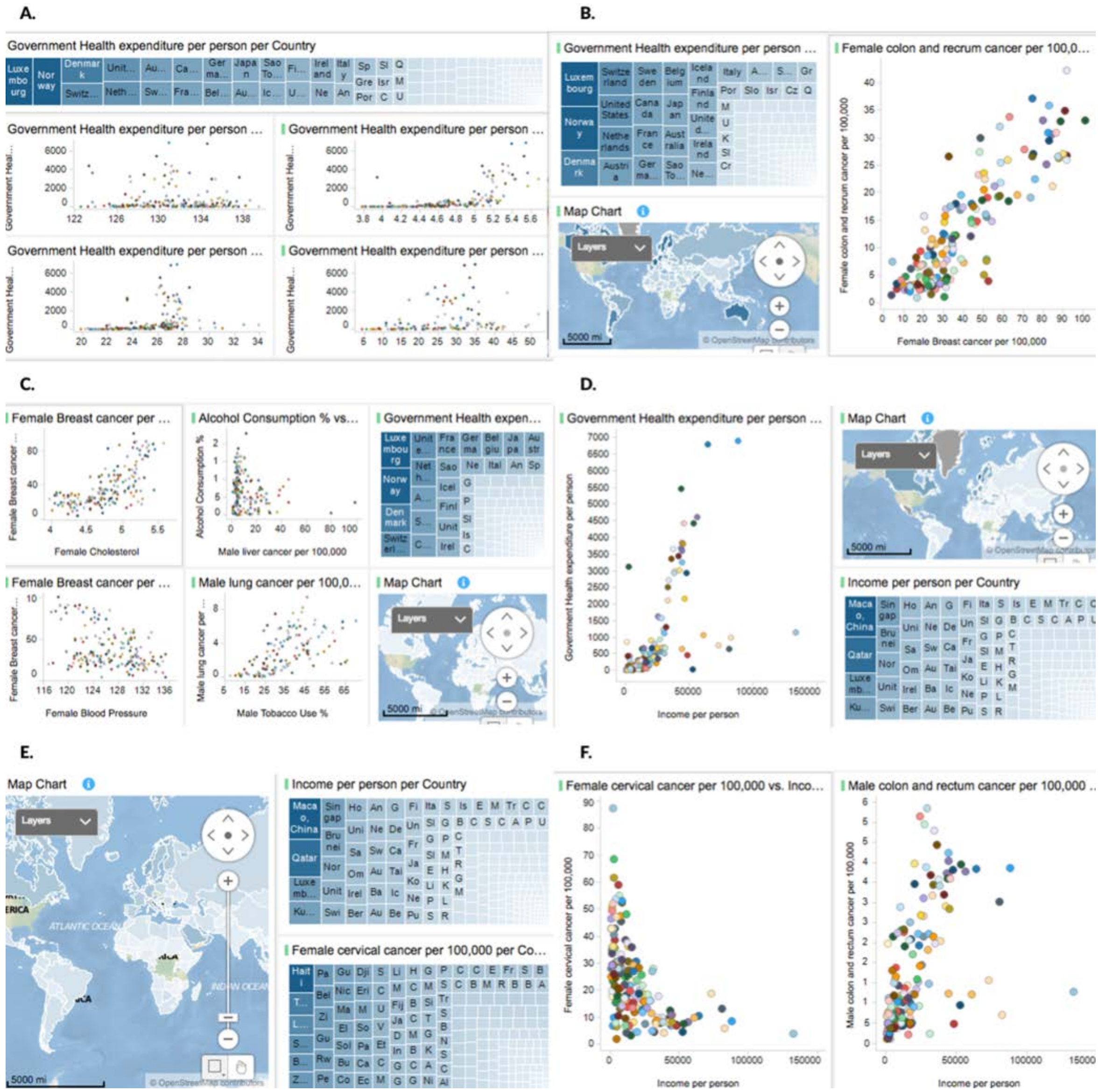

Figure 1. The 6 different data exploration pages, organized in the ordered that the participants explored them in.

individually rather than in pairs in order to allow them to express their personal opinion on how taking part in the study influenced their view on sharing their individual health data with us. All the audio material from the interviews was transcribed. It was thematically analysed [8] together with the videos from the collaboration in the qualitative data analysis tool Atlas.ti. The lead researcher coded the transcripts and videos within Atlas.ti. The codes were then refined and clustered into the three themes presented in the results section.

Participants

The 56 participants ranged from 19 to 35 years of age, 15 were female and 41 male. They received ten euros each as a reward for taking part. The relationships between the participants in each pair ranged from long lasting friendships to complete strangers.

Equipment

For the tablet study sessions, we used a 9.7-inch iPad A1822. For the LHRD sessions we used six 50” 4K Panasonic TX-50AXW804 screens in portrait mode. This resulted in a display with a size of approx. $4.02m \times 1.13m$ (see Figure 2). All three displays were driven by one Microsoft Windows 10 workstation. The study sessions were recorded using two cameras, a Panasonic HC-V520 positioned in order to view the interface the participants were interacting with and a GO-PRO Hero3+ positioned in front of the participants to capture their collaboration and discussion. The post-interviews were recorded using a Tascam DR-40.

RESULTS

The quantitative data indicates that LHRDs are less cognitively demanding than tablets when completing data exploration tasks collaboratively. The most significant difference was

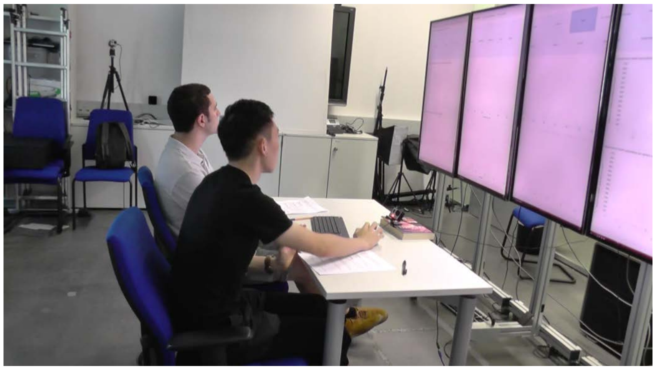

Figure 2. Participants interacting with visualisations in the LHRD condition.

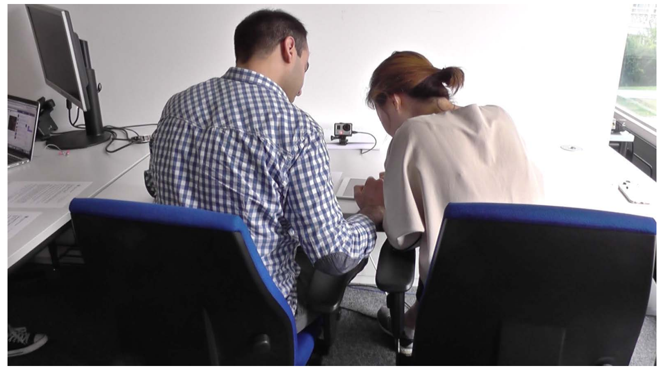

Figure 3. Participants interacting with visualisations in the iPad condition.

observed in the amount of effort and mental demand required to complete the tasks. The video recording and interviews indicated that the use of a tablet allowed for more shared control of the interface.

We found a small but significant positive change in comfort levels after the participants interacted with the visualisations. On average the they felt fairly comfortable sharing their personal health data both before and after, and twenty-five percent reported that they felt an increase in comfort. Despite the fairly low rates of increased comfort levels, the majority of participants felt that sharing data in the form of visualisation made the research team seem more trustworthy as data collectors. Furthermore, the majority of participants indicated that they felt comfortable sharing a range of health data with the research project in the future.

Cognitive Demand and Usability

The LHRD scenario required significantly less mental demand and effort, which made the LHRDs less cognitively demanding overall. The interviews indicate that the two interfaces resulted in both different usability issues and styles of collaboration, which may have influenced the NASA TLX results. Although the participants found it easier to complete the tasks using the large displays, this did not result into significantly different scores for the usability of the system. We conducted a two-sample t-test to compare the perceived workload rated on the NASA TLX questionnaire. The analysis revealed a significant lower perceived workload in the LHRD condition ($M = 5.92$; $SD = 2.52$) than in the tablet condition ($M = 7.20$; $SD = 2.07$); $t(54) = -2.08$, $p < 0.05$. The comparison of the single items of the NASA TLX revealed statistically significant differences for mental demand, $t(54) = -2.511$, $p < 0.05$, and effort, t54 = -2.231, $p < 0.05$ (see Figure 4). To compare the rating of the system usability, which utilised a SUS questionnaire we performed a Wilcoxon rank sum test with continuity correction. The test revealed no statistically significant differences between the LHRD ($M = 74.93$; $SD = 13.80$) and the tablet (M = 72.77; SD = 16.12) condition, $W = 402.5$, $p > 0.05$.

It is difficult to ascribe the lower perceived workload in the LHRD scenario to a specific aspect of the experience. Overall feedback for both interfaces was quite positive but there were different issues that arose in each scenario. Within the LHRD scenario some users liked the large display size because it offered large visuals and a good overview, while others disliked it because it was difficult to find the mouse and it felt too big.

P45: The bigger picture makes it easier to solve the task but at the same time, some time is wasted to navigate, just to find the mouse.

P38: It was helpful that we have a big interface in front of us so that we could see everything next to each other. But for me, the interface was a bit too big.

P36: It aided in understanding the data because it was quite big number of graphs. Quite big amount of points within those graphs.

Within the tablet condition some participants found the screens with higher number of visualisations 'daunting'. At times, the application was not very responsive, and using the selection tool took a bit of practice to get used to. Selecting a section of the visualisations required the user to place their finger on the screen and wait for a grey circle to appear around it, which indicated that the lasso tool had been activated. Beneficially the tablet condition 'helped in the collaboration' by allowing both participants to interact with it by taking turns.

P8: When it comes to pinpointing a certain country it may be harder with the smaller graphs, but for seeing the relationships it was fine.

P12: The UI it helped us to work together, and while one selected the groups of countries or the values, the other one looked at the other graph to see where the correlation is.

Collaborative Data Exploration

When describing the process of working together participants spoke about discussion, data exploration and sharing information. The data displayed in the visualisations was often inconclusive and required the participants to spot trends and relationships between datasets. Their lack of background knowledge and the opinion-based nature of the tasks meant that there were disagreements over the interpretation of the data in many of the pairs.

P30: We didn't maybe think the same about some things that required some previous knowledge, because we were both *a little unsure about this. Both of us had different opinions about it.*

Regardless of the scenario they experienced, all of the participants felt that they understood the data to some extent. The most common mistake the participants made was confuse cause and effect within the data. An example of this is that within the page which looks at the relationship between government health expenditure and two different types of female cancer it seems that countries with increased spending also have more cases of cancer (Figure 1 B). Some of the participants interpreted this as meaning that the countries have to spend more on health because they have more cases of cancer, rather than acknowledge that there may be other reasons for this relationship, such as better testing and screening methods.

As the participants progressed through the tasks many were able to identify that they lacked some background knowledge in order to fully understand what the data meant. Others realised that cause and effect were not clear and that the data in itself is not enough to reach a concrete conclusion. Overall their insights were critical and astute. They were able to spot trends and evaluate the meaning of the visualisations in front of them. Most importantly they were able to spot when either the data seemed incomplete or their own knowledge was insufficient.

P38: I think we needed more facts. Each of us had correct intuitions, and each one of us tried to find the more logical reasons for what we were seeing, so we needed more information to research more, in order to reach what was actually there.

The interviews indicated that the interface of the device the participants used for the completion of the tasks had an effect on their collaboration. The tablet interface allowed for more shared control, 66% of the participants in that scenario reported that they were taking turns to explore the data. The touchscreen interface was located between both participants (see Figure 3) and there may have been more opportunity to take turns. P16 felt that it was 'really balance' and P8 felt that 'each one was working freely with the application'.

In contrast, only 15% of the participants in the LHRD scenario reported that they shared control of the interface. This may have been due to the fact that to interact with the visualisation, the participants had to use a more conventional mouse and keyboard setup (see Figure 2), where it may be easier to take hold of the mouse rather than pass it back and forward. Some of the participants saw this in a positive light like P26 who said that 'since there's only one mouse pointer, only one of us could control the mouse and select stuff. But I also think that's a good thing, because if both started interacting it would just mess everything up'. The participant that did not take control of the mouse could then either contemplate on the meaning of the data or navigate the actions of their partner, as in the case of P36 who recalled:

My partner used the mouse mostly and I spent more time looking at it and just thinking about it and saying things like 'try it like this'.

Although, we observed a lot more unequal control in the LHRD scenario it did not result is a difference in the collaboration survey results. We conducted a two-sample t-test to compare the perceived collaboration scores between the LHRD (M = 85.71; SD = 10.97) and tablet (M = 83.60; SD = 12.36) scenarios and found no statistically significant difference $W = 420.5$, $p > 0.05$. We also conducted a two-sample t-test to compare the overall TCT. The statistical analysis revealed no significant differences between the LHRD ($M = 30min40s$; $SD = 8min52s$) and tablet ($M = 26min40s$; $SD = 7min12s$) condition, $t(54) = 1.87$,$p = 0.07$ (non-significant).

Personal Health Data Sharing

The participants rated how comfortable they felt sharing their health risk data with us before and after interacting with the visualisations. They were asked about their weight, height, if they smoke, drink, exercise and if they have any serious health conditions. They rated their comfortability on a likert scale from 1 being not comfortable at all to 7 being extremely comfortable.

The results indicate a significant positive change in how comfortable participants felt with sharing their personal health data. We used a two-sample t-test to compare the participant's comfortability with sharing their individual health data before and after interacting with the data. Because the health data they shared with us was personal and often participants in the same pair had different comfortability levels, we judged that the analysis of the data should not be dependent on the pairing of the participants. On average, they felt fairly comfortable sharing their data before ($M = 5.21$; $SD = 1.66$), but their comfortability significantly increased after interacting with the data ($M = 5.57; SD = 1.62$), $t(54) = -0.36$, $p < 0.05$. Twenty-five percent of participants changed their level of comfort after the study. There was no significant difference between the tablet and LHRD scenarios.

Participants identified several reasons for their increase in comfort after interacting with the data. The most cited reason was that interacting with the visualizations gave them insight into the importance of the dataset itself.

P36:I think this data could be used to find how should governments spend on health and to better understand the demographics of certain diseases, and, which segment of society they effect more, and why. And it would be, in turn, easier to address those issues

Due to the questions arising about the accuracy of the trends that could be observed in the interface, the participants also saw that more data may increase the accuracy of the dataset. P42 reflected that:

I felt it was important how much data researchers have, and if my data is also a part of the research it can make better results.

Finally, the participants felt that the research offered a 'practical use for the information' they shared. The majority of the participants who did not experience a change in how comfortable they felt sharing their personal data, said that they felt that the data was neither 'that intimate', nor as P15 expressed

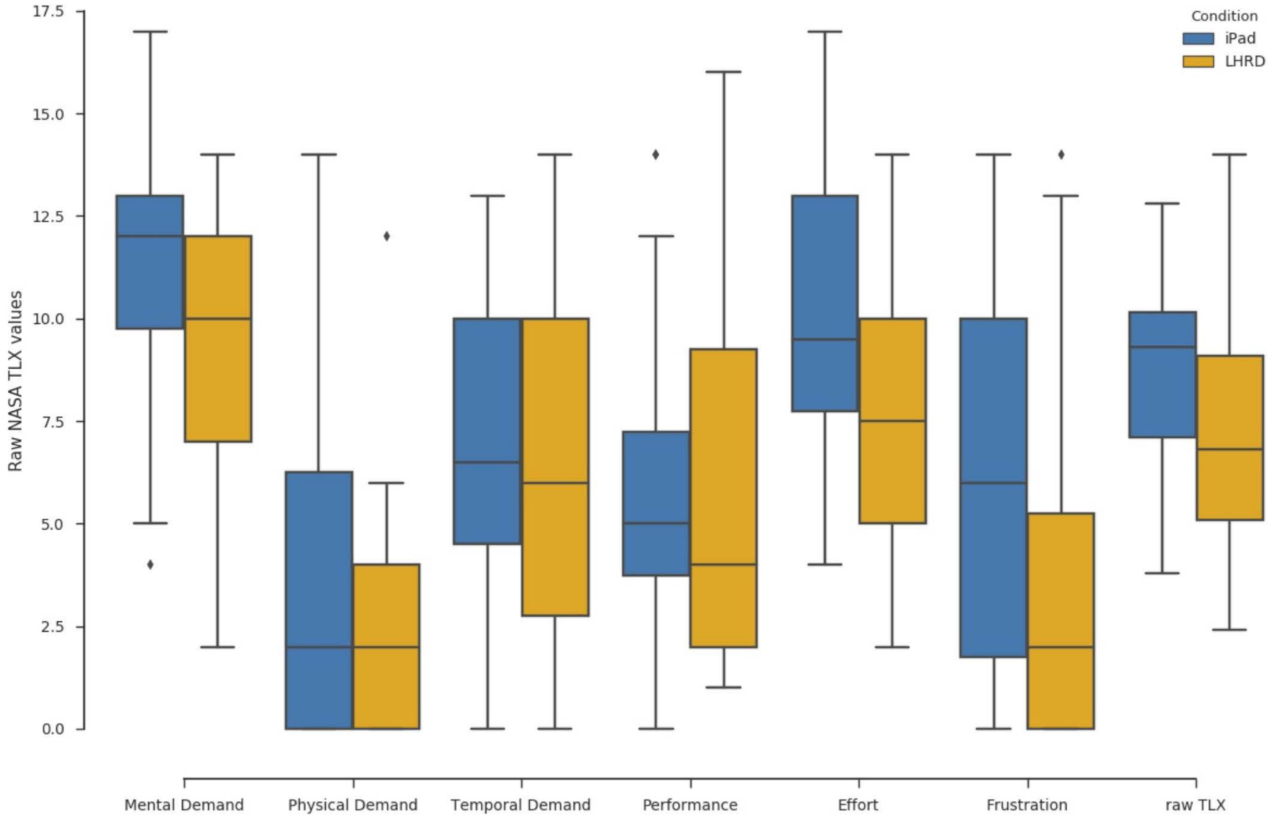

Figure 4. Raw NASA TLX scores comparing the iPad and LHRD condition.

was it very unique, 'This could be like every third person here in Germany'. Others felt comfortable with the lead researcher. P17 said she felt comfortable sharing her health data before interacting with the dataset, "it was more the approach that affected me". Those who had taken part in previous academic research studies felt comfortable and familiar with the process itself. For the participants that had a serious pre-existing health condition the importance of health studies was obvious and they were incentivized to take part in order to help others. P34 who spoke about her personal health issues said:

I share my story just to make this person know what probably can come in the future.

Although the rate of increased comfort was at 25%, the participants experienced a dramatic increase in trust after interacting with the data. Overall 70% of the participants said that they trusted the research project as a result of the study. P9, who indicated that she did not feel more comfortable sharing her own data, expressed a positive perception of both the value of data gathering and the importance of health-related research, she said:

When I did the experiment I found that it is really important and that for example this can help in research. I felt that yes, this is really helpful.

In order to gauge to what extent, the participant's trust and comfort could influence their decision to take part in future health data collection we asked them if they would share data with us in the future. All of the participants were positive about sharing health data in the future but most expressed a series of concerns and conditions. Anonymity, data security, minimum effort, helping others, trust, and gaining knowledge were the most cited conditions to future participation.

P22: I think it would honestly depend on whoever took the data and if I trust the organisation.

P8: I don't mind doing this but it depends on how much effort I have to do in order to report it. I would happily do this if when I report I am reporting this to have a medication from my doctor and it goes automatically to the dataset I would be completely fine with it.

P21: Let's say I give a blood test of mine to a hospital and they do some research on it and they can help people with it, yes, probably I would do it once a month. But in most of the other cases no.

Despite the fact that we did not observe a difference in the results between the LHRD and tablet scenarios the participant's increase in comfort with sharing health data and trust towards the data collection process, indicate that transparency is essential to participation.

DISCUSSION

The LHRD scenario was less cognitively demanding with the collaborative data exploration tasks we designed. Several aspects of the LHRD interface and collaboration may have contributed to this result, such as increased clarity of data points and distribution on the large displays, standard mouse interactions, and less shared control of the interface. The tablet proved harder to interact with due to the small size of the graphs populated with many small data points, and the awkward touch-based selection.

Despite the lack of statistically significant difference between the collaboration survey results for both scenarios, the video recordings and interviews illustrated the development of very different interaction techniques. The tablets allowed for shared control, whereas the LHRDs saw one participant taking over the interaction with the interface. Although, having a shared control of the interface may allow for more personal inquiry and data exploration, this did not have a positive impact on cognitive load.

Finally, the majority of participants felt an increased sense of trust towards the research team after interacting with the data and a quarter of them felt more comfortable sharing their personal health data after interacting with the visualisations. We do not attribute the change in willingness solely to interacting with the data. Instead the participants identified a number of influencing factors, which contributed to their personal increased comfortability, including an increased understanding of data impact, value and use, and comfort with the lead researcher.

Insight gathering: collaboration styles and screen size

Within this study the LHRD condition with traditional mouse and keyboard interface, which did not foster shared control, was less cognitively demanding than the iPad condition, which used a touchscreen and allowed for more turn-taking. Conversely Shaer et al. [35] discovered that touch-screen table-tops foster more collaborative learning than traditional desktop setups. This means that user interfaces where both partners are actively engaged are also less mentally demanding. This is further supported by research, which leverages more active collaboration as a way to enable sensemaking [31].

We believe that there are two factors that contribute to the difference in results between our study and previous work. The first, is that within the work of Shaer et al. [35], the researchers used different software between their conditions, which may mean that within their study the observed increased collaboration was fostered by the software in addition to the interaction. Our study, unlike previous research, emphasized a comparison at the display level with less confounding factors that could influence the results.

The second factor, is that within our study the iPad condition did not allow for easy data exploration due to the complexity of the data and the limited screen space, which could in turn undermine any positive effects of turn-taking and shared control. One of the issues with the iPad scenario was that the touchscreen interactions were difficult due to the complexity and small size of the visualisations. This may indicate that the screen size also plays a decisive role in people's perceptions of cognitive load. This reasoning is supported in the research of Reda et al. [33], where it was discovered that larger displays with significantly increase the number of discoveries reported by users. More research is needed to evaluate if screen size has a greater effect on cognitive load than collaboration styles, within collaborative data exploration tasks.

LHRD potential

The exploration of big data could soon become a common part of our lives as democratic governments push towards a more open and transparent legislation process. Using large displays for advanced analysis is practised in professional environments such as control rooms [22, 44]. LHRDs can help facilitate sensemaking by non-experts and contribute to the outreach of open government initiatives by engaging citizens in data exploration and collection in public spaces. As this study illustrated, current state of the art visualisation software, like Spotfire, can facilitate learning and understanding from non-experts. Large displays may help the public to more easily explore and understand visualised data collaboratively, because they allow for the clear visualisation of complicated and crowded data-sets.

A growing body of HCI research has already focused on the use of large public displays for the increasing of community engagement [17, 20]. As the technology becomes ever more accessible in our urban environment, there is potential to bridge the gap between government and citizens, through data collection and exploration initiatives in public spaces. In the process, we see potential for the development of appropriate communication channels that can close the loop between governmental research and the public.

The future of LHRDs may include large touchscreen displays, voice controls or assistants, and more traditional interfaces like we used in this study. Future research should investigate the most appropriate interaction methods to support data exploration. Touchscreen interfaces on LHRDs would not be subject to the same issues as the iPad condition, where the data visualisation may have been too small for an appropriate collaborative data exploration experience. Instead touchscreen LHRDs may allow for more shared control and more personal as well as collective data exploration. This indicates that despite our results showing that the LHRD scenario with traditional desktop interface was less cognitively demanding than the touchscreen tablet interface, there is room to explore the most appropriate ways of interacting with the LHRDs.

Supporting Data Exploration

The interviews and video recordings illustrated that the participants often confused cause and effect within the data. They could identify when there was a general trend and reported feeling that they understood the data but they made assumptions about the relationships between the variables present. As they progressed some were able to identify a range of additional background information that they need in order to fully understand the data. Despite this issue the majority of participants displayed an exceptional level of understanding of relationships between datasets and could identify specific information that could help them make conclusions about the relationships and validity of the information they were presented with. This indicates that although we have a fairly comprehensive understanding of how to visualise data for insight gathering, there is still sufficient room to explore ways to help the public understand its real meaning.

One way to help understanding may be to include Data Storytelling elements (DS elements), which can help add clarity and leverage sensemaking [12]. For collaborative data exploration the ability to leave notes and trace data exploration history could also enable sensemaking [30]. Furthermore, as indicated by the participants additional background information could help further their understanding. This information could be visualised when a user hovers over a graph axis or a data point. Alternatively, future visualisation tools for non-experts may benefit from incorporating search capabilities within their tools.

Enabling trust to foster data sharing

We found a significant positive influence of data exploration on the participant's comfort with sharing their personal health information. The participants were able to identify trends and felt that they understood its significance and potential use. In understanding how, health data could be used for research and insight gathering they reported that they would trust us with obtaining their personal health data in the future. This indicates that there is potential to increase public participation in democratic societies by distributing well visualised relevant data to the public.

There was no definitive connection between data exploration and the willingness of participants to share their personal health information. Instead we saw a range of perceived reasons, which influenced the increase in willingness to share personal health data. Trust and control were identified as the most important factors that could lead to continuous sustained data sharing. The participants self-reported a series of factors that would influence their willingness to take part in a data gathering project. They wanted the ability to stay anonymous, understand how their data would be used, how it would be made secure, they also felt they needed to know it would help others, and that it would not require much effort on their part. These findings are echoed within the research of Anderson and Edwards [1] who examined the need for a chain of trust between stakeholders in order to encourage clinical data sharing. Furthermore, current European legislation also leverages individual's control of the way their health data is shared [39].

Research in the area of open government investigates the influencing factors for a positive relationship between data distribution and citizen participation [18, 37, 38]. For example, gamification and social elements may provide some incentive for the adoption of participatory applications [37]. What past research has not investigated is the ways in which this transparency and control can actually be enabled. Based on the significantly increased trust and willingness to share health data of the participants, we feel that giving research results back to the original participants in the form of interactive visualized and clear content would provide appropriate downstream accountability. Findings ways to increase accountability to the public can in turn foster trust, which may result in future data sharing. A further challenge within the field of health data sharing would be to find ways to give participants detailed and specific control over how their data is being used. That may include the organisations that would have access to it, the ways in which it is anonymised, and what types of research it could be used for.

CONCLUSION

Through a between-subjects lab study we investigated the effects an LHRD and tablet had on collaborative data exploration and perceptions of comfort with sharing personal health data. The paper makes 3 contributions to the study of LHRDs, collaborative data exploration and data collection. While large displays were perceived as less mentally demanding than tablets within the same set of data exploration tasks on Spotfire, tablets lead to more turn taking and shared control of the interface. The participants in both scenarios felt more comfortable with sharing their own data with us after the study. Most participants felt that interacting with the visualisations, as an example of how their own data could be used, fostered a sense of trust.

ACKNOWLEDGEMENTS

Katerina Gorkovenko completed this work as a visiting researcher at the University of Stuttgart. Her stay and this work were supported by the European Commission's Horizon 2020 Programme under the funding scheme 'FETPROACT-1-2014: Global Systems Science (GSS)', grant agreement # 641191 CIMPLEX.

REFERENCES

1. Nick Anderson and Kelly Edwards. 2010. Building a chain of trust. In *Proceedings of the 2010 Workshop on Governance of Technology, Information and Policies - GTIP '10*. ACM Press. DOI: http://dx.doi.org/10.1145/1920320.1920323

2. Christopher Andrews, Alex Endert, Beth Yost, and Chris North. 2011. Information visualization on large, high-resolution displays: Issues, challenges, and opportunities. *Information Visualization* 10, 4 (aug 2011), 341–355. DOI: http://dx.doi.org/10.1177/1473871611415997

3. Christopher Andrews and Chris North. 2013. The Impact of Physical Navigation on Spatial Organization for Sensemaking. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (dec 2013), 2207–2216. DOI: http://dx.doi.org/10.1109/tvcg.2013.205

4. Pierre Andrews and Flávio Soares Corrêa da Silva. 2013. Using parliamentary open data to improve participation. In *Proceedings of the 7th International Conference on Theory and Practice of Electronic Governance - ICEGOV '13*. ACM Press. DOI: http://dx.doi.org/10.1145/2591888.2591933

5. Johanna E. Awotwi and Charles Amega-Selorm. 2016. Transparency and Accountability After 2015. In *Proceedings of the 9th International Conference on Theory and Practice of Electronic Governance - ICEGOV*

'*15-16*. ACM Press. DOI: http://dx.doi.org/10.1145/2910019.2910081

6. Robert Ball, Chris North, and Doug A. Bowman. 2007. Move to improve: Promoting Physical Navigation to Increase User Performance with Large Displays. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*. ACM Press. DOI: http://dx.doi.org/10.1145/1240624.1240656

7. Aaron Bangor, Philip T. Kortum, and James T. Miller. 2008. An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction* 24, 6 (jul 2008), 574–594. DOI: http://dx.doi.org/10.1080/10447310802205776

8. Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (jan 2006), 77–101. DOI: http://dx.doi.org/10.1191/1478088706qp063oa

9. John Brooke. 2013. SUS: A Retrospective. *J. Usability Studies* 8, 2 (Feb. 2013), 29–40. http://dl.acm.org/citation.cfm?id=2817912.2817913

10. Elizabeth F. Churchill. 2012. From data divination to data-aware design. *interactions* 19, 5 (sep 2012), 10. DOI: http://dx.doi.org/10.1145/2334184.2334188

11. Kellyton dos Santos Brito, Misael dos Santos Neto, Marcos Antônio da Silva Costa, Vinicius Cardoso Garcia, and Silvio Romero de Lemos Meira. 2014. Using parliamentary Brazilian open data to improve transparency and public participation in Brazil. In *Proceedings of the 15th Annual International Conference on Digital Government Research - dg.o '14*. ACM Press. DOI:http://dx.doi.org/10.1145/2612733.2612769

12. Vanessa Echeverria, Roberto Martinez-Maldonado, and Simon Buckingham Shum. 2017. Towards data storytelling to support teaching and learning. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction - OZCHI '17*. ACM Press. DOI:http://dx.doi.org/10.1145/3152771.3156134

13. Michele Gentili, Sara Hajian, and Carlos Castillo. 2017. A Case Study of Anonymization of Medical Surveys. In *Proceedings of the 2017 International Conference on Digital Health - DH '17*. ACM Press. DOI: http://dx.doi.org/10.1145/3079452.3079490

14. Nitesh Goyal and Susan R. Fussell. 2016. Effects of Sensemaking Translucence on Distributed Collaborative Analysis. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16*. ACM Press. DOI: http://dx.doi.org/10.1145/2818048.2820071

15. Nitesh Goyal, Gilly Leshed, Dan Cosley, and Susan R. Fussell. 2014. Effects of implicit sharing in collaborative analysis. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. ACM Press. DOI: http://dx.doi.org/10.1145/2556288.2557229

16. Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX) 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (oct 2006), 904–908. DOI: http://dx.doi.org/10.1177/154193120605000909

17. Luke Hespanhol, Martin Tomitsch, Ian McArthur, Joel Fredericks, Ronald Schroeter, and Marcus Foth. 2015. Vote as you go: blending interfaces for community engagement into the urban space. In *Proceedings of the 7th International Conference on Communities and Technologies - C&T '15*. ACM Press. DOI: http://dx.doi.org/10.1145/2768545.2768553

18. Soonhee Kim and Jooho Lee. 2012. E-Participation, Transparency, and Trust in Local Government. *Public Administration Review* 72, 6 (aug 2012), 819–828. DOI: http://dx.doi.org/10.1111/j.1540-6210.2012.02593.x

19. Søren Knudsen, Mikkel Rønne Jakobsen, and Kasper Hornbæk. 2012. An exploratory study of how abundant display space may support data analysis. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction Making Sense Through Design - NordiCHI '12*. ACM Press. DOI: http://dx.doi.org/10.1145/2399016.2399102

20. Lisa Koeman, Vaiva Kalnikaité, and Yvonne Rogers. 2015. "Everyone Is Talking about It!": A Distributed Approach to Urban Voting Technology and Visualisations.. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. ACM Press. DOI: http://dx.doi.org/10.1145/2702123.2702263

21. Gwanhoo Lee and Young Hoon Kwak. 2012. An Open Government Maturity Model for social media-based public engagement. *Government Information Quarterly* 29, 4 (oct 2012), 492–503. DOI: http://dx.doi.org/10.1016/j.giq.2012.06.001

22. Lars Lischke, Sven Mayer, Andreas Preikschat, Markus Schweizer, Ba Vu, Paweł W. Woźniak, and Niels Henze. 2018. Understanding Large Display Environments: Contextual Inquiry in a Control Room. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*. ACM, New York, NY, USA, Article LBW134, 6 pages. DOI: http://dx.doi.org/10.1145/3170427.3188621

23. Cristiano Maciel, Claudia Cappelli, Cleyton Slaviero, and Ana Cristina Bicharra Garcia. 2016. Technologies for popular participation. In *Proceedings of the 17th International Digital Government Research Conference on Digital Government Research - dg.o '16*. ACM Press. DOI:http://dx.doi.org/10.1145/2912160.2912191

24. Gabriela Marcu, Anind Dey, and Sara Kiesler. 2014. Designing for Collaborative Reflection. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*. ICST. DOI:http://dx.doi.org/10.4108/icst.pervasivehealth.2014.254987

25. Sven Mayer, Lars Lischke, Jens Emil Grønbæk, Zhanna Sarsenbayeva, Jonas Vogelsang, Paweł W. Woźniak, Niels Henze, and Giulio Jacucci. 2018a. Pac-Many: Movement Behavior when Playing Collaborative and Competitive Games on Large Displays. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 539, 10 pages. DOI: http://dx.doi.org/10.1145/3173574.3174113

26. Sven Mayer, Lars Lischke, Paweł W. Woźniak, and Niels Henze. 2018b. Evaluating the Disruptiveness of Mobile Interactions: A Mixed-Method Approach. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 406, 14 pages. DOI: http://dx.doi.org/10.1145/3173574.3173980

27. Albert J. Meijer, Deirdre Curtin, and Maarten Hillebrandt. 2012. Open government: connecting vision and voice. *International Review of Administrative Sciences* 78, 1 (mar 2012), 10–29. DOI: http://dx.doi.org/10.1177/0020852311429533

28. Agnes Owuato Odongo and Gideon Cheruiyot Rono. 2016. Open Government Data as a Right for Effective Citizen Participation. In *Proceedings of the 9th International Conference on Theory and Practice of Electronic Governance - ICEGOV '15-16*. ACM Press. DOI:http://dx.doi.org/10.1145/2910019.2910035

29. Aisling Ann O'Kane and Helena Mentis. 2012. Sharing medical data vs. health knowledge in chronic illness care. In *Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts - CHI EA '12*. ACM Press. DOI:http://dx.doi.org/10.1145/2212776.2223812

30. Benoît Otjacques, Mickal Stefas, Mal Cornil, and Fernand Feltz. 2012. Open data visualization keeping traces of the exploration process. In *Proceedings of the First International Workshop on Open Data - WOD '12*. ACM Press. DOI: http://dx.doi.org/10.1145/2422604.2422612

31. Sharoda A. Paul and Madhu C. Reddy. 2010. Understanding together. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work - CSCW '10*. ACM Press. DOI: http://dx.doi.org/10.1145/1718918.1718976

32. Fateme Rajabiyazdi, Charles Perin, Lora Oehlberg, and Sheelagh Carpendale. 2017. The challenges of individuality to technology approaches to personally collected health data. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare - PervasiveHealth '17*. ACM Press. DOI:http://dx.doi.org/10.1145/3154862.3154923

33. Khairi Reda, Andrew E. Johnson, Michael E. Papka, and Jason Leigh. 2015. Effects of Display Size and Resolution on User Behavior and Insight Acquisition in Visual Exploration. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. ACM Press. DOI: http://dx.doi.org/10.1145/2702123.2702406

34. Bertrand Schneider, Megan Strait, Laurence Muller, Sarah Elfenbein, Orit Shaer, and Chia Shen. 2012. Phylo-Genie. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*. ACM Press. DOI: http://dx.doi.org/10.1145/2207676.2208720

35. Orit Shaer, Megan Strait, Consuelo Valdes, Taili Feng, Michael Lintz, and Heidi Wang. 2011. Enhancing genomic learning through tabletop interaction. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*. ACM Press. DOI: http://dx.doi.org/10.1145/1978942.1979361

36. Cathie Sudlow, John Gallacher, Naomi Allen, Valerie Beral, Paul Burton, John Danesh, Paul Downey, Paul Elliott, Jane Green, Martin Landray, Bette Liu, Paul Matthews, Giok Ong, Jill Pell, Alan Silman, Alan Young, Tim Sprosen, Tim Peakman, and Rory Collins. 2015. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Medicine* 12, 3 (mar 2015), e1001779. DOI: http://dx.doi.org/10.1371/journal.pmed.1001779

37. Sarah-Kristin Thiel. 2016. Reward-based vs. Social Gamification: Exploring Effectiveness of Gamefulness in Public Participation. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction - NordiCHI '16*. ACM Press. DOI: http://dx.doi.org/10.1145/2971485.2996739

38. Sarah-Kristin Thiel, Michaela Reisinger, and Kathrin Rderer. "I'm too old for this!"!": influence of age on perception of gamified public participation.. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia - MUM '16*. ACM Press. DOI:http://dx.doi.org/10.1145/3012709.3016073

39. Thomas Trojer, Basel Katt, Thomas Schabetsberger, Ruth Breu, and Richard Mair. 2012. Considering privacy and effectiveness of authorization policies for shared electronic health records. In *Proceedings of the 2nd ACM SIGHIT symposium on International health informatics - IHI '12*. ACM Press. DOI: http://dx.doi.org/10.1145/2110363.2110425

40. James R. Wallace, Stacey D. Scott, and Carolyn G. MacGregor. 2013. Collaborative sensemaking on a digital tabletop and personal tablets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. ACM Press. DOI: http://dx.doi.org/10.1145/2470654.2466458

41. Sebastian Weise, John Hardy, Pragya Agarwal, Paul Coulton, Adrian Friday, and Mike Chiasson. 2012. Democratizing ubiquitous computing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12*. ACM Press. DOI: http://dx.doi.org/10.1145/2370216.2370293

42. Paweł Woźniak, Nitesh Goyal, Przemysław Kucharski, Lars Lischke, Sven Mayer, and Morten Fjeld. 2016. RAMPARTS: Supporting Sensemaking with Spatially-Aware Mobile Interactions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. ACM Press. DOI: http://dx.doi.org/10.1145/2858036.2858491

43. Paweł Woźniak, Lars Lischke, Benjamin Schmidt, Shengdong Zhao, and Morten Fjeld. 2014. Thaddeus: A Dual Device Interaction Space for Exploring Information Visualisation. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI '14)*. ACM, New York, NY, USA, 41–50. DOI: http://dx.doi.org/10.1145/2639189.2639237

44. Paweł W. Woźniak, Lars Lischke, Sven Mayer, Andreas Preikschat, Markus Schweizer, Ba Vu, Carlo von Molo, and Niels Henze. 2017. Understanding Work in Public Transport Management Control Rooms. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17 Companion)*. ACM, New York, NY, USA, 339–342. DOI: http://dx.doi.org/10.1145/3022198.3026341

45. Beth Yost and Chris North. 2006. The Perceptual Scalability of Visualization. *IEEE Transactions on Visualization and Computer Graphics* 12, 5 (sep 2006), 837–844. DOI:http://dx.doi.org/10.1109/tvcg.2006.184