Mobile and Multimodal? A Comparative Evaluation of Interactive Workplaces for Visual Data Exploration
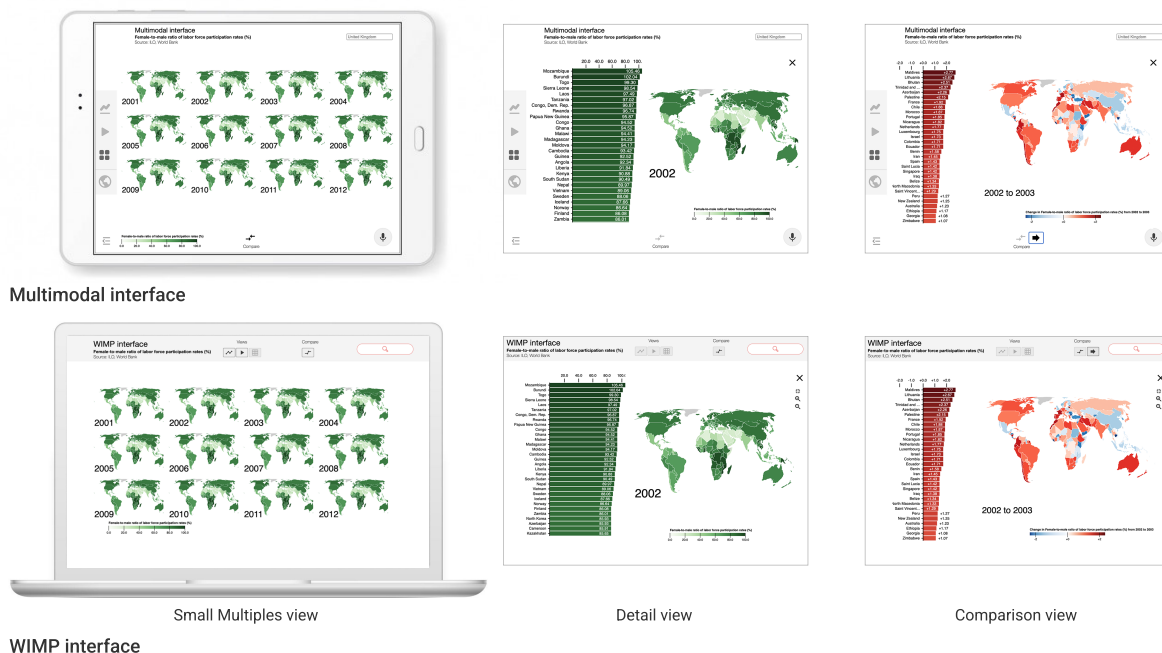
G. Molina León[1,2] , M. Lischka[1] , W. Luo[1], and A. Breiter[1,2]

[1]University of Bremen, Germany
[2]Institute for Information Management Bremen, Germany

Figure 1: *The small multiples view, detail view and comparison view of the multimodal and WIMP interfaces compared in our study.*

Abstract

Mobile devices are increasingly being used in the workplace. The combination of touch, pen, and speech interaction with mobile devices is considered particularly promising for a more natural experience. However, we do not yet know how everyday work with multimodal data visualizations on a mobile device differs from working in the standard WIMP workplace setup. To address this gap, we created a visualization system for social scientists, with a WIMP interface for desktop PCs, and a multimodal interface for tablets. The system provides visualizations to explore spatio-temporal data with consistent WIMP and multimodal interaction techniques. To investigate how the different combinations of devices and interaction modalities affect the performance and experience of domain experts in a work setting, we conducted an experiment with 16 social scientists where they carried out a series of tasks with both interfaces. Participants were significantly faster and slightly more accurate on the WIMP interface. They solved the tasks with different strategies according to the interaction modalities available. The pen was the most used and appreciated input modality. Most participants preferred the multimodal setup and could imagine using it at work. We present our findings, together with their implications for the interaction design of data visualizations.

CCS Concepts

• Human-centered computing → Empirical studies in visualization; Interaction design; • Applied computing → Law, social and behavioral sciences;

1. Introduction

Nowadays people are using mobile devices more often than desktop computers to access the web [Bro21]. Mobile devices support new interaction scenarios and more input modalities, which have the potential to change the way we interact with data. They enable direct manipulation through touch, are lightweight, and portable. However, they also come with challenges, such as a small screen size and less precision due to the "fat finger" problem. Therefore, designing visualization systems for mobile devices has become an increasingly important research goal [LIRC12]. Particularly tablets are a promising medium for visual data exploration. They have a comparable performance to compete with desktop computers and are increasingly used at work [Jes16]. Accordingly, standard visualization techniques, such as bar charts [DFS*13], scatterplots [SS14], and stacked graphs [BLC12], have been adapted to tablets with touch interaction. Furthermore, Drucker et al. [DFS*13] found that touch interaction can lead to better performance and user experience on tablets than interactions based on the standard WIMP (window, icon, menu, pointer) metaphor.

Mobile interfaces are called post-WIMP as they are designed differently according to the screen size and input modalities available. Possible modalities include pen [LSR*15], touch [BLC12], and speech [SLHR*20]. Hinckley et al. [HYP*10] found that combining pen and touch is both powerful and perceived as more natural. Recent work has combined more modalities to explore data in a "more fluid interaction experience" but most of these systems were evaluated with students or software company workers [LSR*15, SLHR*20, SSS20] who would not use them regularly and did not frequently use a tablet. At the workplace, tablets are often used during meetings and could thus become valuable for data exploration. Although mobile visual applications like Tableau Mobile [TS22a] are already on the market, these are mainly designed for touch and do not leverage multimodal interaction. We do not yet know enough about how tablet-based multimodal visualizations could be used in a work setting, and how they differ from their desktop WIMP counterparts regarding performance and user experience. How would domain experts analyze data on a tablet? How would their performance vary compared to working on the standard desktop setup? How would the experts make use of multimodal interaction? Would they approach their tasks differently?

To compare these workplace setups, we created a visualization system with two interfaces. Following previous work, we designed one interface for tablets supporting multimodal interaction and another for desktop computers supporting WIMP interaction. We consider each interface to be the representation of a setup as its design is based on the corresponding device and input modalities. We compare how domain experts perform exploratory data analysis in those setups by focusing on two research questions:

RQ1 How do the devices and the interaction modalities (mouse and keyboard vs. touch, pen, and speech) affect the performance of the domain experts, in terms of accuracy and response time?

RQ2 How do the devices and the interaction modalities affect the user experience?

We investigate these questions by conducting a within-subjects user study with experts from the social science domain, who explore data as part of their everyday work life. Data exploration plays an important role in the social sciences, in which the increasing availability of open data about government policies and development trends has motivated a growing interest in interactive data visualizations. These data are often spatio-temporal by nature: each data point represents the value of an indicator such as life expectancy or unemployment rate, associated to a country and to a time step. As detailed in section 3, we worked together with social scientists from the field of comparative politics to design a system for exploring development indicators, and created two different interfaces for tablets and desktop computers. Our goal was to investigate how the domain experts could work with multimodal visualizations on a tablet, and how their experience differs from conducting the same tasks on a more familiar WIMP environment. According to the interests of the experts, we focused on supporting the exploratory analysis of spatio-temporal data. We followed the task typology of Andrienko and Andrienko [AAG03a, AA06] for defining exploratory tasks that fit their workflow.

Research on multimodal visualizations has focused on qualitatively evaluating the designed systems so far [SSS20, SLHR*20]. We complement their work by looking also at quantitative metrics such as completion time and accuracy, in comparison with WIMP-based visualizations. We combine these metrics with the analysis of interaction logs and qualitative feedback to compare the performance and experience of the experts in both conditions. Accordingly, we conducted a semi-remote user study with 16 social scientists. Participants were significantly faster and made less errors with the WIMP interface, but were slightly more accurate solving synoptic tasks on the multimodal interface. We found that participants interacted significantly more on the tablet, and pen interaction was particularly appreciated and beneficial. The interaction analysis revealed that the smaller screen size of the tablet did not lead the experts to zoom more often but rather to approach the tasks differently. The experts had different strategies across devices and usually chose specific input modalities for individual actions. Ten participants preferred the multimodal interface, and 15 could imagine using it at work. According to our results, social scientists are interested in working with multimodal visualizations on a tablet, and they could perform as well or better than on a desktop computer after getting familiar with the input modalities.

With this paper, we contribute our quantitative and qualitative findings on how domain experts explore data differently with multimodal visualizations on a tablet, in contrast to using their desktop WIMP counterparts. We identify the different interactions patterns and strategies, and accordingly, provide recommendations for the interaction design of multimodal visualizations for tablets.

2. Related work

In this section, we present previous work on the topics of exploratory analysis of spatio-temporal data and post-wimp interaction for data visualization, and how our research connects to it.

2.1. Exploratory Analysis of Spatio-Temporal Data

Andrienko et al. [AAG03a] devised a typology of exploratory tasks for spatio-temporal data focused on time identification and comparison, at the elementary search level (individual time steps) or at

the general level (intervals). The latest classification considers tasks either *elementary* (about individual elements) or *synoptic* (set of elements) [AA06]. The authors describe exploratory data analysis as discovering properties of the dataset as a whole, mainly through synoptic tasks. They recommend using different visualization techniques depending on the task. We defined the visualizations of our system based on their recommendations. Furthermore, we designed a series of elementary tasks and synoptic tasks for our experiment based on the work tasks of the domain experts.

Boyandin et al. [BBL12] conducted a qualitative study on exploring temporal changes with flow maps through animation and small multiples. With animations, participants made more findings related to geographically local events and changes between subsequent years. With small multiples, they made more findings about long time periods. Thus, the authors suggest using both techniques to increase the number and diversity of the findings. Brehmer et al. [BLIC20] confirmed this recommendation on mobile phones. We designed the views of our system following their suggestions.

2.2. Post-WIMP interactions for data visualization

In the last decade, visualization researchers have been increasingly investigating the design of visualization "beyond the desktop" [LIRC12]. One of the most well-known studies on this topic is the work by Drucker et al. [DFS*13] on designing and comparing two interfaces of a bar chart application for tablets. The researchers designed one interface based on WIMP elements and another interface focused on using touch gestures. In their study, participants were significantly faster and preferred the gesture-based interface. Inspired by their work, we make a similar comparison but take both tablets and desktop PCs into account, adding pen and speech input.

Sadana and Stasko [SS16] designed a multiple coordinated views application combining WIMP elements and touch gestures. One of the main challenges was to define a set of consistent gestures across views. We designed our system with similar design principles (see section 4). We have multiple views, and we aimed at having consistent interaction techniques over specialized ones.

Oviatt et al. [ODK97] investigated first how multimodal interaction could support map-based tasks. They found that participants preferred pen interaction to draw symbols and would write with the pen before using speech, in combined interactions. Similarly, we combine pen and speech input, and supplement it with touch, given its relevance for tablets. More recently, Jo et al. [JLLS17] surveyed 13 studies on leveraging pen and touch interaction. They found that the five most common touch gestures in tablet-based studies were drag, tap, pinch, long press, double tap, and lasso selection. Accordingly, we limited the touch gestures of our multimodal interface to those. On the WIMP interface, we use standard interactions such as click, double click, and drag.

Much of the visualization research on leveraging multiple input modalities has focused on the combination of touch, pen, and speech. Srinivasan et al. [SS18] created a system for exploring networks with speech and touch. Saktheeswaran et al. [SSS20] compared said interface with its unimodal counterparts. The participants preferred multimodal input due to having more freedom of expression, and the option to combine modalities. Srinivasan et al. [SLS21] created a system for a large vertical display combining the three modalities and found that they complemented each other well in complex operations. Srinivasan et al. [SLHR*20] surveyed 18 visual systems to collect interaction techniques using touch, pen, or speech input. Based on that survey, they proposed a set of multimodal interactions for visualizations on tablets. We used their results as a base for choosing the multimodal interaction techniques.

3. Social science domain

As part of a multidisciplinary collaboration, we worked with social scientists to support them in the exploratory analysis of their spatio-temporal data. The scientists are members of a research project which explores the evolution and diffusion of social policy across the globe from 1850 until today. The policies are measured by indicators, such as health care expenditure or unemployment rate, and help the experts assess the development of nations [CLL*10]. Accordingly, the experts want to answer questions such as "how does health expenditure vary across world regions?", or "how did the unemployment rate in every country change over time?". The social scientists were already working with data visualizations provided by international organizations such as the World Bank [TWBG22], but wished for custom visual tools that would facilitate the recognition and the comparison of spatio-temporal patterns in the data.

We conducted a series of co-creation workshops and contextual interviews with the social scientists to explore visualization opportunities. We chose co-creation as a design methodology to empower the domain experts to actively shape the tools they wished for and to continuously validate the design [LHS*14, MLB20]. In the workshops, we found that the main interest of the experts was facilitating the first steps of their exploratory analysis, where they would look for countries and time spans of interest to focus on. Given that the indicators they work with often have varying temporal coverage and country samples, they were looking for options to explore the spatio-temporal coverage of the data, to recognize relevant patterns, and to compare data points over space and time. During the two-year collaboration, we co-designed web-based visual tools to explore their data (e.g. [MLLB20]). Furthermore, we observed that many of the experts owned and used a tablet at work to take notes and draw diagrams. Motivated by this observation and previous work on ubiquitous visual analytics [BMR*19, SLHR*20], we decided to provide a tablet-based interface for one of the tools. We conducted the user study more than a year after the last workshop. Some of the experts participated in both. The study was where they saw the evaluated system for the first time.

3.1. Data

The spatio-temporal data the experts work with is often relative data. Such a dataset typically consists of a set of triples of country, year, and value, e.g. (Chad, 2001, 11.3%). For the study, we chose datasets relevant to the social scientists published by the World Bank [TWBG22], one of their primary data sources. On each version, we visualized one of two datasets covering a 12-years time span (2001–2012): child mortality rate per 1000 live births [RRD13], and female-to-male ratio of labor force participation rates [OOTR18]. We counterbalanced the interface order and dataset assignment following a Latin square design.

Table 1: *Question examples classified according to the tyopology of Andrienko and Andrienko [AA06].*

Task question	Type	Sub-type
How high was the child mortality in Peru in 2001?	Elementary	Direct lookup
Which country had the highest female-to-male labor ratio in 2009?	Elementary	Inverse lookup
Is the child mortality in Myanmar lower, higher or equal to the one in Cambodia in 2009?	Elementary	Direct comparison
How was the female-to-male ratio in Western Europe in 2004?	Synoptic	Behavior characterization (space)
How did the child mortality develop in Northern Africa until 2007?	Synoptic	Behavior characterization (space over time)
In which African country did the female-to-male ratio increase most in the first five years?	Synoptic	Pattern search (time)
In which continent did the child mortality decrease most over the whole time period?	Synoptic	Pattern search (space over time)
How did the 2003 female-to-male ratio in South America compare to the one in Southern Africa?	Synoptic	Direct comparison (space)

3.2. Tasks

We defined the study tasks based on the exploration tasks that the social scientists described in the workshops, and on examples from the related study of Duncan et al. [DTPG20]. We needed clearly defined tasks to compare the performance and user experience of the social scientists across setups, and therefore, decided against an open exploration. According to the typology of Andrienko and Andrienko [AA06], we cover the following task types for the exploratory analysis of spatio-temporal data: (1) Direct lookup (elementary), (2) Inverse lookup (elementary), (3) Direct comparison (elementary), (4) Behavior characterization (synoptic), (5) Pattern search (synoptic), and (6) Direct behavior comparison (synoptic).

For the synoptic tasks, multiple variants were possible according to the reference sets: space, time, or space over time. Different combinations of the sets would lead participants to approach each task differently. In the study, we presented 13 tasks per dataset and interface: five elementary and eight synoptic tasks. We selected more synoptic tasks because those were predominant among the examples given by the experts. Given that interactivity played no significant role on the effectiveness of solving elementary tasks in previous work [DTPG20], the predominance of synoptic tasks suggested that this case study was suitable for comparing input modalities. Examples for each of the main question types are shown in Table 1. The full lists of questions are included in the supplementary material. Following the conceptual framework of Peuquet for spatio-temporal dynamics [Peu94], we formulated the exploratory tasks based on variations of its triad elements: *what*, *where* and *when*. In particular, we focus on varying the *where* and *when*.

4. Visualization and Interaction Design

In the following sections, we describe our design principles, the design of our multimodal system, and its WIMP counterpart. We aimed to create two interfaces with equivalent functionality and standard interaction techniques to make a fair comparison.

We present our design principles below, based on design reflections and findings on strategies to visualize spatio-temporal data [RFF*08, BBL12, BLIC20, PABP20] as well as on the interaction design of tablet-based visualizations [DFS*13, SS16, JLLS17].

DP1 *Leverage standard interaction techniques of multimodal systems.* Srinivasan et al. [SLHR*20] surveyed work on multimodal visualizations to determine what the standard interaction techniques are. To make a fair comparison, we mapped our inter-

action techniques according to those. We surveyed relevant examples to define the interaction techniques of choropleth maps because they were not included in the survey.

DP2 *Leverage standard interaction techniques of WIMP interfaces.* Given that line charts, bar charts and choropleth maps are commonly used visualizations, there are multiple well-known tools and examples that offer similar interaction techniques. We surveyed them to define the techniques of our system.

DP3 *Use standard touch gestures.* Familiar gestures are easier to remember and are usually preferred on touch-based visualizations [DFS*13, SS16, JLLS17]. We avoid complex gestures because discoverability is an issue on touch interfaces [BLC12, DFS*13] and such gestures can be hard to remember.

DP4 *Achieve interaction consistency.* Users expect that a gesture triggers similar results on different features of a system. Previous work on multiple coordinated views has emphasized the need for consistent gestures across views [SS16]. Accordingly, we put together a set of consistent interactions through all views.

DP5 *Introduce WIMP elements when necessary.* On the multimodal interface, we added redundant WIMP elements to ensure a good experience, following the findings of Drucker et al. [DFS*13]. For example, we enabled speech commands to switch views, but also included a side menu to do the same, to make sure that critical interactions could not be limited by speech recognition errors.

We created a first version of the system based on the requirements we elicited in the workshops. After finishing that version, we conducted an expert review with two HCI researchers who own and use a tablet regularly. After improving the system according to their feedback, we conducted an exploratory study with seven participants to further refine the system. In the following sections, we present the final design that resulted from those iterations.

4.1. Views and visualization techniques

We follow the suggestions of Andrienko et al. [AAG03b] on the visualization techniques to solve the elementary and synoptic tasks that we tackled. Overall, the system includes the following views:

1. *Lines view.* We included a multi-line chart to support comparison and behavior characterization tasks, focused on time [AA06]. This view gives an overview of the whole dataset and shows time trends (see Figure 3).
2. *Animation view.* We included animated choropleth maps for visualizing the temporal behavior of the spatial behavior [AA06]
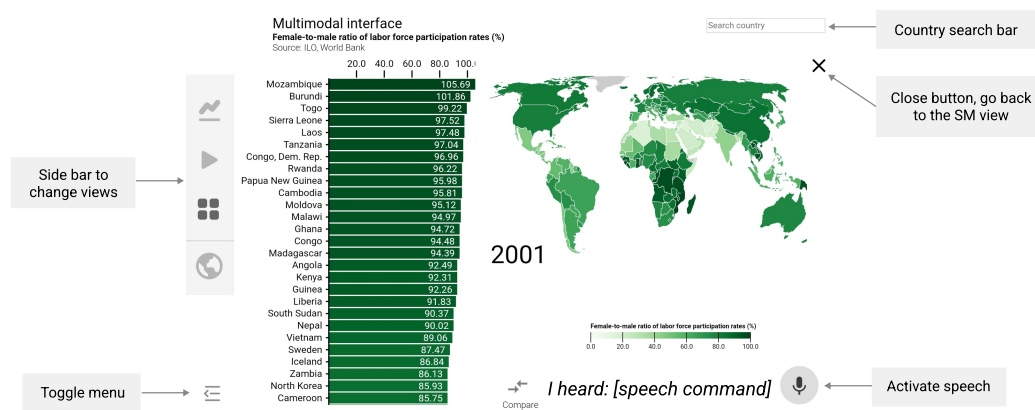
Figure 2: *Detail view of the multimodal interface. It presents the country values of a chosen year. On the left, the bar chart shows the 2001 values of each country, sorted in descending order by default. On the right, the choropleth map shows the data with its geographical location.*

(see Figure 3). We chose choropleth maps because many datasets used by the social scientists represent *relative* values.

3. *Small multiples view*. Small multiples provide an overview of the data and allow to visually compare it at different time points (see Figure 1). For propagation tasks, small multiples of choropleth maps perform better than other alternatives [PABP20].

4. *Detail year-based view*. We created the detail view for exploring the data distribution for a specific year. It includes a choropleth map to recognize the spatial distribution, and a sorted bar chart that helps identifying the countries that perform best or worst. The detail view of the multimodal interface is shown in Figure 2.

5. *Comparison view*. This view facilitates the comparison of two time steps. A choropleth map and a bar chart show the data values derived from calculating the difference between the two steps. The comparison view is depicted in Figure 1.

4.2. Interaction techniques

To decide on the interaction techniques of the maps, we surveyed interactive mapping tools such as Google Maps [Goo21] and Apple Maps [Inc21]. Additionally, we inspected tools like Datawrapper [Gmb21] that support the authoring of the visualization techniques we included. We defined our initial set of interaction techniques following the common features we recognized across tools, taking into account our design principles. We restricted our touch gestures to the standard set: tap, double tap, drag, swipe, and pinch.

The tablet interface was designed for touch, pen, and speech interaction, while the PC version for mouse and keyboard interaction. An overview of the interaction techniques is presented in Table 2. We show both interfaces with a few exemplary interaction techniques in the supplementary video. For navigating between views, we included buttons on both interfaces to make the interactions succinct and easy to discover (**DP5**). Both versions support brushing and linking. Overall, the WIMP actions are based on the tools we surveyed (**DP2**), and the multimodal actions follow the recommendations of Srinivasan et al. [SLHR*20] (**DP1**).

On the tablet, tapping is possible with either the pen or a finger.

In the study of Srinivasan et al. [SLHR*20], participants had a hard time differentiating pen and touch. Although the authors attributed this to a lack of experience with pen input, several of the domain experts we co-designed the system with were regular tablet users. They also coincided on enabling actions that could be possible with either the pen or a finger without limitations. We decided to allow for both to compare the setups according to how the experts would actually use them in their everyday life.

4.3. Implementation

We implemented both versions of the system as a web prototype with D3.js [BOH11] working with a Samsung Galaxy Tab S3 and an EIZO 23.8-inch desktop monitor. We used the standard HTML5 web speech recognition API [Mic19]. The prototype is available at: https://cocreation.uni-bremen.de/workplaces.

5. User study

Our goal was to compare the performance and the user experience of the domain experts with each interface. We wanted to gain a better understanding of how the interaction design influences the data exploration. Based on previous work and our research questions, we established the following hypotheses for the experiment:

H1 *The experts will need more time on the multimodal interface*. Natural User Interfaces (NUIs) are believed to be more engaging than WIMP interfaces, enabling a "more natural" interaction [SLHR*20] that encourages people to explore. Furthermore, multimodal interfaces are still rare and even if touch interaction with a smartphone is nowadays common, the combination of touch, pen, and speech is still new to many. This is why, we believe participants will need more time on the tablet.

H2 *The experts will make fewer errors on the WIMP interface*. We expect participants to be more accurate on the WIMP interface because it is the type of interface they already use at work. On the tablet, direct manipulation may lead to difficulties with precision, as it happens with the "fat-finger problem" [DFS*13]. Using a pen may compensate for this limitation of touch input be-
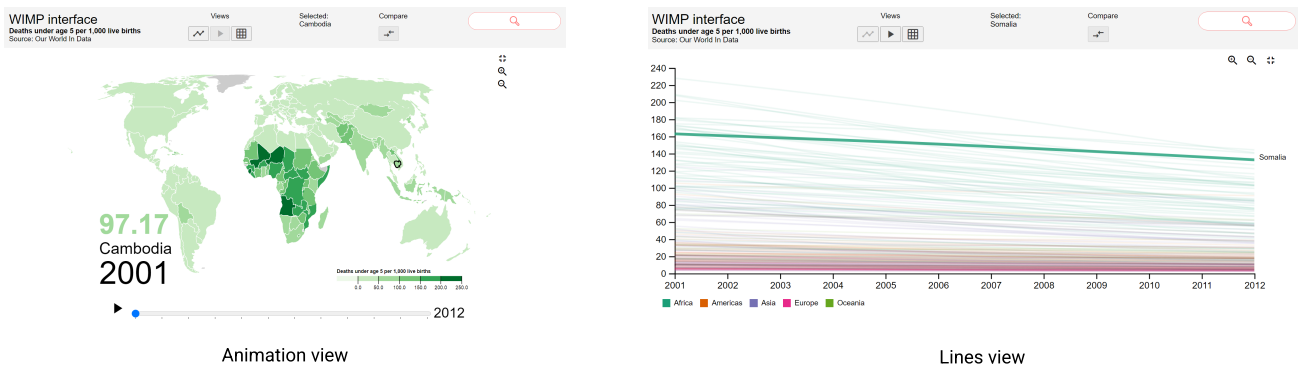
Figure 3: *Animation view (left) and Lines view (right) of the WIMP interface, showing the child mortality dataset.*

Table 2: *System actions with their corresponding interactions on each interface. Taps can be with either a finger or the pen. If applicable, the interaction includes the references to previous research or existing systems that it is based on.*

Action	WIMP interaction	Multimodal interaction
Go to a main view (Lines, Animation, SM)	Click a view button.	Speech command, e.g. "Go to line chart" [SLHR*20]. Tap on a view button.
Go to detail view	Click on a small multiple (SM) [vdEvW13].	Tap on a SM.
Go to comparison view	Click on Compare button. Then, click on two SMs. Then, click on View comparison button.	Tap on Compare button. Then, tap on two SMs. Then, tap on View comparison button.
Select a country	Click on map [TS22b]. Click on bar [TS22b]. Type country name in search bar [SBT*16].	Tap on map [SLHR*20]. Tap on bar [DFS*13, SLHR*20]. Speech command, e.g. "Select Italy" [SSS20, SS18]. Draw lasso [SSS20, SLHR*20].
Deselect a country	Click on selected country on map [TS22b]. Click on selected bar on bar chart [TS22b].	Tap on selected country on map. Tap on selected bar on bar chart [TS22b, JLLS17].
Select a year	Click on SM map.	Tap on SM map.
Deselect a year	Click on selected SM map.	Tap on selected SM map.
Deselect all	Click on empty space [TS22b].	Speech command "Deselect all" [SS18]. Tap on empty space [TS22b, SLS21].
Get country value	Hover on the map to show tooltip [Gmb21, TS22b]. Hover on a country line [Gmb21, TS22b].	See select (only possible for selected countries) [TS22b]. Drag pen over the x-axis of the line chart [SLHR*20].
Zoom	Click on any of the zoom buttons [Gmb21, TS22b]. Move mouse wheel [Goo21, TS22b].	Pinch gesture to zoom in or out [SLHR*20, Gmb21].
Pan	Drag the view with the mouse [Goo21, TS22b].	Drag with one finger [SLHR*20, Goo21].
Sort	Double click on x-axis of bar chart.	Swipe left or right on x-axis of bar chart [DFS*13].

cause pens have proven to convey precise spatial information on map-based tasks [Ovi97]. Still, we believe that familiarity with WIMP interfaces will lead to better spatial accuracy overall. As Duncan et al. [DTPG20] found that interactivity has a larger impact on accuracy for synoptic tasks on cartograms, we expect the accuracy difference to be large for the synoptic tasks.

H3 *Participants will prefer the multimodal interface.* In previous studies, participants preferred multimodal over unimodal interaction. However, these comparisons included only pairs of modalities: pen and speech vs. pen-only and speech-only [ODK97], and touch and speech vs. touch-only and speech-only [SSS20]. Recent work has shown that multimodal interaction can enhance the user experience and improve usability [SS18]. We continue the research by considering pen, touch, and speech together, and expect multimodal interaction to be preferred due to a more engaging user experience.

5.1. Experimental Design

We applied a within-subjects design, where each expert interacted with both interfaces to explore a dataset. We used different datasets for each interface and all combinations of interface and dataset were in counterbalanced order, as described in subsection 3.1.

To measure performance and user experience, we prepared 13 tasks for the experts to solve on each setup (see subsection 3.2). Each task consisted of a question about the given dataset with three possible answers, similar to previous studies [BLIC20, DTPG20]. Only one answer was correct. We formulated the questions in a way so that only the years, and the countries or regions, changed between datasets. We aimed to mention all continents equally often, to avoid focusing on a region that participants may be familiar with.

We created an online survey in which we measured the response time as the time between arriving at the task page and clicking on

the "Next" button. We measured accuracy as the error rate. Each wrong answer was counted as one error. The study had four parts: (1) Consent and demographics, (2) Introduction and tasks with the first interface and dataset, (3) Introduction and tasks with the other interface and the other dataset, and (4) Comparison survey. The tasks, surveys, and data are included in the supplementary material.

First, we explained the motivation of the study to the participants and asked for their written consent to record the session. They proceeded to answer a series of demographic questions and to perform a color vision test to make sure that they could correctly distinguish colors, similar to Duncan et al. [DTPG20]. Then, we shared a web link for accessing the system and provided them with a slide deck that we had prepared for the corresponding interface (see supplementary material). The slides described the views and the interaction techniques. We asked participants to read them and to perform all interactions on the system while reading. We did this to make sure that every participant received the same information. Then, we gave them five minutes to interact freely. If there were any questions, we discussed them. When participants confirmed that they felt confident enough, we proceeded with the tasks. We asked them to think aloud while solving the tasks to better understand their interactions. They did it on both interfaces so the performances were comparable. Afterwards, we asked participants to rate their satisfaction with the System Usability Scale (SUS) questionnaire [B*96]. Subsequently, we asked what they liked, what they disliked and what they missed about the corresponding setup. At the end of the study, we asked participants what interface they preferred, and to mention one to three reasons for their choice.

Due to the COVID-19 pandemic, we conducted the study semi-remotely. We met participants shortly before and after the session to provide and collect the tablets. During the session, we communicated through a video conference tool. They answered all questions on their office computers, and we recorded their interactions via a tablet app and the video conference tool. The computers were provided by their employer and included a 23.8-inch monitor.

6. Results

We recruited 16 participants (eight female), their average age was 32 years. All participants were social scientists from diverse disciplines, mainly political science and sociology. They worked on topics such as international trade and welfare policies. Fourteen of them were researchers, and all had a Master's degree.

Eight participants reported that they work with data visualizations weekly. Ten participants interacted with touch devices daily, while only one interacted with pen-based and speech-based systems daily. For five participants, this was the first time using a pen as an input device. For seven, it was the first time using speech input. All participants spoke English fluently but none was a native speaker. Nine of them owned a tablet. Since this was more than half of the participants, we adjusted our experiment design to conduct between-subjects comparisons of tablet owners and non-owners.

6.1. Performance

Participants took longer to solve the tasks with the multimodal interface. Their accuracy was slightly better on the WIMP interface.

We detail these results in the following sections. In the statistical tests, we considered a difference significant when the p-value was below 0.05, and we report Pearson's correlation coefficient r as the effect size to provide a measure of the importance of the effect.

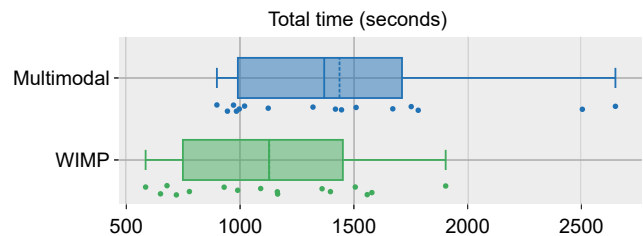### 6.1.1. Response time

Figure 4: *Total response time per interface.*

Figure 4 shows the total response time for each session per interface. The mean response time per task was 86.8 seconds with the WIMP interface and 110.54 seconds with the multimodal interface. After confirming that the difference between the response times was normally distributed, we ran a one-tailed t-test. The result showed that the participants were significantly faster on the WIMP interface with a medium to large effect ($t(15) = 1.83, p = 0.043, r = 0.43$). Therefore, **H1** is supported. Looking at individual tasks, the time difference was larger on T4 and T11, which were synoptic tasks. Based on these observations, we compared the time of elementary tasks with the time of synoptic tasks. There was no significant difference between the response times across task types.

In addition, we compared tablet owners and non-owners to find out whether owning a tablet had an impact on their response time with the multimodal interface. Tablet owners took longer ($M = 24.88$ minutes, $SD = 9.05$) than non-owners ($M = 22.75$ minutes, $SD = 9.39$), but not significantly ($W = 25, p = 0.54, r = -0.21$).
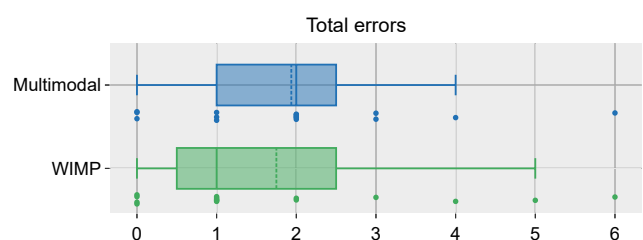
6.1.2. Accuracy

Figure 5: *Distribution of the errors per interface.*

Participants made less errors with the WIMP interface than with the multimodal interface. They solved 86.54% of the tasks correctly with the former and 85.1% with the latter. The error distribution is shown in Figure 5. Given that the distribution was not normal, we ran a Wilcoxon signed-rank test to investigate the differences. Participants were not significantly more accurate on the WIMP interface ($W = 33.5, p = 0.39$), so **H2** is not supported by our experiment. We tested whether the interface order had an effect and found

no significant difference ($t = -0.31, p = 0.76$). There was also no significant effect of the dataset ($t = 1.32, p = 0.21$).

We also tested whether owning a tablet had an impact on the accuracy of the participants with the multimodal interface. Tablet owners actually made slightly more errors ($M = 2.22$) than non-owners ($M = 1.57$) but the difference was not significant ($W = 35.5, p = 0.70, r = 0.13$). Additionally, we tested whether the accuracy of the participants was significantly better on the WIMP interface for each task type, to investigate whether our results fit the results of Duncan et al. [DTPG20]. On average, participants made more errors on the synoptic tasks ($M = 1.28$) than on the elementary tasks ($M = 0.56$), which corresponds to their difficulty level. For the elementary tasks, participants made less errors on the WIMP interface, but not significantly. For the synoptic tasks, participants made less errors on the multimodal interface, but the difference was not significant ($U = 128.0, p = 0.49, r = -0.05$).

6.2. Interactions

We recorded the interactions of the participants with screen video and audio recording. Then, we logged and coded the interactions per session. For each interaction, we documented the participant ID, the device, the dataset, the task, the view, the input modality, the action (see Table 2), the command or gesture, and the outcome. We successfully logged the interactions of 14 participants in 28 videos (two videos per person). For two participants, there were technical issues that did not allow to record their interactions properly.

We logged 4087 individual interactions. Each interaction corresponds to an attempt to perform an action, e.g. pinch to zoom. We classified its outcome as either *successful* (i.e. the system reacted to the action as it should have), *erroneous* (i.e. the system did not react as it should have), *invalid* (i.e. the interaction is not valid in the current view or state) or *unsupported* (i.e. the interaction is not one of those the system recognizes). Overall, the success rate of the interactions was 0.94. The erroneous rate was 0.03, and the rates for unsupported and invalid were 0.02 and 0.01. Most errors happened during selection on the tablet due to speech recognition errors. A common issue was that participants were unsure about the English pronunciation of country names, and this led to errors.

We logged 2197 interactions on the multimodal interface and 1890 on the WIMP interface. On the WIMP interface, participants performed 1697 (89.8%) mouse interactions and 193 (10.2%) keyboard interactions. On the multimodal interface, the pen prevailed: 1544 (70.3%) interactions were pen-based, 554 (25.2%) were touch-based and 99 (4.5%) were speech-based. The dominance of the pen over touch is surprising because 10 participants used pen interaction rarely or never, but most used touch daily. According to the qualitative feedback, participants liked the pen because of its high precision, and the ability to select by drawing.

We show the total of interactions per modality and participant on the tablet in Figure 6. Every participant tried each input modality at least once. However, most had one dominant modality. Of these 14 participants, 11 mostly used the pen to interact, two mostly touch, and one almost equally used both pen and touch. Among the 11 participants who mostly used the pen, four had never interacted with pen-based systems, two interacted with it less than once per
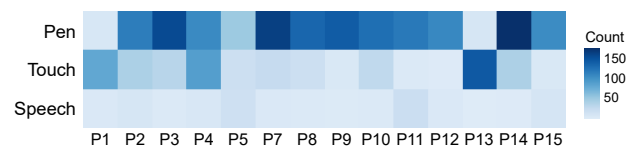

Figure 6: *Total of interactions per input modality and per participant on the multimodal interface.*

month, and five at least monthly. This suggests that the tendency to use the pen was independent of the frequency of pen use in their everyday life. Of the 11 participants, seven owned a tablet.

6.2.1. Actions and views per interface

Previous work suggests that multimodal interaction is more engaging and consequently leads to more interactions [SS18]. Thus, we compared the interactions per participant across devices. On average, participants interacted significantly more with the multimodal interface ($M = 156.93, SD = 40.27$) than with the WIMP interface ($M = 135.00, SD = 36.19$), $t(13) = 1.85, p = 0.046, r = 0.45$.

Table 3: *Interactions per interface and input modality of the most common actions.*

| Interface | Multimodal | | | WIMP | |
Action	Pen	Touch	Speech	Mouse	Keyboard
Go to Animation view	47	1516	0	56	0
Go to Comparison view	247	41	0	156	0
Go to Detail view	101	19	0	109	0
Go to Lines view	39	21	0	72	0
Go to SM view	213	41	0	246	0
Select country	353	87	95	234	193
Deselect country	93	9	0	69	0
Deselect all	89	21	3	56	0
Get country value	9	8	0	286	0
Zoom	0	114	0	125	0
Pan	202	93	0	37	0

We show what type of actions participants performed with each input modality in Table 3. Looking at the actions across devices, participants selected more often on the tablet, and performed the *Get a country value* action more often on the desktop PC. This makes sense given that country values were visible while hovering, but getting values on the tablet required selecting first. *Pan* was the second most often action on the tablet, mostly on the bar charts.

Regarding the views, participants interacted most often on the *Detail* view on the tablet (29.30%), and on the *Small multiples* view on the desktop (26.95%). The larger size of the visualizations of the detail view may have been more important on the smaller screen. Participants interacted more with the *Animation* view on the tablet (21.98%) than on the desktop (13.98%), but used the *Lines* view more on the desktop (19.60% vs. 6.59%). The screen size and the possibility of hovering are the most likely reasons for this.

6.2.2. Interactions per task

More than half of the participants interacted with the pen on every task, either alone or in combination with speech or touch input. In

contrast, only one person used touch on every task. We compared the interactions based on the accuracy of the participants on the tablet. Participants who solved a task correctly interacted more with the interface. Looking at the combination of input modalities used per task, we found that for the tasks solved correctly, the most common interaction was the pen only (35.48%) followed by the combination of pen and touch (30.97%). On tasks answered wrongly, the most common combination was pen and touch (33.33%). Interacting with the pen alone is thus associated to better accuracy.

By analyzing how participants successfully solved each task, we found the following patterns across devices:

1. *On the tablet, participants used most views with larger maps. Pen selection on the map, panning on the bars.* For tasks about specific countries, participants often went to the Detail or Animation view for selecting them with the pen, based on where they thought the country was. They would sometimes combine this with zooming on the map to make a more precise selection, and with panning on the bar chart to get an idea of the relation of that country to others. This was the most common pattern on the tablet. For tasks where two time steps were involved, a variation of this pattern would take place on the Comparison view.
2. *For time intervals, most used the line chart on the PC, and the comparison view on the tablet.* When comparing the views used for solving temporal development tasks across, it was noticeable that people preferred to use the *Comparison* view on the tablet and the *Lines* view on the PC. This reveals that participants had different ways to solve the same task across devices.
3. *Hovering was key to solve most tasks on the PC.* On the WIMP interface, participants selected less often and compensated with hovering. Some participants solved most tasks with the *Lines* view and mainly interacted by using the search bar to select a country, and then hovering to inspect its values.

6.3. User experience

We asked participants to rate their experience based on the standard SUS questionnaire. Afterward, we asked them a few questions about their impression of the system. As mentioned above for **H3**, we expected multimodal interaction to provide a better experience. Participants considered that the system features were better integrated on the multimodal interface, which is noteworthy considering the multiple input modalities. Furthermore, participants scored the multimodal interface as quicker to learn. We show the scores for the questions with largest difference across interfaces in Figure 7.

After interacting with each version of the system, we asked participants what they liked and disliked about it. On the multimodal interface, four participants particularly liked the ability to select with the pen due to its high precision, and the option to draw for selecting. This corresponds to the pen being the most used input modality on the tablet (see subsection 6.2). Multiple participants mentioned that the speech input was important for searching for countries whose geographical location they did not know. When asked what they disliked, participants mentioned the speech recognition errors, and selecting countries with their fingers by accident while zooming. On the WIMP interface, participants especially appreciated the possibility to hover to get the tooltip, and the ease of
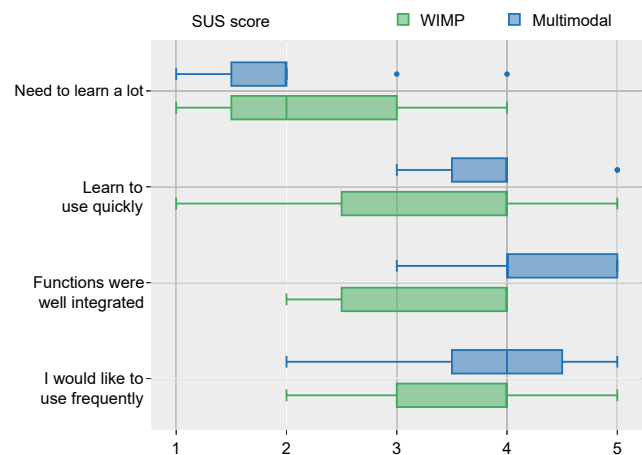
Figure 7: *Score distribution of the four SUS statements for which the answers differed most between the two interfaces. The score range was from 1 (strongly disagree) to 5 (strongly agree).*

searching for countries with the keyboard. The most common issue was that hovering on the line chart led to highlighting the closest line and this line would sometimes overlap with the line of interest.

In addition, we asked participants whether they could imagine using the system at work. Fifteen answered positively. They highlighted the advantage of performing "quick checks" and "fast comparisons" of their data to answer questions such as "how has the child mortality changed due to the civil war in Syria?". Fifteen participants would use the multimodal interface at work and 14 would use the WIMP interface.

6.4. Preferences

Ten of 16 participants preferred the multimodal interface. Therefore, our results support **H3**. Participants focused on different input modalities to justify their choice. Three participants argued that zooming with touch gestures felt easier than with the buttons of the WIMP interface. Three participants especially liked the lasso selection with the pen, and for P4, it felt faster than the mouse. According to P3, the speech input made the multimodal interface "really superior." For three participants, the multimodal interface felt "more intuitive", and one felt more confident with it. For P5, touch interaction was "much more fun than just keyboard and mouse."

Six participants preferred the WIMP interface. Their main reasons were the familiarity of working with mouse and keyboard and the larger screen of the PC. P2 did not own a smartphone, and therefore, felt that he was not skilled enough with touchscreens to perform well on the tablet. P7 preferred the multimodal interface but pointed out that she would rather use the WIMP version at work because tablets with good performance are scarce at the workplace.

7. Discussion

Our results indicate that the different combinations of devices and interaction modalities affect the performance (**RQ1**) and user expe-

rience (**RQ2**) of the domain experts during exploratory data analysis. The social scientists were significantly faster solving tasks in the familiar WIMP context, but they were similarly accurate on the tablet, and were even better solving synoptic tasks on it. At first sight, it may seem that the smaller screen size of the tablet caused participants to be slower because they may have spent more time zooming and panning. However, the logged interactions presented in Table 3 indicate that participants did not zoom more on the tablet than on the PC (114 times on the tablet vs. 125 on the PC). Moreover, the interaction patterns show that participants had different strategies to solve the tasks across conditions. For example, they used the *Lines* view more often on the PC because they could easily hover on the lines. Therefore, we conclude that the interaction modalities were the most decisive factor on the interaction choices of the experts. This is confirmed by their qualitative feedback where they justified their preferences based on the modalities available.

Although most participants had interacted rarely or never with a pen, they used it for most interactions and were successful with it. Pen and speech interaction were especially helpful to select with precision and more comfortable when the country location was unknown. That suggests that each input modality fits best to specific actions, and its benefits depend on the task at hand. In visualizations of large datasets, the precision of the pen may be most valuable to interact with each data point. Furthermore, most participants already used a tablet in the office, but rather for simple tasks such as taking notes. Our results suggest that if multimodal tools are given, domain experts would consider including them into their workflow.

Our results are not as positive for the tablet interface as the ones of Drucker et al. [DFS*13], but given the success of the participants with pen interaction and their comparable accuracy overall, we believe that data exploration on the tablet may become more beneficial and preferred at work, after getting familiar with it. We consider the lack of a significant difference on accuracy as a positive result because pen and speech interaction are still not as common as touch, which makes the accuracy and user experience results promising. Moreover, Drucker et al. compared both interfaces on a tablet while our WIMP condition included a PC because we wanted our results to reflect the real-world experiences of the experts. For the same reason, we used a PC and a tablet with different screen sizes. While the interaction analysis and the qualitative feedback suggest that the modalities were the most decisive factor on the participant choices, the interaction patterns also reveal that participants tended to use the views with larger visualizations on the smaller display. Thus, our findings do not compare the interaction modalities only, but rather the combination of devices and modalities. They describe how the experience is shaped by both factors.

7.1. Recommendations for interaction design

Our findings suggest three main recommendations for the design of multimodal visualization systems for tablets:

1. Pen interaction was dominant regardless of previous experiences. Thus, the pen should be able to perform most interactions, and all critical interactions should be possible with it.
2. Participants described and appreciated each input modality based on the actions they preferred to perform with it. The pen

was notably helpful to select small countries, confirming the findings of Oviatt [Ovi97]. Thus, we recommend pen interaction for selecting in map-based visualizations. Given the better performance of touch with bar charts [DFS*13], we conclude that performance depends on the modality that suits better the corresponding mix of visualization and interaction techniques.
3. According to the qualitative feedback, speech input was very appealing despite its problems. This is consistent with the findings of Saktheeswaran et al. [SSS20] on multimodal interaction being less error-prone than speech-only. Leveraging speech interaction may lead to a more engaging experience, but other modalities should support the same actions to guarantee usability.

7.2. Limitations

We defined the experiment tasks according to the work of the experts we collaborated with. A study with open exploration would help verify whether interacting significantly more on the tablet is associated with being more interested in exploring multimodally in general. Furthermore, we defined our interaction techniques based on previous work, but adding more complex techniques that combine the three input modalities sequentially or simultaneously would help to learn more about how people interact multimodally.

Due to the COVID-19 pandemic, we conducted the study semi-remotely. We provided the tablets, but the experts used their office computers. Although the monitors were of the same model, this means that the experiment was not fully controlled, and further investigation is needed to confirm our findings. We also asked participants to think aloud. We acknowledge that this may have influenced the results, yet without any clear observable bias in one direction. Furthermore, having a larger sample, and using products provided by a third party, would help to test the reliability of our results.

8. Conclusions

We investigated how devices and input modalities affect the performance and user experience of domain experts while solving exploratory tasks on spatio-temporal data. Participants used pen interaction for high precision tasks without having much experience, used touch for zooming, and speech for selecting countries. Our work suggests that combining touch, pen, and speech, is a promising option for visual data exploration, and that different modalities fit better to specific tasks and lead to different interaction patterns.

Although we designed the system for domain experts, we think that exploring such data is also relevant for the general public. Leveraging multimodal interaction may make the exploration process more engaging, but the data literacy of the audience should be taken into account. Furthermore, exploring data with different input modalities may be an opportunity to make visualizations more accessible. The relevance for accessibility should be studied further.

References

[AA06] ANDRIENKO N., ANDRIENKO G.: *Exploratory analysis of spatial and temporal data: a systematic approach.* Springer Science & Business Media, 2006. 2, 3, 4

[AAG03a] ANDRIENKO N., ANDRIENKO G., GATALSKY P.: Exploratory spatio-temporal visualization: an analytical review. *Journal of Visual Languages and Computing 14*, 6 (2003), 503 – 541. doi:10.1016/S1045-926X(03)00046-6. 2

[AAG03b] ANDRIENKO N., ANDRIENKO G., GATALSKY P.: Tools for visual comparison of spatial development scenarios. In *Proceedings on Seventh International Conference on Information Visualization, 2003. IV 2003.* (2003), pp. 237–244. doi:10.1109/IV.2003.1217985. 4

[B*96] BROOKE J., ET AL.: Sus-a quick and dirty usability scale. *Usability evaluation in industry 189*, 194 (1996), 4–7. 7

[BBL12] BOYANDIN I., BERTINI E., LALANNE D.: A qualitative study on the exploration of temporal changes in flow maps with animation and small-multiples. *Computer Graphics Forum 31*, 3pt2 (2012), 1005–1014. doi:10.1111/j.1467-8659.2012.03093.x. 3, 4

[BLC12] BAUR D., LEE B., CARPENDALE S.: Touchwave: Kinetic multi-touch manipulation for hierarchical stacked graphs. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces* (New York, NY, USA, 2012), ITS '12, Association for Computing Machinery, p. 255–264. doi:10.1145/2396636.2396675. 2, 4

[BLIC20] BREHMER M., LEE B., ISENBERG P., CHOE E. K.: A comparative evaluation of animation and small multiples for trend visualization on mobile phones. *IEEE Transactions on Visualization and Computer Graphics 26*, 1 (2020), 364–374. doi:10.1109/TVCG.2019.2934397. 3, 4, 6

[BMR*19] BADAM S. K., MATHISEN A., RÄDLE R., KLOKMOSE C. N., ELMQVIST N.: Vistrates: A component model for ubiquitous analytics. *IEEE Transactions on Visualization and Computer Graphics 25*, 1 (2019), 586–596. doi:10.1109/TVCG.2018.2865144. 3

[BOH11] BOSTOCK M., OGIEVETSKY V., HEER J.: D³ data-driven documents. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (2011), 2301–2309. doi:10.1109/TVCG.2011.185. 5

[Bro21] BROADBANDSEARCH.NET: Mobile vs. desktop internet usage (latest 2021 data), 2021. URL: https://www.broadbandsearch.net/blog/mobile-desktop-internet-usage-statistics. 2

[CLL*10] CASTLES F. G., LEIBFRIED S., LEWIS J., OBINGER H., PIERSON C. (Eds.): *The Oxford Handbook of the Welfare State.* Oxford University Press, Oxford, 2010. doi:10.1093/oxfordhb/9780199579396.001.0001. 3

[DFS*13] DRUCKER S. M., FISHER D., SADANA R., HERRON J., SCHRAEFEL M.: Touchviz: A case study comparing two interfaces for data analytics on tablets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2013), CHI '13, Association for Computing Machinery, p. 2301–2310. doi:10.1145/2470654.2481318. 2, 3, 4, 5, 6, 10

[DTPG20] DUNCAN I. K., TINGSHENG S., PERRAULT S. T., GASTNER M. T.: Task-based effectiveness of interactive contiguous area cartograms. *IEEE Transactions on Visualization and Computer Graphics* (2020), 1–1. doi:10.1109/TVCG.2020.3041745. 4, 6, 7, 8

[Gmb21] GMBH D.: Datawrapper, 2021. URL: https://www.datawrapper.de/. 5, 6

[Goo21] GOOGLE: Google maps, 2021. URL: https://www.google.com/maps. 5, 6

[HYP*10] HINCKLEY K., YATANI K., PAHUD M., CODDINGTON N., RODENHOUSE J., WILSON A., BENKO H., BUXTON B.: Pen + touch = new tools. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 2010), UIST '10, Association for Computing Machinery, p. 27–36. doi:10.1145/1866029.1866036. 2

[Inc21] INC. A.: Apple maps, 2021. URL: https://apps.apple.com/us/app/maps/id915056765. 5

[Jes16] JESDANUN A.: The consumer tablet is dying; long live the business tablet, 2016. URL: https://www.ctvnews.ca/sci-tech/the-consumer-tablet-is-dying-long-live-the-business-tablet-1.2829557. 2

[JLLS17] JO J., L'YI S., LEE B., SEO J.: Touchpivot: Blending wimp & post-wimp interfaces for data exploration on tablet devices. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2017), CHI '17, Association for Computing Machinery, p. 2660–2671. doi:10.1145/3025453.3025752. 3, 4, 6

[LHS*14] LANDSTORFER J., HERRMANN I., STANGE J.-E., DÖRK M., WETTACH R.: Weaving a carpet from log entries: A network security visualization built with co-creation. In *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)* (2014), pp. 73–82. doi:10.1109/VAST.2014.7042483. 3

[LIRC12] LEE B., ISENBERG P., RICHE N. H., CARPENDALE S.: Beyond mouse and keyboard: Expanding design considerations for information visualization interactions. *IEEE Transactions on Visualization and Computer Graphics 18*, 12 (2012), 2689–2698. doi:10.1109/TVCG.2012.204. 2, 3

[LSR*15] LEE B., SMITH G., RICHE N. H., KARLSON A., CARPENDALE S.: Sketchinsight: Natural data exploration on interactive whiteboards leveraging pen and touch interaction. In *2015 IEEE Pacific Visualization Symposium (PacificVis)* (2015), pp. 199–206. doi:10.1109/PACIFICVIS.2015.7156378. 2

[Mic19] MOZILLA, INDIVIDUAL CONTRIBUTORS: Web speech api, 2019. URL: https://developer.mozilla.org/en-US/docs/Web/API/Web_Speech_API. 5

[MLB20] MOLINA LEÓN G., BREITER A.: Co-creating visualizations: A first evaluation with social science researchers. *Computer Graphics Forum 39*, 3 (2020), 291–302. doi:10.1111/cgf.13981. 3

[MLLB20] MOLINA LEÓN G., LISCHKA M., BREITER A.: Mapping the global south: Equal-area projections for choropleth maps. In *2020 IEEE Visualization Conference (VIS)* (2020), pp. 91–95. doi:10.1109/VIS47514.2020.00025. 3

[ODK97] OVIATT S., DEANGELI A., KUHN K.: Integration and synchronization of input modes during multimodal human-computer interaction. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 1997), CHI '97, Association for Computing Machinery, p. 415–422. doi:10.1145/258549.258821. 3, 6

[OOTR18] ORTIZ-OSPINA E., TZVETKOVA S., ROSER M.: Women's employment. *Our World in Data* (2018). https://ourworldindata.org/female-labor-supply. 3

[Ovi97] OVIATT S.: Mulitmodal interactive maps: Designing for human performance. *Human–Computer Interaction 12*, 1-2 (1997), 93–129. URL: https://www.tandfonline.com/doi/abs/10.1080/07370024.1997.9667241. 6, 10

[PABP20] PEÑA-ARAYA V., BEZERIANOS A., PIETRIGA E.: A comparison of geographical propagation visualizations. In *CHI'20-38th SIGCHI conference on Human Factors in computing systems* (2020), Association for Computing Machinery, pp. 223–1. doi:10.1145/3313831.3376350. 4, 5

[Peu94] PEUQUET D. J.: It's about time: A conceptual framework for the representation of temporal dynamics in geographic information systems. *Annals of the Association of American Geographers 84*, 3 (1994), 441–461. doi:10.1111/j.1467-8306.1994.tb01869.x. 4

[RFF*08] ROBERTSON G., FERNANDEZ R., FISHER D., LEE B., STASKO J.: Effectiveness of animation in trend visualization. *IEEE Transactions on Visualization and Computer Graphics 14*, 6 (Nov 2008), 1325–1332. doi:10.1109/TVCG.2008.125. 4

[RRD13] ROSER M., RITCHIE H., DADONAITE B.: Child and in-
fant mortality. *Our World in Data* (2013). URL: https://
ourworldindata.org/child-mortality. 3

[SBT*16] SETLUR V., BATTERSBY S. E., TORY M., GOSSWEILER R.,
CHANG A. X.: Eviza: A natural language interface for visual anal-
ysis. In *Proceedings of the 29th Annual Symposium on User Inter-
face Software and Technology* (New York, NY, USA, 2016), UIST '16,
Association for Computing Machinery, p. 365–377. doi:10.1145/
2984511.2984588. 6

[SLHR*20] SRINIVASAN A., LEE B., HENRY RICHE N., DRUCKER
S. M., HINCKLEY K.: *InChorus: Designing Consistent Multimodal
Interactions for Data Visualization on Tablet Devices*. Association for
Computing Machinery, New York, NY, USA, 2020, p. 1–13. doi:
10.1145/3313831.3376782. 2, 3, 4, 5, 6

[SLS21] SRINIVASAN A., LEE B., STASKO J.: Interweaving multimodal
interaction with flexible unit visualizations for data exploration. *IEEE
Transactions on Visualization and Computer Graphics 27*, 8 (2021),
3519–3533. doi:10.1109/TVCG.2020.2978050. 3, 6

[SS14] SADANA R., STASKO J.: Designing and implementing an inter-
active scatterplot visualization for a tablet computer. In *Proceedings of
the 2014 International Working Conference on Advanced Visual Inter-
faces* (New York, NY, USA, 2014), AVI '14, Association for Computing
Machinery, p. 265–272. doi:10.1145/2598153.2598163. 2

[SS16] SADANA R., STASKO J.: Designing multiple coordinated visual-
izations for tablets. *Computer Graphics Forum 35*, 3 (2016), 261–270.
doi:10.1111/cgf.12902. 3, 4

[SS18] SRINIVASAN A., STASKO J.: Orko: Facilitating multimodal in-
teraction for visual exploration and analysis of networks. *IEEE Trans-
actions on Visualization and Computer Graphics 24*, 1 (Jan 2018), 511–
521. doi:10.1109/TVCG.2017.2745219. 3, 6, 8

[SSS20] SAKTHEESWARAN A., SRINIVASAN A., STASKO J.: Touch?
speech? or touch and speech? investigating multimodal interaction for
visual network exploration and analysis. *IEEE Transactions on Vi-
sualization and Computer Graphics 26*, 6 (2020), 2168–2179. doi:
10.1109/TVCG.2020.2970512. 2, 3, 6, 10

[TS22a] TABLEAU SOFTWARE I.: Tableau mobile, 2022. URL: https:
//www.tableau.com/products/mobile. 2

[TS22b] TABLEAU SOFTWARE I.: Tableau software, 2022. URL:
https://www.tableau.com. 6

[TWBG22] THE WORLD BANK GROUP: The World Bank. https:
//data.worldbank.org/indicator, 2022. 3

[vdEvW13] VAN DEN ELZEN S., VAN WIJK J. J.: Small multiples, large
singles: A new approach for visual data exploration. *Computer Graphics
Forum 32*, 3pt2 (2013), 191–200. doi:10.1111/cgf.12106. 6