Increasing the bandwidth of Interactive Visualizations, using complex display environments and targeted designs.

Anastasia Bezerianos

Habilitation à Diriger des Recherches

presenté par

Anastasia Bezerianos

Increasing the bandwidth of Interactive Visualizations, using complex display environments and targeted designs

Discipline: Informatique
Spécialité: Interaction Homme-Machine

February 16, 2020

Jean-Daniel Fekete	Inria	Examinateur
Eric Lecolinet	Télécom Paris	Examinateur
Laurence Nigay	Université Grenoble Alpes	Rapporteur
Margit Pohl	TU Wien	Examinateur
Nicolas Sabouret	Université Paris-Saclay	Examinateur
Claudio Silva	New York University	Rapporteur
Jack van Wijk	Eindhoven U. of Technology	Rapporteur

Habilitation à Diriger des Recherches préparée au sein de l'équipe HCC du Laboratoire de Recherche en Informatique de l'Unviersité Paris Saclay, et de l'équipe ILDA commune au centre de recherche Inria Saclay, à l'Université Paris-Saclay, et au CNRS

Acknowledgements

I would first like to thank the members of my Habilitation committee. The three *rapporteurs* for their insightful and thoughtful reports: Laurence Nigay who provided me with excellent advise and has supported HCI research in France for years; Claudio Silva who helped me see my career through the eyes of experienced visualization researchers; and Jack van Wijk whose reflective and deep report moved me more than I can say. And my *examinateurs*: Jean-Daniel Fekete who introduced me (directly and indirectly) to visualization research and has mentored me through the entire process; Eric Lecolinet who helped me think more about the connection between visualization and HCI; Margit Pohl who has supported my students and myself over the years; and Nicolas Sabouret for being an exemplary president. Thank you for you all for your precious time and feedback, the questions you asked me gave me food for thought and inspiration for the future. I would also like to thank David Auber and John Stasko, for their letters supporting my application to the Habilitation à Diriger des Recherches.

Next I want to deeply thank my colleagues over the years, that shaped how I see and conduct research. A big thank you to the members of the ILDA team Caroline Appert, Olivier Chapuis and Emmanuel Pietriga for being amazing colleagues and friends, and making coming to work a joy. Thanks to my long time colleagues and very dear friends Nadia Boukhelifa, Fanny Chevalier, Pierre Dragicevic, Jean-Daniel Fekete, Nathalie Henry-Riche, Petra Isenberg, Yvonne Jansen, Bongshin Lee, Evelyne Lutton, Theophanis Tsandilas and Jeanne Vezien, who have spent many hours with me, discussing questions on research and life. My deep appreciation to Lonni Besançon, Tanja Blascheck, Johannes Fuchs, Vanessa Peña-Araya and Nicole Sultanum for all the energy they bring in any project they work on.

Thanks to the heads of the HCC and InSitu/ExSitu teams Michel Beaudouin-Lafon and Wendy Mackay for making HCI research a core discipline in the University. To my amazing HCC colleagues Baptiste Caramiaux, Sarah Fdili Alaoui, Cédric Fleury and Tobias Isenberg. To the ILDA team students and engineers for making work everyday so much fun: Eugénie Brasier, Mehdi Chakhchoukh, Emmanuel Courtoux, Marie Destandau, Dylan Lebout, and Qi Ze. And to all the brilliant current and past members of the University and of the ILDA, ExSitu and Aviz teams that I've had the chance to work with either in research or in teaching in the past: Jessalyn Alvina, Ignacio Avellino Martinez, Benjamin Bach, Nicole Bidoit, David Bonnet, Marianela Ciolfi Felice, Yuheng Feng, Romain DiVozzo, Niklas Elmqvist, Stelios Frantzeskakis, Olivier Gladin, Pascal Goffin, Carla Griggio, Sarkis Halladjian, Steve Haroz, Jiayi Hong, Stephane Huot, Samuel Huron, Stacy Hsueh, Mohamad Alaul Islam, Thibaut Jacob, Germán Leiva, Ilaria Liccardi, Can Liu, María-Jesús Lobo, Mathieu Nancel, Charles Perin, Alexis Pister, Romain Primet, Jean-Philippe Rivière, Marzieh Rafiei, Reyhaneh Raissi, Hugo Romat, Mickaël Sereno, Natkamon Tovanich, Paola Valdivia, Frédéric Vernier, Romain Vuillemot, Xiyao Wang, Wesley Willet, Lijie Yao, and Yi Zhang.

Beyond the teams in my institution, I'd like to thank my many colleagues in France and around the world: Marie-Aude Aufaure, Gilles Bailly, Enrico Bertini, Sheelagh Carpendale, Anthony Collins, Christopher Collins, Maxime Cordeil, Tim Dwyer, Fabian Fischer, Steven Franconeri, Waqas Javed, Judy Kay, Daniel Keim, Gregor McEwan, Nathalie Mejean-Perrot, Cassio Melo, Matthias Miller, Teresa Onorati, Themistoklis Palpanas, Catherine Plaisant, Yann Riche, Markus Rittenbruch; and the administrators who have helped me so much: Murielle Benard, Stéphanie Druetta, Katia Evrat, Aurélie Lerasle, and Alexandra Merlin.

Most importantly, thank you to the PhD students I've had the absolute pleasure to supervise over the years: Evanthia Dimara, Micheline Elias, Anna Gogolou, Raphaël James, Arnaud Prouzeau and Tong Xue.

Before closing, I want to thank all my future colleagues and students that will inspire my next research, I can't wait to start!

Last but not least, I would like to thank George Katsirelos and my family for their patience and support.

(This section was updated after the HDR jury).

Contents

1 Introduction

The digital data our society generates increase dramatically every year. Open Data initiatives [ABK+07] from science [Wea16] and government [Uba13] promote the sharing of these data, that easily reach petabytes each year. For example the Large Hadron Collider alone has generated more than 200 petabytes[1] and the city of Paris has openly given access to more than 200 datasets[2]. In addition, we get increasing access to personally collected data, through API's and SDK's of tracking technologies in mobile phones and fitness trackers[3]. These give us today access to an unprecedented amount of data.

Nevertheless, understanding data remains challenging, as their volume far exceeds what one person can reasonably consume and understand. Statistic measures don't tell the whole story. For example very different datasets can have the same statistics as demonstrated by the Anscome's Quartet [Ans73] and by the more recent work from Matejka and Fitzmaurice [MF17]. Automated mining and machine learning processes can search for patterns, nevertheless it is to this day an open problem how to explain [AB18, RSG16] and communicate their suggestions[4]. Even when explanations are possible, automated approaches do not necessarily help the user actually understand their dataset (i.e., the deeper global relationships present in their data), may learn spurious correlations [RSG16], and still require training (supervised and unsupervised) so they are unreliable when patterns of interest are not frequent in the dataset. Thus sense-making approaches based on visually inspecting and interacting with data to make sense of it [PC05], remain an extremely valuable alternative.

Information visualization as a field attempts to combine human computer interaction, visual design and perception theory, in order to propose visual data representations that amplify cognition [CMS99] and aid data understanding [Car03]. Alone or combined with data processing methods, these representations support data exploration and visual analysis [TC05].

Coming from a Human-Computer Interaction (HCI) background, this definition of interactive visualization resonated with me, as it invokes the more general definition of a user-interface. An interface is traditionally seen as the means for humans and machines to communicate (or humans between them through machines) [DFAB03]. When I started research in visualization during my post-doc in 2008, I drew a parallel to this definition, seeing *information visualization as a means by which humans can communicate with their data and the machines that store and process them*. If we consider visualization as a communication channel, with humans on the one side and machines on the other, the higher the bandwidth the more effective the visualization is. We are faced with the limits of this channel given the amount of data at our disposal. Visually presenting large amounts of information remains a challenge, with previous work [FP02] showing that a million items shown on a traditional screen is close to the practical limit.

As a continuous goal of visualization research is to find new ways to amplify cognition [CMS99] and increase the amount and understanding of data communicated to us, this goal can be seen as an attempt to increase the communication bandwidth between machines and humans. There are limits at both sides of the communication channel (human and machine). I tend to characterize visualization research on a high level based on what direction of the channel it attempts to amplify.

[1] https://home.cern/news/news/computing/cern-data-centre-passes-200-petabyte-milestone
[2] https://opendata.paris.fr/
[3] Android developer.android.com/wear, Apple developer.apple.com/healthkit, Fitbit dev.fitbit.com
[4] Workshop on Visualization for AI Explainability http://visxai.io/

On the one side, the human viewer is constrained in the amount of visual information they can decode and process. Finding these viewer limits are a core topic of visual perception studies. These studies may compare different traditional visual variables for encoding data (like length, area, angle [Cle85, HB10] and color [JLMV06]), or newly introduced ones such as our own work on sketchiness **[BBIF12]**. They can study the perception of specific visualizations like scatterplots [RB10], barcharts [TSA14] or linecharts **[IBDF11]**. They can consider how to best communicate uncertainty [HQC+19]. Or examine the impact of different visualization platforms like tabletops [WSFB07], wall-displays **[BI12]**, mobiles [BLIC19] and smartwatches **[BBB+19]**, or even tangible visualizations [JDF13]. Increasing human computation power could also be achieved by allowing multiple viewers to collaborate and share their expertise, for example by having them collaborate in real time using large displays [JH14, LCBLL16, IFP+12], or in distributed desktop settings [VWvH+07, HVW07]). Other approaches attempt to automate, or semi-automate, user tasks to reduce their workload, for example by generating annotations [HDA13] or appropriate visualizations [DD19].

On the other side of the channel, as the amount of data increases, machines cannot always process and render it in real time. And the display technology may not have the pixel density to display it. New display technologies can show increased amount of data, for example large or multiple displays [BNB07, RJPL15, RWM+15] have a higher pixel count compared to traditional monitors. And mobile [KJM+07] or augmented reality [MSD+18] devices can show data in contexts that were not available before. Combining visualization with automated methods can also help guide the data mining algorithms towards interesting visual patterns **[BBTL15]**[BL09a], making better use of computational resources. And progressively calculating and revealing results can address computational scalability issues, and help viewers get quickly insights based on partial data and early approximations [FSME14, ZGC+17, MFDW17].

Most work though falls in the middle, considering both human and computational or rendering constraints. New visual representations can distill the important aspects of the data, thus rendering potentially less ink but more salient information for the viewer, amplifying both machine and human bandwidth. Examples include hybrid graph-matrix network visualizations [HFM07], compressed timeline representations [KL06], or facetted views of text corpora [CVW09]. And more recent work studies how to better guide viewers to known insights through visual storytelling [RHDC18]. Moreover, visual analysis tools [Kei02, KMS+08] are often designed with specific domain-users and their needs in mind, focusing on accelerating the discovery of salient information in medicine [TLS+14], urban data [FPV+13], etc. Most of these approaches provide meaningful overviews of data [Shn96]. For example through appropriate sampling or aggregations, like edge bundling in graphs [HvW09] and clustering of time series [vV99]. Or through combinations of multiple and coordinated data views that can reveal and focus on different aspects in the data [Wea04, BCC+05]. At the same time, appropriate general-purpose interactions (such as filtering [SGL08] and zooming [BSM04]) or ones targeting specific visualizations or data (like interactive horizon graphs [PVF13] or direct manipulation for time-varying data [KC14]), can help viewers focus on parts of their data (details on demand [Shn96]) without being overwhelmed. When the amount of data is hard to see in an overview, other approaches inverse the visual Information-seeking mantra of "Overview first, zoom and filter, then details-on-demand", by start from subparts of their data and slowly constructing a bigger view (e.g., [vdEvW14]).

Among these strategies, my own research attempts to increase this communication bandwidth in the following two ways:

▷ Move away from traditional desktops towards larger displays that can both render larger amounts of data and can accommodate multiple viewers (Chapter 2);

▷ Find appropriate visual representations and their limits, to show salient information that can be understood and acted upon (Chapter 3).

Figure 1.1 summarizes the focus on my work in a word cloud of my publication titles.

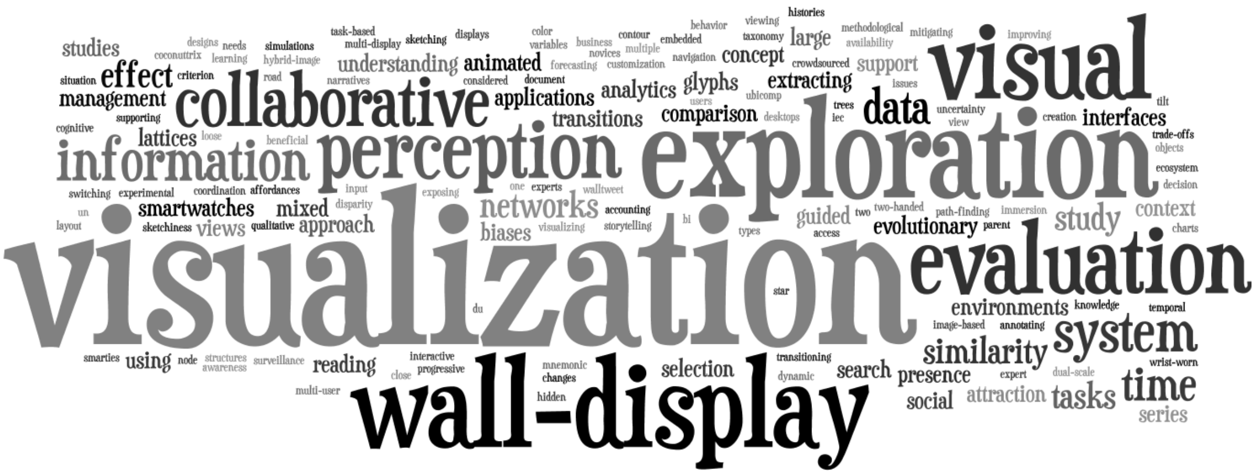


Figure 1.1: Word cloud of terms from my paper titles, indicating the focus of my research on collaboration and wall-display research, on perception and evaluation, and on visual exploration systems.

The high-level view of visualization research I have described above does not cover all possible work. Enhancing data exploration, and amplifying and understanding the limits of data visualizations, is by no means the only goal of visualization research. There is an increasing number of excellent work that considers visualization both from a creation and a large-scale communication perspective. Researchers have focused on making visualization authoring easier through the use of toolkits [BOH11, Fek04] and the definition of common grammars of graphics [SRHH16]. Other work targets larger audiences, possibly without programming expertise, introduce new authoring paradigms [NM17, SE19], combining visualization creation with natural language [SBT+16], and make more accessible the construction of creative and personalized visualizations [KIHR+19]. Recent work has also started to reflected on visualization and its impact more generally, looking at visualization education and literacy [ARC+17, BRBF14], examining the impact of visualization use at large scale [VWvH+07], attitudes of data-poor viewers towards visualization [PAEE19], and ethical implications of visualization research [Cor19].

Before summarizing the content of the next chapters that describe my own approaches to increase the human-machine bandwidth, I describe the methodologies followed in my research.

1.1 Research methodology and inspiration

I started out my research career in Human-Computer Interaction, with a PhD on interaction designs for wall-display environments, in an institution and lab[5] where HCI was a well established domain. This gave me the opportunity to access several classes that covered HCI methods for design, evaluation and analysis. Due to the interdisciplinary origins of HCI, the domain includes a plethora of methodologies [DFAB03, S+19] that are adapted from domains such as experimental psychology and human factors (more quantitative research, involving controlled experimental evaluations with users), sociology and anthropology (often qualitative research, such as observations, interviews, contextual inquiries, usually trying to understand user needs), design and software engineering (focusing on the design process, such co-design and iterative prototyping), etc. A common chapter in most HCI textbooks is one covering these methods and how to choose which one is best.

As such, HCI-trained researchers such as myself tend to be aware of, and open to, multiple methodologies for answering research questions and analyzing their findings. I have brought these methods to my

[5] DGP at University of Toronto `http://www.dgp.toronto.edu/`

research in visualization, choosing different methodologies depending on the goal of the work (mentioned in each section and summarized in the end of each chapter). My research is user-centric, involving target user groups in the design process, testing in the validation process, or both. But my methodology ranges from qualitative controlled experiments when trying to isolate different factors that affect performance, to participatory design and qualitative observations when introducing a new visual analysis system. When I started to actively conduct research in interactive information visualization in 2008 (work on improving the readability of social networks [**HBF08**]) the field had stated to reflect deeply about evaluation methods, with the BELIV workshop that had just appeared in the AVI and CHI conference (and more visualization-targeted manuscripts on evaluation methodologies started appearing [Car08, LBI+12, SMM12]). This active reflection on evaluation methods for visualization systems, and my belief that I could contribute to it, was one of the reasons my research focus shifted towards visualization research.

Visualization research has since come a long way when it comes to evaluation and validation of our work. Much of the visualization research mentioned in the previous section varies in research methodology, i.e., the steps we take and the evidence we provide. These are generally dictated by the goals of the individual piece of research. For example, in our own work considering traffic control centers [**PBC16a**] we started by interviewing analysts to understand their user needs before designing and testing visualization alternatives. Whereas when we studied the impact of glancing speed on visualization reading on smartwatches [**BBB+19**], we systematically compared in very controlled settings different speeds in terms of accuracy. Nevertheless, I would argue that our research *goals*, as stated in our respective research outcomes (publications, reports, systems) are not always the *inspiration* behind our work. More generally, it is not always clear where ideas for visualization research come from.

Manuscripts on visualization design studies (e.g., [SMM12]) promote a user-centered methodology to understand real-world data challenges and propose solutions. These include a thorough understanding domain data and users needs, before introducing more appropriate visualizations. But what is the inspiration behind these new designs, that often seem to magically appear, is not always clear. Was it knowledge of the literature on specific data types that sparked an idea for a new encoding, was it a combination of previous visualizations? Work that searches for theories and fundamental understanding (e.g., studies on the limits of the human visual system or comparisons of visual encodings) often starts with specific tasks or data representations. For these studies it is natural that the phenomena is studied in controlled conditions (e.g., lab experiments). But is it always the task/data that is the origin of this investigation? Or is it cross-pollination of findings and practices from other fields (e.g., vision science) that have inspired the original question? Is it inspired by observations of how people use existing visualization systems?

As our research reporting is often "story driven" (i.e., present a coherent story), the knowledge of what inspired our research questions and solutions may be lost (impossible for third parties to reconstruct or even forgotten by the authors). Nevertheless, such knowledge could act as a valuable tool for future visualization researchers attempting to form their own research agenda and process. In the following chapters I first describe the goal of my research and my research methodology, and then reflect on the inspiration behind my own work in the hopes to start a dialogue about better capturing and communicating visualization inspiration. In the final chapter I discuss how approaches from other disciplines could help us better document and share our research inspiration and evolution.

1.2 Manuscript Overview

This section briefly summarizes the remaining chapters of the manuscript, acknowledging my many collaborators that made this work possible, and providing representative papers of that work (attached at the end of the manuscript). Publications that I have co-authored are rendered in **bold** in the manuscript.

I split my work in two parts, each describing a different aspect of my work that tries to increase the communication bandwidth between human and the data we can process, the first is related to the use of large visualization platforms, and the second on visualization design and understanding.

Collaborative displays (chapter 2). Large, high-resolution displays such as interactive walls and tables, can extend the rendering limits of desktop displays due to their large real-estate and high resolution. They can also naturally allow multiple users to collaborate and explore data simultaneously. It is thus not surprising that data visualization and monitoring applications (e.g., [WSK+00, JLMV06, RWM+15]) are now migrated to digital walls, tabletops, or combinations of multiple such large displays. Moving interactive visualizations to large and multiple interaction surfaces creates many opportunities, but also raises challenging questions. For example, there are still open questions on what are appropriate visualizations for such environments, how to best help users explore and interact with these data, and more generally how do they support collaboration.

This chapter focuses on my work that explores interaction and visualization challenges in these novel display environments, examining first how to support interaction in a setting where mice and keyboards are not necessarily appropriate as viewers are often standing and mobile. The work on Smarties **[CBF14]** looks at prototyping interaction support and was done in collaboration with O. Chapuis and S. Franzeskakis. The work on SketchSliders **[TBJ15]** allows analysts to sketch their own interfaces. It was done in collaboration with T. Tsandilas and T. Jacob and received an Honorable mention award in CHI 2015.

Next, the chapter presents studies on how the large surface area of such displays can affect how we see and understand visualizations, and opportunities for new types of visualizations. The work on perception magnitude studies **[BI12]** was done with P. Isenberg, and the work on creating hybrid-image visualizations **[IDW+13]** that can encode different information depending on viewing distance was a collaboration with P. Isenberg, P. Dragicevic, W. Willet and J.-D. Fekete.

The last part of the chapter summarizes work on understanding and supporting collaborative visual analysis in these novel display environments **[PBC16a, PBC17a, PBC17b, PBC18, PBC16b, PBC15]** was conducted during the PhD of A. Prouzeau that I co-supervised with O. Chapuis. I also summarize briefly other relevant work on collaboration with colleagues from the INRA institution **[BBT+19]** (N. Boukhelifa, I.C. Trelea, N. Méjean and E. Lutton), and University of Sydney and NICTA **[CBM+09]** (A. Collins, G. McEwan, M. Rittenbruch, R. Wasinger, and J. Kay).

The chapter closes with reflections on the methodology and the inspiration behind this work.

• Representative paper (Interaction): T. Tsandilas, A. Bezerianos, T. Jacobs (2015). SketchSliders: Sketching Widgets for Visual Exploration on Wall Displays. Proceedings of ACM CHI 2015 - the ACM SICGHI Conference on Human Factors in Computing Systems, (10 pages), **Honorable mention** (top 5% of papers).

• Representative paper (Perception): A. Bezerianos and P. Isenberg (2012). Perception of Visual Variables on Tiled Wall- Sized Displays for Information Visualization Applications. In IEEE InfoVis 2012 - the IEEE Transactions on Visualization and Computer Graphics (Proceedings Scientific Visualization / Information Visualization 2012), 18(12): 2516- 2525, (10 pages). [25% acc. rate]

• Representative paper (Collaboration): A. Prouzeau, A. Bezerianos, O. Chapuis (2016). Evaluating multi-user selection for exploring graph topology on wall-displays. In TVCG - the IEEE Transactions on Visualization and Computer Graphics, 14 pages.

Appropriate representations (chapter 3). Designing interactive visualizations is not trivial. As visualization designers we need to ensure that the representations or systems we propose work well with real data and support user tasks that may be complex. And while the utility of systems designed around real user needs is undisputed for the users themselves, it does not necessarily help us understand the mechanics of visually perceiving and understanding the presented information. Visualization research can approach the question of what are appropriate interactive visualizations from different perspectives. For example it can start by considering the end-user and their needs when introducing new designs. Or it can start by the properties of the data and the tasks we want to perform on them. It can be motivated by the querying algorithms or other available technologies. Or it can seek deeper understanding of fundamental properties and impact of visual representations in order to provide design guidelines.

This chapter presents my work on creating appropriate visual representations, that originate from different perspectives. It first describes work that attempted to address needs for specific users, starting with Business Intelligence Analysts [**EB11**, **EB12**, **EAB13**], work done during the PhD work of M. Elias that I co-supervised with M.-A. Aufaure. The chapter then briefly describes our work with genealogists [**BDF+10**] (in collaboration with P. Dragicevic, J.-D. Fekete, J. Bae, B. Watson), and neuroscientists [**GTPB19**] that part of the PhD work of A.Gogolou that I co-supervised with T. Palpanas and T. Tsandilas. Our work with Business Intelligence analysts on storytelling received the Brian Shackel Award in INTERACT 2013.

Next, I report on work around fundamental questions on how we make decisions using visualizations, which is the PhD work of E. Dimara [**DBD17a**, **DBD18**, **DBD14**] that I co-supervised with P. Dragicevic, and later work on a taxonomy of cognitive biases and how we can mitigate them [**DFP+20**, **DBBF19**] (with E. Dimara, P. Dragicevic, S. Franconerri, C. Plaisant, G. Bailly). I also summarize investigations on perception for specific visual representations such as glyphs [**FIBK17**, **FIB+14**] (with J. Fuchs, P. Isenberg, D. Keim and E. Bertini) and line charts [**IBDF11**] (with P. Isenberg, P. Dragicevic, and J.-D. Fekete). The work on attraction effect with E. Dimara received an Honorable mention in IEEE VIS 2018.

Finally, the chapter describes work that started with particular visualization systems GraphDice [**BCD+10**] and EvoGraphDice [**CBL12a**, **BTBL13**], both very different, but inspired by an existing tool [EDF08]. I explain how we adapted and used them in different contexts: for social network analysis (with F. Chevalier, P. Dragicevic, F. Chevalier, N. Elmqvist, J.-D. Fekete) and in exploring very complex multi-dimensional data with the aid of automatic learning (with N. Boukhelifa and E. Lutton). The chapter also reports on research that started when with my colleagues we considered specific tasks such as correlation and transmission in spatio-temporal data [**PPB20**, **PBP20**] (with V. Peña-Araya and E. Pietriga).

The chapter concludes with reflections on challenges faced, on the very different methodologies followed, as well as inspirations behind the work.

• Representative paper (work with end-users): M. Elias, M.-A. Aufaure and A. Bezerianos (2013). Storytelling in Visual Analytics Tools for Business Intelligence. INTERACT 2013 -IFIP International Federation for Information Processing, Part III, LNCS 8119, (18 pages). **Brian Shackel Award** (Best Paper).

• Representative paper (fundamental questions): E. Dimara, A. Bezerianos, P. Dragicevic (2016). The Attraction Effect in Information Visualization. In IEEE InfoVis 2016 - the IEEE Transactions on Visualization and Computer Graphics, 23(1), (10 pages). **Best paper Honorable Mention** (4 best papers of conference).

• Representative paper (system): N. Boukhelifa, W. Cancino, A. Bezerianos and E. Lutton (2013). Evolutionary Visual Exploration: Evaluation With Expert Users. Computer Graphics Forum (EuroVis 2013), Eurographics Association, 2013, 32 (3), (10 pages).

• Representative paper (data & task): Vanessa Peña-Araya, Emmanuel Pietriga, Anastasia Bezerianos (2019). A Comparison of Visualizations for Identifying Correlation over Space and Time. In IEEE InfoVis 2019 - the IEEE Transactions on Visualization and Computer Graphics, 26(1), 10 pages.

Perspectives (chapter 4). The last chapter concludes with a set of future directions and some closing remarks on visualization inspiration and research.

2 Collaborative display environments: improving infrastructure & understanding their use

Computer input and output technology has evolved in the last decades, allowing for computer environments that range from small mobile devices (phones and smartwatches) to very large displays of several meters (high-resolution wall-displays, digital whiteboards and digital tabletops), with trends towards even smaller and larger sizes. Collaboration environments, such as industrial design or command and control centers, incorporate wall displays of over 5m x 2m in width and height, often coupled with digital tabletops and smaller whiteboard size displays.

Wall-sized displays (e.g., [BLHN+12]) offer several benefits for data analysis: their large physical size and high pixel count allow for the simultaneous viewing, comparison, and exploration of large amounts of data. As such, they have been identified early on as intriguing platforms for data analysis [BNB07, RJPL15, RWM+15]. These previous studies provided early evidence of an increase in the number and quality of insights gained when using these high pixel displays compared to traditional displays [RJPL15, RWM+15].

Despite their promising nature for data analysis, wall-sized displays present unique challenges when it comes to how to interact and view content on them. Given their high resolution, data may be viewed at close proximity to see details, or from a distance to gain an overview [AEYN11, BNB07]. The importance of physical navigation in front of such displays has been emphasized when it comes to improving spatial memory [JSH19] and user performance in sense-making tasks [BNB07]. This need for mobility poses challenges in how to best support *interaction* during data exploration, since using keyboards and mice hinders mobility. Moreover, it raises the question of how our *perception* of data may be affected given that viewing perspective can change as we move in the physical space. These aspects are discussed in section 2.1 and section 2.2 (work conducted with colleagues from Inria and Univ. Paris-Sud).

The large physical size of wall-displays also makes them well suited to *collaborative analysis* as they can comfortably accommodate multiple viewers around them [JH14, LCBLL16, LKD19], with researchers often studying collaboration quality and movement strategies. Nevertheless, questions remain when in comes to their integration in work settings and their impact in analysis tasks. The last section 2.3 of this chapter discusses the use of wall-displays in collaborative data analysis and command and control contexts (work in most part done within the PhD of A. Prouzeau).

2.1 Interaction with wall-displays

High-resolution wall-sized displays [AEYN11, BLHN+12] allow viewing a large amount of visual information, and thus have applications in a wide range of domains related to visual data analysis and exploration, such as healthcare [RTR+16] or command and control [SBMR12]. Nevertheless, choosing appropriate techniques to explore data in such environments is not a simple matter. Viewers are often mobile, thus traditional mice and keyboards are less adapted for input.

My doctoral dissertation work in University of Toronto started looking at this input challenge. In particular, I considered how to interact when at close proximity to such displays, so as to gain benefits of direct touch or pen input. While it is possible to envision a touch-based interface for visualization applications, we need to consider that on wall-displays not all parts of the display are easily visible or reachable through touch without significant physical movement (too high up or at the other side of the display). To address the problem of quick access to remote areas of a wall display, in my PhD I designed several interaction techniques that either bring interactive proxies of remote content close to the user [**BB05b**], or aid layout management and context switching by providing fully interactive views (portals) to different remote areas of the display [**BB05a**, **Bez07**].

Nevertheless, during visual analysis in front of wall-displays, it is natural to also move away from the display to get an overview of complex visuals, and coming up-close to see details [AEYN11]. To help such transitions, and interaction while at a distance, with my colleagues we consider mobile interaction alternatives (work conducted since joining U. Paris-Sud).

Existing work has considered mobile devices as a means to interact with content on the wall display, such as pointing [MI09] or panning+zooming [NWP+11]. Nevertheless, this work is fairly limited to basic navigation or pointing, ignoring complex interactions that take place when exploring visualizations that often require dedicated widgets (e.g., filtering using sliders, loading datasets, click and drag when selecting data points, annotation and tagging of content, etc.). We thus need to provide mobile interfaces that can support complex interactions. Recent work introduced custom-made mobile interfaces to match specific visualizations seen on a wall display. For example Kister et al. [KKTD17] show parts of a larger network visualization on a mobile phone or tablet that can also act as input, and Horak et al. [HBED18] render personal views of the data on an interactive smartwatch. Nevertheless, this requires considerable effort both in design (the visualization designer needs to carefully consider how the mobile and wall visualizations blend together), and development (develop both for the wall and the mobile device). Our approach is different, we instead support a flexible way to provide complex interactions using mobile devices, in a way that is easy to setup, develop, and use with different wall-display applications [**CBF14**].

Exploring complex datasets may require access to a large number of interactive controls in order to manipulate multiple dimensions and adjust their visual parameters. Rendering all of them on a single interface is problematic in our context, given that the mobile device is limited in space. We suggest instead allowing analysts to decide what controllers they need, and allowing them to customize their parameters. Work has also combined mobile devices with tangible controllers for data filtering [JDF12a], that analysts map on the fly to desired dimensions when exploring visualizations seen on a wall display. Or introduced tangible sliders [RCJN17] other deformable controllers [RCP+16] for eyes free interaction. Our approach is not tangible, but allows for flexibility in customizing parameters of controllers [**TBJ15**] .

We next discuss our two pieces of work that allow for **flexibility in customizing the interaction interface** from the **designer's perspective (Smarties)** and from the **analyst's perspective (SketchSliders)** when it comes to mobile interaction with wall-displays during visualization analysis.

2.1.1 Smarties [CBF14] - Input flexibility for visualization designers

Smarties [**CBF14**] is motivated by our own frustration, as designers of visualizations for wall-displays, with the lack of input support. With O. Chapuis and S. Franzeskakis we designed the Smarties toolkit to ensure we can customize the input for our different wall applications, without having to reprogram the interface on the mobile device side.

Smarties is an open-source framework that combines a mobile input interface that can be customized from the wall-display side, and a communication protocol between multiple mobiles running the interface and the wall application. The library hides completely the communication between wall application and mobile device(s) from the developer, allowing for fast development of input support for wall applications. It also provides collaborative interaction support with just a few lines of code (the framework supports multiple synchronized devices). Finally, it allows the customization of the mobile interface with very simple instructions in the wall application side, without having to modify the code on the mobile devices.

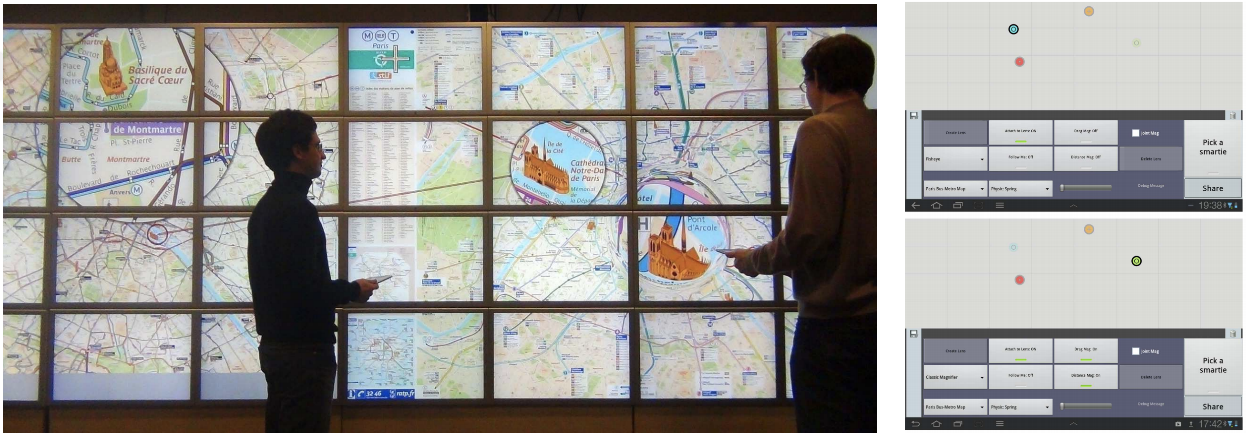


Figure 2.1: **Smarties.** On the left an example of a wall-display with multiple lenses on a map visualization (a magnification lens, a DragMag and a fisheye). Each analyst is holding a mobile tablet. On the right, the mobile client interfaces running on the tablets. In each tablet, the round colored pucks are attached to the position of the different lenses. The active puck is the blue one for the device on top, and the green for the bottom since each analyst is interacting with a different lens. The widgets added for this specific application are seen on the widget area at the bottom of the tablet interface.

The mobile input interface includes several components for complex interaction and is highly customizable by the designer without writing code for the mobile device. By default the mobile interface is divided into two parts (Figure 2.1): the top is a trackpad-like area that acts as a proxy of the entire wall on the top; and the bottom is a widget area. In the widget area the designer can include any number of specialized widgets such as buttons, text fields, menus, or sliders, all with programmable interaction behavior (e.g., program a button for gathering selected items, or a slider for filtering data). The proxy area for the wall can act as a touchpad (with support for multi-touch gesture recognition). It can also include pucks, round interaction elements that by default act simply as cursors on the wall-display. With a few lines of code pucks can instead as shortcuts to data items or visual content on the wall display (e.g., a shortcut for acting on a set of selected items, or for repositioning a lens). Pucks allow for persistent work (e.g., create multiple selections of items), can be stored and shared with other users.

Smarties are natively multi-user. Without any additional programming, multiple devices can connect to the same wall-display application and are by default synchronized. So puck location is reflected to all devices and so is the state of all widgets (e.g., sliders for filtering). By default analysts can share their work with each other by exchanging packs (that can be attached to content such as item selection, lenses, etc.), although the designer can decide to deactivate sharing.

In our original publication we demonstrated, through 3 application examples, how the system supports very different applications. These included a map visualization with lenses, a clustering application, and a windowing application. Each had different interaction needs, developed using different wall display software technology. Since then we have also used the toolkit in many of our own projects as the main input, for early prototyping, as a control interface for running experiments, etc.

Smarties focuses on the input side only, offering an integrated system for adding mobile input to wall-display applications, with a library that hides the complexity of the protocol for communicating with the mobile device, for quicker input prototyping. It also comes with a ready to use (but customizable) input interface running on mobile devices, that can handle advanced input in the form of widgets and multi-touch gestures, as well as collaborative interaction. It thus allows visualization designers to easily add and customize the interaction controllers required for their wall-display visualization. Nevertheless, we acknowledge that as it is a general purpose input toolkit, it cannot provide visualization-specific rendering on the mobile device side the way dedicated mobile software does (such as [KKTD17, HBED18]).

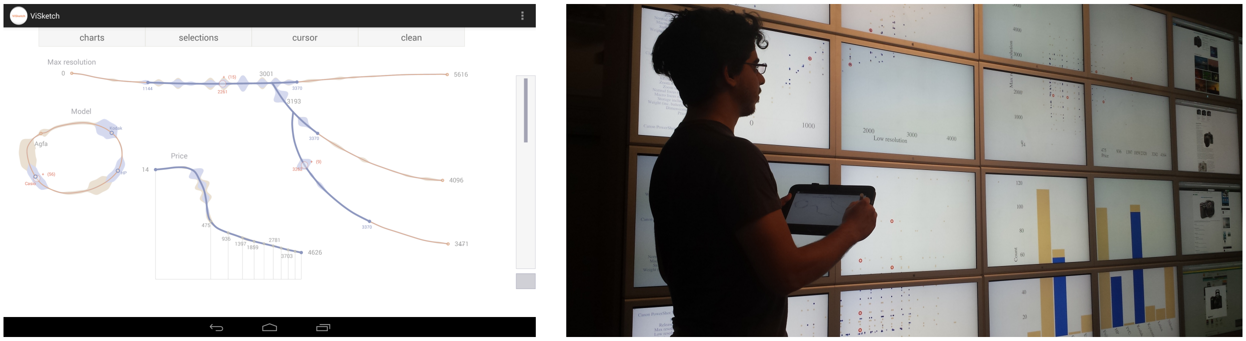


Figure 2.2: **SketchSliders**. On the left a screenshot from the user's tablet, where they sketched the controllers they need to explore a multi-dimensional dataset. They created a circular slider for periodic data (left), a branching slider (top), and a transformation slider (bottom). Slider variations in width express the data distribution. Light blue segments indicate selected points in the distribution and dark blue lines are the ranges for these range sliders. Values of importance (bends, user selections, endpoints of ranges) are shown on the sliders. In the branched slider, the lower branches provide a more detailed view of a dimension. The transformation slider provides is a focus+context widget (different interaction accuracy) and applies a transformation on the data. The right image is that of the analyst using the mobile device running SketchSliders to interact with visualizations on a wall-display.

2.1.2 *SketchSliders [TBJ15] - Input flexibility for visualization analysts*

Our work on SketchSliders **[TBJ15]** with T. Tsandilas and T. Jacob also decouples control and visualization in the wall environment, putting the interaction on a mobile device. Instead of having users interact with a large set of predefined exploration widgets decided by visualization designers, we now let analysts customize their exploration interface by *sketching* the controllers that best suit their needs. Visual exploration and analysis tasks [AS04] often include filtering and range-selection. These can be expressed with selections and dynamic queries [AWS92] performed using slider widgets mapped onto a data dimension. So to demonstrate our approach we focus on range slider controllers (SketchSliders). Users can sketch directly the interactive sliders they require to conduct data exploration on the wall-display. These controllers can be sliders of arbitrary forms, including circular, branched and transformation sliders, that support complex queries over multiple data dimensions and multiple levels of control granularity (Figure 2.2).

Sketching has been used in visualization as a rendering style in our own work **[BBIF12]** and by others [WII+12a], for expressing desired patterns when querying data [HF09, MA18, RLL+05, Wat01], or for rapidly generating charts [CMvdP10] and data stories [LKS13]. These approaches do not use sketching as input for data exploration as we do. More relevant is SketchVis [BLC+11] that allows users to create visualizations using sketching, but also provides simple sketch interactions to switch between different data views, select specific data categories, or apply simple functions such as averages or maximum. We focus on more complex exploration, studying in detail one type of interactive controller, the range slider.

In our approach, any drawn line can be turned into a slider through a simple menu. When this is done, the analyst is prompted to map the slider to one of the possible dimensions of a visualization on the wall display. We support several slider designs, that are inspired by Wizard-of-OZ and participatory design sessions with three visualization experts. Their feedback reinforced our choice of focusing on range sliders as the most versatile controller. Next we list the possible benefits of using sketched sliders raised by our experts, and explain how they were included in our prototype that runs on mobile devices.

• *Customization/Special Shapes.* Each stroke is unique and can encode information inside it (e.g., bending the slider in a given point of interest). Thus in our system any line can be turned into a slider, and simple crossing gestures can be used to add multiple range controllers on them. Detected bents/angles are highlighted and the slider value at that point is shown. Analysts can sketch sliders in any shape they want, including circular sliders to be used in the exploration of periodic dimensions.

• *Granularity.* The length of sketched slider affects the level of its control on the data. Long sliders allow fine-grained control and more precise filtering of data, while smaller ones allow coarse-grained control. Users can "graft" a fine-grained slider on top of a course-grained one when finer control is required.

• *Parametrization.* Sketched sliders can support multiple ranges, giving users the possibility to filter the visualization in a discontinuous manner. Another way of parametrizing a slider is to write by hand possible slider extremums or link specific data values to specific locations of the slider, which controls completely the mapping between the slider and the data.

• *Reusability.* By deactivating controllers (that remain accessible for later use) they keep copies of their sketched components, allowing them to explore alternative aspects of their data without losing past work.

• *Annotation.* A sketching environment naturally supports annotating and bookmarking of important information that is crucial for long term analysis tasks. SketchSliders combine free-form notes taken by users, as well as traces of the interaction exploration (controllers and values), that led to specific insights.

• *Transformation.* Inspired by our analysts sketched curves to communicate mathematical functions, we also allow the creation of transformation sliders. We apply focus+context transformation functions that affect the visualization of the plots on the wall-display: peaks of a slider curve represent areas of focus while valleys represent areas of context. The transformation applies both to the visualization on the wall and to the slider itself (values are sparser around peaks and denser around valleys).

Beyond these characteristics we included the following functionality. First, the slider width renders the density distribution of data in the form of a violin plot [HN98], inspired by scented widgets [WHA07]. Second, we provide a view in the mobile device, that acts as a trackpad in order to create a cursor on the wall-display and select data items. The cursor on the wall-display is a circular area cursor with *excentric labeling* [FP99], for previewing and selecting data points directly on the plots. Users can pinch with two fingers to resize the area cursor and reduce or increase the active area of selection.

To validate our prototype we conducted a user study with another six visualization experts, with experience from 2 to 15 years (median 10). They conducted an open-ended exploration task in front of the wall-display, interacting with a mobile tablet running SketchSliders. The experts re-iterated the customization value of sketched widgets *"There is something very compelling about sketching your own tools"*, and the freedom it *"this is vertical [indicating a dimension on a plot on the wall-display] so I drew a vertical slider"*. They also highlighted the flexibility of the approach *"I can focus either with branches or transformation, you don't have that in other interfaces"*. They also commented on how the sketching can help structure their analysis process and thoughts: *"[I] focus on one dimension, and see the result on other dimensions [using the slider distributions], which reduced my cognitive load - helped me filter not just my data, but also my controllers"*. This sentiment was shared by another expert *"existing interfaces are too cluttered and it is hard to decide where to focus on, here I can focus on one thing, and drawing a slider is a way for myself to decide what will be most likely of interest next"*. More generally, they found the combination of wall-display and sketching controllers as *"very well integrated"* and that *"the setup really works well for me, they are complementary, as you cannot show all the information on the tablet, and I don't really want to directly interact with the wall all the time"*.

We showed how with a small gesture vocabulary users can creatively use sketching to create their own controllers of various sizes and shapes, focus on parts of the data by changing the control resolution in dense areas of data, explore variations of controllers by grafting alternative paths, and bookmark important results and points. Due to the nature of sketching, users can naturally customize the controller's appearance and its effect on the exploration. For example, they can draw larger sliders and branches for a finer control, circular sliders for periodic data, or shapes that describe transformation functions to focus on a smaller range of the data. We were motivated by a scenario where users view large datasets on wall displays, and require mobile interaction support. Nevertheless, our experts commented that sketching customized controllers can be useful even in desktop settings, where visualizations are usually laden with numerous inflexible controls. This is a clear topic for future research.

The SketchSliders work received an **Honorable Mention** (top 5% of all papers) in ACM CHI 2015.

2.2 Perception on wall-displays

Physical navigation is an important means of accessing visualizations on wall displays [BNB07, EALN11, YN06]. Viewers choose close or far viewpoints to zoom in and out, and pan physically by moving left and right to see different parts of the display. In my PhD work I examined the notion of change blindness [Ren02] in large display environments, examining if viewers can identify and understand dynamic visual changes happening potentially outside their visual field. For example given the size of the display, if they are close to the wall they may not be able to see the edges of the display. In that work I proposed a set of techniques that identify changes that may have been missed, store these changes, and reveal them to the user at a later stage **[BDB06]**.

But even for non-dynamic visual information, there are still questions surrounding the visual perception of information in such environments. As viewers fluidly and frequently switch viewing distances and angles, this may lead to discrepancies between the actual displayed information and its perceived appearance. This question has been in isolated display cases, but not walls-displays. For example, in digital tabletops, researchers have assessed the relationship of view position and 2D object rotation on coordination, and comprehension [KCSG04], and the perception of simple charts from varying viewing angles [AJI10]. While Wigdor et al. [WSFB07] investigated how varying screen orientation from a horizontal to up-right position influenced the accuracy of perception, but they only considered static participants and smaller display sizes. We studied instead how viewers' visual perception changes depending at different positions in front of the wall **[BI12]**. These discrepancies have implications for a single analyst comparing content at different display locations, but also in situations where more than one analysts are trying to establish common ground.

Nevertheless, differences in perception based on viewer position are not necessarily bad, if the visualization designer controls them. With my colleagues we purposely altered visualizations in way that when presented on a wall-display, they show different information depending on viewing distance **[IDW+13]**. These visualizations, that we call hybrid-image visualizations, can be used to enhance overview tasks from a distance and detail-in-context tasks when standing close to the display.

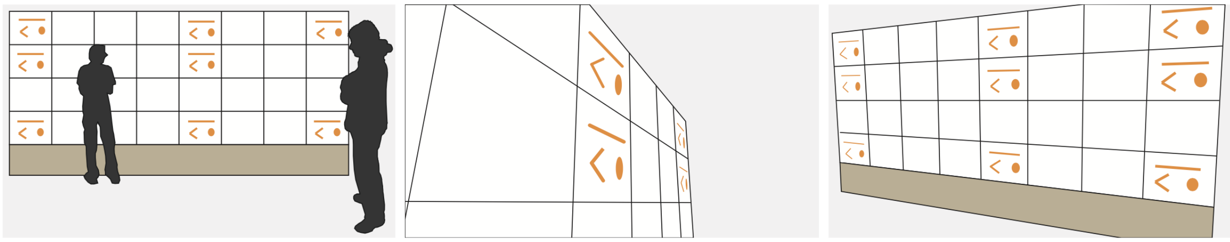
We next discuss in more detail these pieces of work, focusing on **how we perceive visual information** depending on different viewing positions in front of a wall-display, and on **how we can encode additional visual information** on visualizations that change with viewing position.

2.2.1 *Magnitude Tasks [**BI12**] - Understanding visual perception on wall-displays*

As mentioned, when dealing with large wall-displays, physical navigation becomes an important means of accessing visualizations [BNB07, EALN11, YN06]. Viewers choose close or far viewpoints to zoom in and out, and pan physically by moving left and right to see different parts of the display. Thus, they fluidly and frequently switch viewing distances and angles which may lead to systematic discrepancies between the actual appearance of displayed information in *physical* space (as can be measured by rulers) and its psychophysical appearance in a person's *visual space*.

Understanding these discrepancies and where / when they occur is important for visualization design, as fundamental data analysis tasks involve the correct assessment and comparison of elementary visual variables such as areas, angles, positions, slopes, or lengths [Cle85]. To read a bubble chart, for example, one has to compare the sizes of circles to one another and to a legend, as well as to relate positions in a $2D$ coordinate space. Figure 2.3 gives an example of how the appearance of three visual variables is affected when seen from different viewpoints and viewing angles on a wall-display. As researchers working with this new technology, we were curious to see whether comparisons such as these are affected by the oblique viewing angles which occur when viewing data from different positions in front of a wall-sized display.

Information understanding has been studied in other types of displays. For example, in digital tabletops, researchers have assessed the relationship of view position and 2D object rotation on coordination, and comprehension [KCSG04], and the perception of simple charts from varying viewing angles [AJI10].



(a) The viewpoints of two observers. (b) Wall as seen by the left observer. (c) Wall as seen by the right observer.

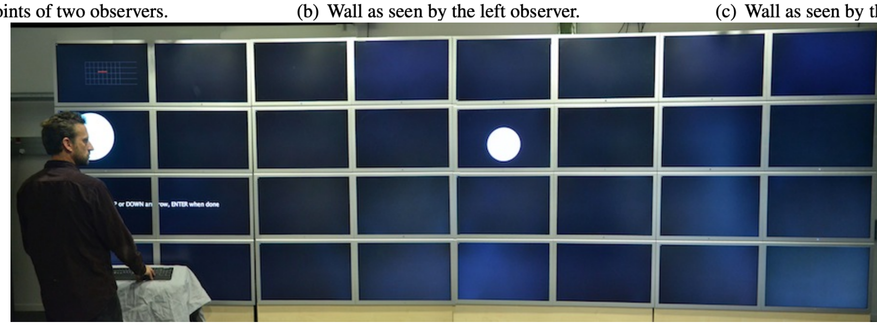


Figure 2.3: Two observers looking at the same angles, lengths, and circles displayed on a large wall display, from different positions (Top). Experimental setup, the large interactive stimulus is close to the viewer that can adjust it to match the remote modulus (Bottom). In a first study participants were not allowed to move and we tested two distances from the screen (Near distance shown here). In a follow-up study they could move freely.

In work related to ours, Wigdor et al. [WSFB07] investigated how varying screen orientation from a horizontal to up-right position influenced the accuracy of perception. With my colleagues we examined visual perception on wall-displays under conditions where participants were both static and could move [**BI12**]. Our findings differ from the similar study on tabletops. These perceptual differences across displays indicates that we cannot generalize findings from one display to another, nor can we predict how information will perceived when spread out across displays in complex environments.

With P. Isenberg, we began addressing these questions by studying how perception of elementary visual variables (Angle, Area, Length) was affected by varying viewing distances and angles. We contribute two studies. The first assessed *static viewing* conditions and identified different parameters that can help predict the perceived magnitude of the tested visual variables. The second contributes an understanding of the influence of allowing participants to *move* in front of the display.

In the first study (static) participants were placed at either 60cm (close - desktop monitor distance) or 320cm (far - distance where the entire wall is in the visual field). In the second study (moving) participants could move freely. In both studies our participants had to look at a remote modulus object (Angle, Area, Length) and reproduce it in front of them, using an interactive stimulus shape whose magnitude they can adjust. Before running the experiment we had calculated the visual angle of all remote modulus objects and formed hypothesis based on these visual angles. For example, we expected areas to be underestimated on average, while angles oriented towards the biggest axis of distortion (left-right distance) would be overestimated (their line segments look smaller and they will seem more obtuse). We also expected errors in estimation to increase with left-right distance, an effect that has not been observed in tabletops [WSFB07], as we are the first to test such large left-right distances.

In the static experiment (15 participants), our analysis shows no significant difference for completion time, but there is a difference for two error metrics: one calculating the amount of error, the other whether the error was an over-estimation or under-estimation of the true magnitude. The absolute error follows the ordering reported in tabletops [WSFB07], with angle being the most error-prone visual variable and length the least. Error also increases when viewers are close to the screen, and when the distance to the remote object increases, with angle being again most affected. Moreover, the absolute estimation error decreases with the increase of object size. The rate of decrease is more steep for angle (and somewhat less for Area), until errors become similar across visual variables for the largest object sizes.

The nature of over- or under- estimation is different per visual variable: angle is consistently overestimated, whereas length and area are less consistent in their tendencies. The generally observed nature of overestimation is less pronounced in lower areas of the wall. Nevertheless, as we move upwards on the wall the overestimation becomes more pronounced for angles and areas. This indicates that lower parts of the wall-display may be perceived differently, a finding that may relate to physiology literature that identified differences for visual activities in the upper and lower visual fields [Pre90].

In the follow-up walking study (with 9 participants who had taken place in the first study), participants could move freely. Here accuracy for estimations improves when participants are allowed to move, although the task completion time was more than twice as long in the static study. Three moving strategies emerged during our experiment. An overview strategy: walking to the center of the display, but at a far distance. A move-to-target strategy: walking until they arrived almost in front of the remote modulus. And a step-back strategy: moving slightly backwards from their original position (1m) to look at the remote modulus. This step-back strategy was the least effective (almost twice the amount of errors).

Overall, we demonstrated that our perception is affected depending on where information is displayed on the wall. When viewers are up-close, their judgement accuracy for angle encodings, and to a lesser degree area ones, starts to drop for targets placed roughly at the center of our wall (3m from them), whereas length is more robust. Distortion also seems asymmetric between the higher and lower parts of the display. Appropriate movement strategies can mitigate these effect, stressing the importance of physical navigation.

2.2.2 *Hybrid Image Visualization [IDW+13] - Encoding additional information on wall-displays*

We next consider situations where we'd like to enhance differences in perception depending on viewing positions around a wall-display. Our work is again motivated by collaborative viewing situations, where several viewers are situated in front of high-resolution wall-displays and can pan and zoom into the data simply through locomotion. Physically zooming-out (to get an overview) can have several effects on the perception of the data encodings. First, parts of the data will be perceptually grouped, forming visible clusters and will be perceived as a unit [Gol99]. Second, parts of the data previously visible will be lost as they reach the limits of visual acuity or contrast sensitivity [War04]. And finally, the perception of color [Sto12] and quantitative estimation of magnitude of visual variables can change (as our previous work demonstrated [**BI12**]). Thus the effectiveness of visualizations viewed from afar depends on the visual aggregation of particular data encodings and whether the data analysis task can benefit from visually aggregated data. To design effective visualizations for wall-sized displays, it is important to choose encodings that allow viewers to get an effective overview of the data from afar, while at the same time detailed information is available to viewers at a close distance from the display.

Some existing designs could be well adapted for such viewing, taking advantage of perceptual aggregation, where local features come together to form global patterns when seen from afar. For example, in a scatterplot, individual dots with carefully chosen background contrast and size, will begin to form clusters naturally at a distance. Previous work has encoded extra information on node-link diagrams by using specifically laid-out text to render both the links and the nodes of the graph [WMP+05]; has used automatic typographic maps where typography forms certain map features [AMJ+12]; or FatFonts [NHC12] where the typeface of the numbers is such that their area (amount of 'ink') is proportional to the represented number. Natural groupings and aggregations have also been extensively used in art, such as the famous Arcimboldo paintings of heads made up of vegetables or fruits, photomosaics [BDBFG07] where are made up by small colored fragments that are themselves images (tiles), and Calligrams where stylized text is arranged so it creates a visual image related to the text. One disadvantage of techniques

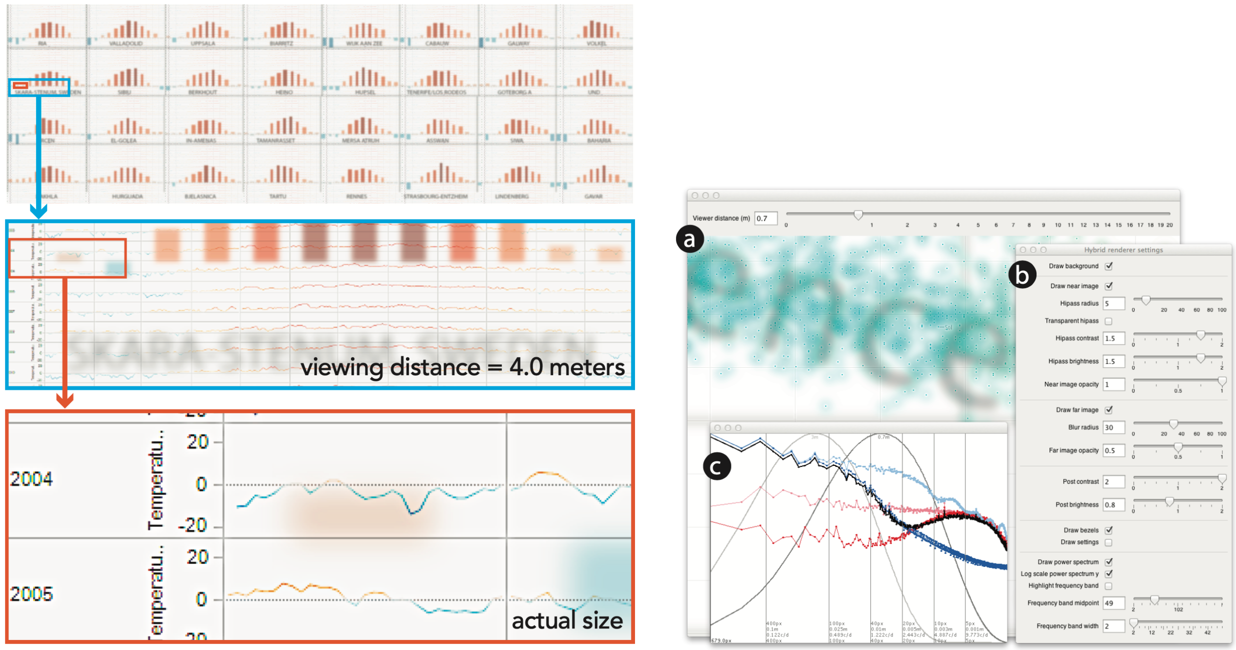


Figure 2.4: On the left hybrid-image visualization containing different representations at each scale. From a distance (top), bar charts show the average temperature by month between 1990-2012 for 32 cities. Up close, viewers can inspect and compare the individual temperatures by day using linecharts. On the right, a screenshot of our preview tools for reviewing and refining hybrid-image visualizations. Designers can preview the image as seen from different distances (a) and can use the control panel (b) and power spectrum visualizer (c) to refine their visualizations and tune how images are processed and blended.

based on visual aggregation is that the encodings are hard to control (layouts need to be crafted meticulously) and as several researchers have pointed out [EALN11, Sto12], we need more research on how features visually aggregate.

Other approaches separate global and local features, specifically assigning regions for fine and coarse-scale information. NodeTrix [HFM07] is an example that encodes small matrix visualizations inside the nodes of a large network graph. Lenses and more specifically focus-(and/in)-context techniques and radar views, provide windows that show detailed local information next to a global view. In these techniques the display is reserved for either close or far viewing distances, reducing the display space for either type of information.

In our Hybrid-image visualizations, with many colleagues from Inria, we attempt to resolve both these issues [**IDW+13**]. These visualizations (i) provide global information from afar and local information up close, in a way local and global views can take advantage of all the display space; and (ii) are easy to layout by bypassing the need to visually aggregate local views to create global ones. Our approach is inspired by hybrid images [OTS06]. Hybrid images blend two images after applying frequency filters to achieve distance-dependent perception. This works because the human perceptual system analyzes images at multiple scales through a collection of band-pass filters, each narrowly tuned to a specific frequency band [CR68, OS97]. We explore the use of hybrid-image visualization for data analysis in wall-display environments. A hybrid-image visualization can show two representations in the same view and without tracking viewers in the space: one is a detailed representation seen up-close to the display and one is a global representation gives overview information when viewed from a distance.

Through a set of examples (Figure 2.4 left) we show how hybrid images can be used to overlay two visualizations, one designed for near and one for far viewing, in a way that from close to the display, the encodings for far viewing distances practically disappear. Meanwhile, the information meant for close inspection does not hinder the overview from a far viewing distance. Hybrid image visualizations have the advantage that the whole display space can be used for drawing both local and global information. And the two overlaid visualizations can present very different encodings since there is no need for visual aggregation. They also do not require interactivity to work (for example compared to lens-based techniques that create areas of high/low magnification [TGK+17]) and thus do not disrupt other users.

Our work contributes a detailed discussion of the perceptual background and practical considerations when creating hybrid-image visualizations, example visualizations and a summary of encoding techniques, and a set of tools for creating them. Nevertheless, it also opens up questions related to how well these visualizations work in practice, and on whether this nice division between close and far viewing remains when they become interactive (as human perception of motion needs to be considered).

2.3 Collaboration around wall-displays

In our previous work mentioned here on input infrastructure and perceptual questions when it comes to viewers' movement, there is always the underlying assumption that more than one user may be present in front of the wall-display. For example Smarties subsection 2.1.1 support multiple users, while the magnitude study subsection 2.2.1 and hybrid visualizations subsection 2.2.2 consider the perspective of different users in the room. The importance of collaborative visual analysis has been acknowledged in the visualization community (see [IES+11, IIH+13] for reviews), and wall-displays are indeed a very attractive environment to conduct such work given their large physical size that can accommodate multiple people. Previous work has focused mostly on tabletops, and smaller vertical displays (SDG and whiteboards). For example, when it comes to understanding coordination around collaborative tabletops, researchers have explored how colleagues shift from tight to loose work coupling [TTP+06], how they divide the space (territoriality) [SCI04], how they coordinate document search [IFP+12], etc.

There are fewer works that consider co-located collaborative work on larger wall-displays. Studies on the behavior of a pair of users in a sense-making task [JH14] comment on how participants fluidly moved from parallel and group work using different parts of the display. Others [LCBLL16] look at different collaboration styles and interaction in a classification task. Or at how territories are formed when larger groups work together [ARV+12], commenting that the personal, storage and group territories that emerge are more dynamic and fluid than in tabletops, due to users being mobile. More recent work [LKD19] observes user behavior and movement when analyzing coordinated views of crime data.

Nevertheless, there are many questions that we have yet to answer when it comes to collaborative visual analysis using wall-displays. In our work we have focused on aspects related to coordination. More specifically, we try to determine **if wall-displays they better than regular desktops** when it comes to coordination, and **how coordination is affected** by factors such as the interaction technique used.

Other remaining question relate to challenges and opportunities when considering possible contexts of use for these displays. For example what are application domains that can benefit from them and how could they be integrated in more complex display ecologies that include individual devices. The second part of this section focuses on the **context of command and control rooms**, and how analysts could transition from using wall-displays as awareness monitors, to collaboration platforms.

These questions are the focus of work described next, conducted during the PhD thesis of A. Prouzeau, that I co-supervised with O. Chapuis. The chapter ends with a mention of my other work on collaboration.

2.3.1 Wall-displays vs. desktops [PBC17b]

Empirical studies support the idea that large displays foster collaboration. Nevertheless, studies that compare a shared interactive surface to individual devices have mainly provided qualitative results. For example, when comparing a tabletop with personal tablets for a sensemaking task, the table supports better prioritization and data comparison, and leads to more participation equality [WSM13]. In work comparing two users working on a large display, versus one on a large display and one on a regular screen, in a trip planning task [HKR+05], participants feel that collaboration was more enjoyable and efficient with the large display. And that communication was more difficult when having separate screens. when comparing a large display, with a large display combined with three desktops [WSS+09], the large display alone provides more awareness of partners' activities, but can lead to distraction. With the three desktops there is less distraction, but collaboration is more demanding. This previous work comparing a single large collaborative display with smaller multiple displays shows that the single large display provides more group awareness and aids communication. The downside is that using the shared surface can distract colleagues working together, and thus possibly impact performance.

This previous work focuses mainly on subjective measures and on characterizing the nature of collaboration when participants perform high-level tasks. Results related to performance mainly refer to quality of results. As researchers who conduct both qualitative and quantitative research, we want to investigate if we can provide *quantitatively* evidence of differences in coordination when using these displays. We thus decide to measure performance and coordination differences when pairs use a large display, compared to them using two desktops that share a common view (Multi-Display Groupware - MDG) **[PBC17b]**. The two desktops are motivated by setups where collaborators use individual workstations (e.g., command and control centers), that are often distant and cannot support deictic communication, but allow for verbal communication. In our setup the individual workstations share a common view of the virtual content.

A challenge in conducting quantitative research is the formation of hypothesis, and the isolation of possible confounding factors. To do so, we have to chose tasks that are abstract enough to avoid confounds (e.g., related to previous knowledge and familiarity with the data), being simple enough to repeat across conditions, while being complex enough that coordination is required to solve them (as this is what we study). We are inspired by previous work on studying collaboration under route planning tasks with constraints. Examples of such tasks include planning a route in a subway map [HKR+05], or creating bus routes that have to pass through specific locations and at the same time not overlap [TTP+06]. Our task is similar to the latter. Within a grid, each participant has to form a path between two "end-nodes". To encourage pairs to coordinate and make decisions, we enforced constraints in their planning: the two paths were required to cross at two specific nodes but nowhere else, and could not overlap (they could not share an edge). This type of constrained path-planning is an abstraction of resource-routing and planning tasks common in real situations, such as traffic control centers, for example during accidents operators can guide first responder teams to the location of the accident, and at the same time reroute regular traffic. Our abstracted path-planning task is simplified (simple layout and no road context) to reduce effects due to complex layout and due to context knowledge. It also has specific constraints to encourage coordination. To avoid learning, we varied the locations for constraints and starting nodes. All these characteristics together ensure that the task can be performed by participants without domain knowledge, and the findings can be generalizable as we limit effects caused by factors not related to collaboration.

Figure 2.5 shows both our experimental setup and an example task with its solution.

We expected our pairs to develop collaboration strategies over multiple trials, that likely differ across setups (wall-display vs. two synchronized desktops), eventually reducing the need for coordination and decision making that are essential in collaboration [McG84]. As such, we did not provide any training to our participants, but rather compared the learning phase across settings, as this is where pairs need to communicate and coordinate to improve their strategy. Learning rate has been used in the past as a measure of coordination [GG98a]. To study possible trade-offs between the setups, we also measured other metrics that could shed light to differences in collaboration, such as the amount of communication between pairs and their coordination strategies. Participants were recruited in pairs that knew each other (32 participants) and each pair used only one setup (between subjects design), as we did not want them to apply coordination strategies across setups.

Our results do not indicate a significant difference in learning between setups, but pairs are overall faster using desktops. With desktops, pairs divided the task as much as possible, requiring less com-

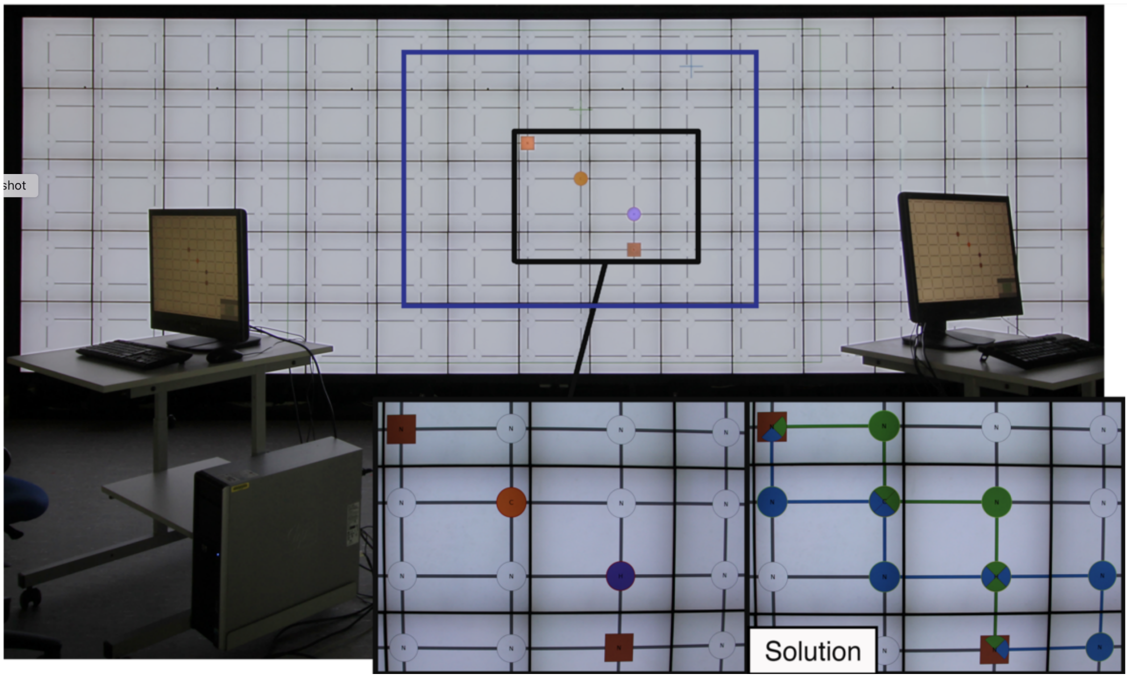


Figure 2.5: **Wall vs. Desktops Comparison.** Setup of the experiment, with both conditions: wall-display, and two desktops with a common view (the wall-display was turned-off in the two desktops condition and vice versa). The left cut-out shows a close-up of a task, and right a possible solution where the green and blue user's paths have met in exactly two places.

munication, but this seems to affect the quality of work. Indeed, the quality of the solution (defined as the number of corrections needed to reach an optimal solution that meets the imposed constraints), is more consistent with the large display. Pairs communicate and plan more ahead of time in this setup. Thus with the large display, participants adopt strategies that included more planning and coordination, which leeds them from the beginning to consistent, good quality results. In summary, it seems that when participants are faced with a new task, they do not adapt more quickly using the large display, but they can produce better results from the start. This observation may have important implications in situations like crisis management, and command and control centers, where collaboration on large displays could provide better quality solutions in unexpected crisis events.

2.3.2 *Impact of interaction techniques [**PBC17a**]*

Continuing our investigation of coordination during collaboration, with my colleagues we start considering different factors that may affect it. Given our own background work on interaction techniques in wall-display environments, we naturally thought of interaction as one factor that may affect how colleagues coordinate on a shared display. We initially prototyped our application and study using our Smarties toolkit [**CBF14**] while exploring different interaction techniques. Nevertheless, in the end we focus on touch-based interaction techniques. Our reasoning is that direct touch provides awareness of others and their actions, and thus would be a better candidate for coordination.

In this work [**PBC17a**], we focus on analysis of complex graph/networks. Graphs exist in various application areas such as in social networks, in molecular biology proteins interaction networks, in transport networks, etc. Graph structures are frequently represented as node-link diagrams, but they can be too wide to view comfortably on regular screen monitors [VLKS+11], and thus well adapted for viewing them in wall-displays. Moreover, collaborative analysis has been identified as one the next challenges of

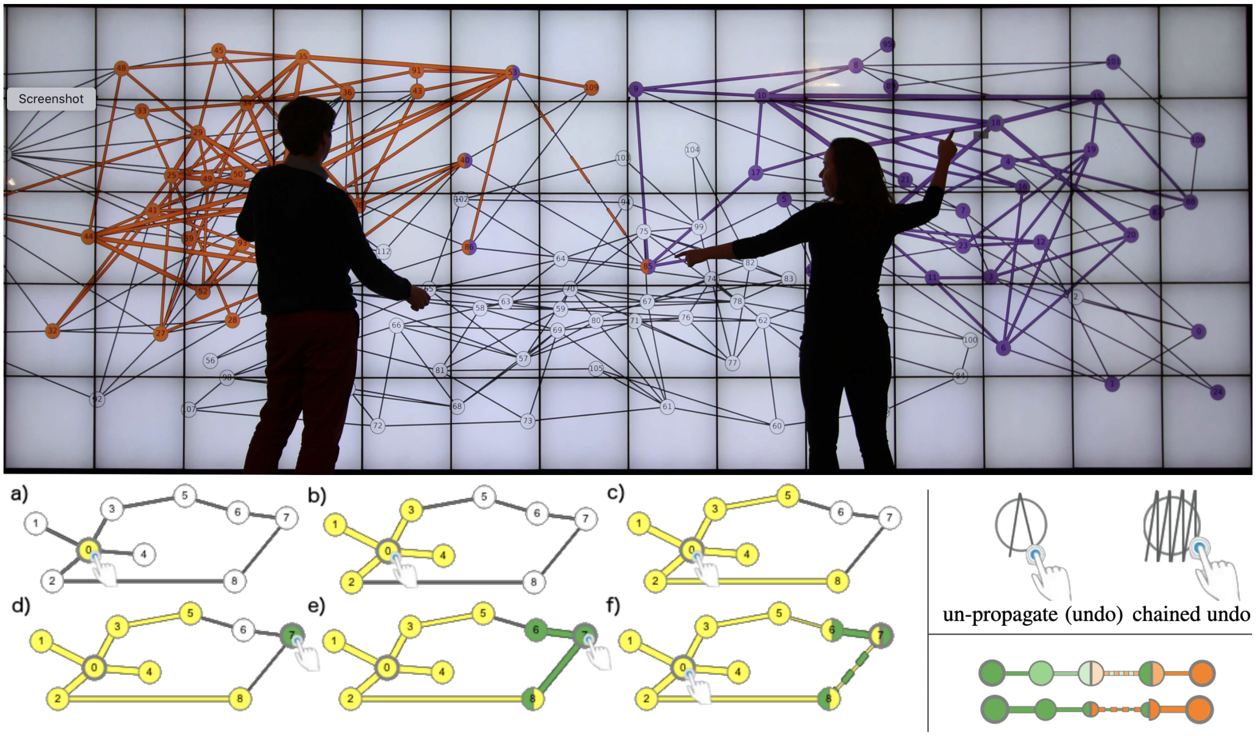


Figure 2.6: **Propagation Selection.** (Top) Two users in front of the wall-display using propagation selection. (Bottom) Details on propagation selection. On the left we see multiple propagations: (a) a first tap on node 0 selects it; (b) a second tap propagates the selection to immediate neighbors; (c) and a third tap to 2nd degree neighbors (notice the difference in link width according to distance); (d) a tap on node 7 selects it with a new color; (e) a second tap selects its neighbors, one of which (node 8) is shared with the first propagation and has both colors; (f) a fourth tap on node 0 propagates the first selection a third time, resulting in nodes 6, 7, 8, and link $8-7$ being shared between propagations, with the color and width on shared link $8-7$ alternating. On the right we see gestures to undo one propagation step on a node, or to do chained undo for backtracking multiple steps. The lower part shows variations for displaying propagation distance using color intensity (top) and node-link size (bottom).

the analysis of graphs [VLKS+11]. Existing graph analysis systems support mainly remote collaboration (e.g., [ZK14]). Less work has targeted co-located analysis of graphs on wall-displays. For example our previous work **[ICB+09]** retrofitted an existing graph visualization application to run for multiple users. Nevertheless, that work only supports multiple analysts with mice and keyboards, which cannot accommodate physical navigation around the wall-display. In the work presented in this section, we study the analysis of graphs by multiple users moving freely in front of wall-displays. Figure 2.6 shows our setup.

In particular, we looked at how pairs use a wall-display to solve topology based tasks, that are components of more complex graph analysis tasks [LPP+06]. Our main goal was to study how the choice of interaction technique supports or hinders pairs collaborating on these tasks. We focus on techniques for *selection*, as it is a pre-requisite to many interactions such as filtering, comparisons, details on demand, etc.

We adapt two general purpose graph selection techniques for use by multiple users on a touch-enabled wall-display. Our baseline selection technique is an extension of basic node/edge selection for multiple users, using simple clicks on nodes and links to select them. The nodes/links selected by multiple users are colored in a combination of their colors. This technique is easy to master, and has a limited, and thus fairly localized, visual footprint on the wall-display, that does not interfere with colleagues' work.

The second interaction technique is a propagated selection, that extends for multiple users the idea of transmitting a selection to neighboring nodes/edges [HB05, MST+14]. With each consecutive tap on a node, the selection progressively propagates to the neighboring edges of first-degree (Figure 2.6-bottom). Appropriate visuals indicate how far the selected links are from the node of origin, and how many users have selected them. This propagation interaction highlights the connectivity structure of the graph, but may have a large visual footprint that disturbs colleagues (depending on the connectivity of the graph, the selection can quickly propagate to the entire graph). We conducted two user-studies to compare these interaction techniques. As there is little work on graph analysis on wall-displays in general, we also studied an individual user context, to tease out effects due to collaboration and ones due to the techniques.

In the first study, we chose a well-defined topology task, the identification of the shortest path between two nodes, a task used often in graph studies [TTP+06, DHMM13]. Finding the shortest path can be fairly involved in complex graphs, as it requires an understanding of both the local context of nodes (identifying neighbors), as well as more global structure information since a shortest path is not necessarily small in absolute distance. It is also a task that is not clearly divisible, as a more global understanding of the graph structure is required. Thus it is unclear if multiple users working together would fare better than single users. And very importantly for our purposes, it does not bias against basic selection, as it is a task not trivial to do with propagation: propagation highlights a large number of possible paths (transmission can quickly cover well connected graphs), revealing issues with visual clutter caused by propagation. We also chose carefully the graphs used in the study. We consider two graph types with different connectivity: planar graphs (common in transport networks) and small-world graphs (common in social networks). In both cases we varied the number of nodes and links to create difficult and easy variations for each type.

We recruited 16 participants, that conducted the tasks alone and in pairs. Our results show that propagation is faster in both individual and multi-user contexts, and propagation is also more accurate in multi-user contexts. In single user contexts basic selection requires a lot of walking (slowing users down). And in multi-user contexts, with basic selection it is difficult to acquire an overview of all choices considered by one's partner, and thus maintain a global view of the work. On the contrary, with propagation it was easier to verify at a glance the work of one's partner and check for errors.

In a second study, participants used propagation to conduct other topology tasks related to graph connectivity, such as finding the shortest distance between two nodes (the number of links rather than which links), finding common neighbors, articulation points, and connected components (communities) [LPP+06]. Participants had taken part in the previous study, and were not given instructions on how to use the technique for these new tasks. All pairs were able to devise correct strategies for the majority of tested tasks (apart from articulation point), and even in this case they were able to identify good candidates for a solution (even if some participants could not provide proof as to why).

Overall, we observed that even when tasks are not clearly divisible, pairs divide the wall spatially. For many topology tasks identified in the literature, and used in our experiments, there is no clear strategy to divide them in space, as they require a global understanding of subgraphs that may extend across the display. Nevertheless, irrespective of task and technique, pairs divided the wall spatially. Even when not optimal, they each took responsibility of one part of the wall, and then combined their work. While space division has been observed in spatially divisible tasks [JH14, LCBLL16, THSG04], we have not seen it before in tasks that are not clearly spatially divisible. As designers, we should anticipate this division of space and encourage tighter collaboration when tasks are not spatially divisible. Techniques like propagation selection can encourage such tight collaboration (when compared to simple basic selection). As propagation has a large visual footprint, it requires higher coordination when used by multiple users, to avoid disruption of others' work. This tighter coordination leads to an increase in accuracy overall, irrespective of graph types and difficulties. When using basic selection, that has a small visual footprint, accuracy dropped for pairs, most noticeably in complex graphs. Here pairs tend to work independently and loose awareness of each other's work, which proved detrimental for the quality of non-divisible task.

There is thus a coordination / visual disruption tradeoff. Techniques with a large visual footprint (like Propagation) can visually disrupt and affect the partner's work, but can also promote coordination, providing higher degree of workspace awareness [DB92, GG02]. While techniques with small visual footprint (like Basic) are less disturbing, but pairs can loose track of each other's work due to the wall size and graph complexity. We should support both types of techniques, allowing colleagues to transition between them depending on how divisible their task is and what degree of coordination they require.

2.3.3 *Collaboration in traffic control centers [PBC16a, PBC18]*

Our previews work reinforced the trade-off between awareness of others for coordination vs. visual disruption, in particular when it comes to the quality of their work. There are situations where quality is extremely important, such as crisis management and command and control room situations. Previous work has highlighted the importance of awareness of others in these settings [MPBW07]. The goal of our work described next is to promote awareness in such settings and investigate how to best integrate wall-displays in them. I have mentioned command and control rooms as motivations behind our decisions (when we studied individual desktops vs. wall-displays [**PBC17b**]), as they represent extreme collaboration environments, where many operators may be present at any given time, and the quality of their collaboration is considered critical [MPBW07].

Understand user needs

With my colleagues we fist set out to understand the needs and practices of this user group, focusing in particular on *traffic*, as we had access to road-traffic operators through previous contacts. In 2016 we visited the two control centers in the city of Paris, one that monitors all traffic inside the city, and one that monitors the Periphérique, a motorway surrounding Paris. Together, they are responsible for 1500 Parisian intersections and its tunnels, with more than 2 million cars and 2.5 million pedestrian movements daily. We observed two operators in each of the centers, and interviewed two of them in depth (one per center) as well as an operator supervisor. All interviews lasted approximately 1h. We distill next findings important for our research (full report is available in A. Prouzeau's thesis [Pro17]).

• *Room Layout*. Both control centers are furnished with a large shared visualization wall-display showing the monitored network, surrounded by smaller screens with live camera feeds. Road segments are colored depending on traffic congestion from green (no congestion), to yellow, orange, and red (high congestion). Gray is used to indicate segments with faulty loop (traffic) detectors. Individual workstations are located in front of the wall, and they are also displaying the network visualization, alerts and other statistical information. This setup motivated our comparison of wall vs. desktops [**PBC17b**]. Due to the small scale and resolution of their monitors (w.r.t. the scale of the monitored network), operators tend to focus on localized areas of the network in their workstations, using mouse and keyboard to navigate. While they use the wall as an awareness monitor to acquire the "big picture" of the network state.

• *Forecasting and Traffic Plans.* An automated system manages the traffic lights, with a library of traffic-light plans (a collection of consistent traffic light durations that are automatically chosen depending on the current traffic situation, the day of the week and the time of day). To optimize traffic flow, operators can change traffic-light duration, activate/deactivate lanes, reroute drivers using variable message signs, and evacuate tunnels. Our interviewees explained that due to their experience, operators can accurately predict the impact of their actions and interventions (e.g., traffic rerouting) but only in a local scale, such as a crossroad. It is difficult to assess the impact of actions at a more global scale, for example it is often unclear how a change in a crossroad can affect the entire network. One operator gave a recent example, when a tunnel had partially flooded in both directions for 11 hours, but the operators did not risk closing it down as they did not have a clear picture of potential global effects on the rest of the network.

• *Coordination.* Crisis management is a good example of a situation where colleagues switch between parallel and collaborative work [LB16] - monitoring and intervention respectively. Our interviewees explained that they generally work independently, but during crisis they need to coordinate with other organisms, such as firefighters when a fire is suspected in a tunnel, police in cases of accidents or rerouting for special events (such as state visits), and first responders. Representatives of these organizations are often present in the center full-time (as is the case for police representatives), or may be invited in periods of crisis. In such times, the awareness of where others are focusing on is very important, and so is the sharing of resources. One example provided by an operator is the flood of the Loire crisis [FBDM+15], that involved more than 23 agencies including police, fire brigades, first responders, public transportation, managers of water/power/communication/road networks, and flood forecasting services. Each agency had their own tasks to perform (monitoring the road traffic, coordinating firefighters or first-responders, managing public transport, etc.). But they also needed to occasionally coordinate, for example an operator of the power company needed to guide her team to a damaged power unit in the fastest way possible, avoiding traffic and first responder units.

Our interviews and observations revealed two situations where wall-display technology could be beneficial: (i) provide a means for operators to **visualize predictions / forecasting** of the impact their actions can have on the network; and (ii) consider additional **awareness mechanisms to aid in coordination** in situations where operators have both their individual desktops and access to a large visualization wall.

Traffic forecasting using wall-displays [PBC16a]

This project is not related directly to coordination (our main research goal). Instead it stemmed from the needs of our users. As we were creating a prototype to test and demonstrate coordination in traffic-control rooms, we decided to explore it as well. Our interviews suggest that it would be beneficial to incorporate visualization of predictive models with real-time monitoring tools, as the impact of actions is often hard to predict. Operators should be provided with likely outcomes of their interventions both globally on the entire network, and locally on specific sectors or intersections. We thus considered combining forecast visualizations running different simulation models, with the general monitoring visualization of real traffic [**PBC16a**]. These forecasting visualizations are only needed periodically (so not constantly visible) when operators need to plan around an incident.

Large displays in current control rooms tend to be of low resolution and designed to be seen only from afar, as awareness monitors. We suggest instead replacing them with high-resolution walls that are interactive and can be seen up-close. For our prototype we decided to show the visualization of the prediction models on the wall-display, and test the limits of how many such prediction models users can comfortably monitor at a given time.

We created a prototype software for initializing predictive simulations based on the possible actions available to the operators. Operators can control the time-frame of the simulations, and the speed to play-back the results (described in detail in [**PBC16a**]). To show simulation results, we chose difference maps (as Lampe et al. [LKH10]). The colors of roads do not indicate an absolute measure of traffic density, but rather a positive or negative distance from a baseline situation (real traffic). We chose a blue-brown diverging color scheme, adapted for viewing on the wall-display, to highlight differences [VP04]. Our forecasting simulation model [CHP04] is an extension of the well-known Nagel and Schreckenberg one [NS92], but can be substituted by others.

We propose two techniques for viewing multiple simulations in combination with real traffic:

- *Multiple views* [JE12, Tuf86, WBWK00]. These show the global state of the network and are thus well adapted for situations where operators need to see the impact for the entire network.
- *DragMagic* an extension of DragMags [WL95] and magic lenses [BSP+93], to visualize localized sectors, and are better for showing local effects in a specific location of the network (Figure 2.7).

Nevertheless, the situation is more complex when operators need to consider several areas of interest (critical areas) on the network. Due to the higher number and sparsity of areas of interest, this task is neither clearly local nor global, and thus it is unclear which technique fares best. DragMagic likely works well for few areas of interest, but as their number increases they approximate the entire network, and as such MultiViews may be better. Finally it is unclear how hard it is to follow multiple simulations running

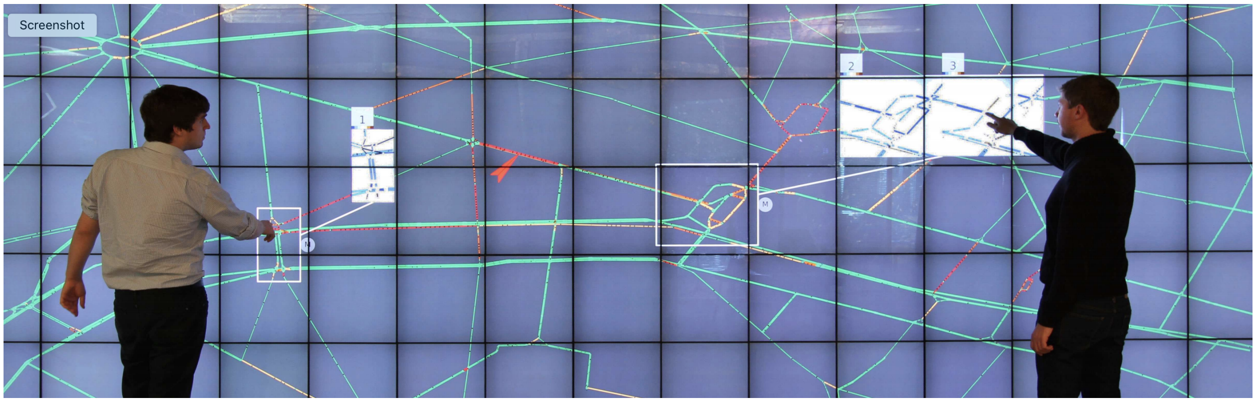


Figure 2.7: **Traffic forecasting.** Visualization of traffic in central Paris, with two DragMagic lenses (white rectangles) showing one (left) and two (right) simulations associated with different possible interventions on the traffic. The simulation visualizations use difference color maps to highlight differences from the real traffic. Blue hues indicate improvement, brown ones deterioration in traffic.

at the same time in order to decide between alternatives, using either technique. We thus designed an experiment to compare viewers' performance using DragMagic and MultiViews for this intermediate case, varying the number of simulations and areas of interest.

We designed a control experiment, where participants had to monitor 2-6 simulations, focusing in increasing number of areas on the map, that varied between $3 - 7$. To study how well participants could follow the simulations, we interrupted their monitoring at pre-specified intervals and asked questions on the present state of the areas of interest (best simulation currently), and on their past history (simulation with globally the best performance so far). These questions were inspired by tests accessing situation awareness (i.e., participants' awareness of the past and current state of the system) [ICH+13].

Our task is perceptual in nature (track color changes), and as such we run it with 16 non-expert volunteers for measuring perceptual situation awareness (similarly to previous work [ICH+13]). Our results show that DragMagic is easier to master (learning was faster), but that both techniques are good design options in terms of accuracy. We were surprised to observe that even for several simulations and areas of interest, the accuracy and time of the task remained almost constant. This is contrary to the model of Plumlee and Ware [PW06], that predicts a clear growing relation between time and the number of comparisons needed to perform the task, in particular when it comes to the increase of areas of interests. This can be explained by the temporal nature of our task. Based on their comments, participants continuously compared simulations in the time between our interruptions to ask questions, and were thus able to identify and ignore ahead of time non-promising simulations, providing answers more quickly. Thus, the Plumlee and Ware model does not extend to tasks that have a temporal continuity.

Expert traffic-control operators also provided encouraging feedback and suggestions after seeing the prototype, appreciating in particular the use of DragMagic to follow forecast simulations while keeping the context of real traffic. They also stated that the specific visualizations could be useful to compare real-time and historical data to help identify possible problematic situations.

Aiding awareness in control centers using wall-displays [PBC18]

As we saw from our interviews, crisis management is a good example of a situation in which operators often switch between parallel and collaborative work [LB16]. Previous work has already investigated the use of multiple devices in such situations. For example, Chan et al. [CASM16] developed an emergency operation center that combines a tabletop, a wall display and several tablets and wearable devices.

However, this previous work does not focus on providing awareness to aid transitions between the different displays when operators need to move from parallel work to closely coordinated actions. We decided to investigate how a wall-display could be integrated in such contexts as a means to improve workspace awareness (awareness of the work of others) but also as a surface where operators can collaborate closely. We propose techniques specifically aimed at identifying opportunities for close collaboration and for improving the transition between parallel and close work. In a similar vein, Bortolaso et al. looked at transitions in a military command and control context [BOP+14], focusing on tabletops. They introduced different types of lenses on the tabletop and on the tablets, to support workspace awareness in different collaborative configurations. We follow a similar approach on a wall-display setup, and additionally focus on techniques to facilitate transitions between multiple devices.

Based on our interviews with traffic-control operators, as well as documented crises (derailment of a freight train in a tunnel in Baltimore [Sty01], a helicopter crash in the center of London [BBC], and the flood of the Loire river in France [FBDM+15]) we draw design goals for such collaborative environments:

• *Different roles.* The environment should accommodate colleagues with different roles (from different agencies), and access to different information. Other types of collaborative contexts, such as sensemaking and brainstorming, can also bring together individuals with different expertise [GG98b].

• *Sharing.* Colleagues should be able to share only data useful for the situation. During crisis management operators have access to various data [LB16] that are specific to their roles, and that are important to share with others when coordinating. But access to data that are not relevant to the situation can confuse them and alter their understanding of the situation [Con93]. Thus it is important to be able to share specific data only. A need to only share content relevant to the collaboration also affects other collaborative analysis situations where colleagues bring with them their own data.

• *Current Opportunities and Past actions.* Ideally, colleagues should be able to serendipitously identify opportunities for collaboration and coordination. Thus it is important to have a good mutual awareness of where others are working on, and on what. Moreover, when colleagues work concurrently in different sub-tasks they partially loose awareness of others [GG95]. So when they want to transition to closer coordination tasks, they may need contextual information about the recent work and focus of their colleagues.

To this end we built a prototype, meant to support multiple users with different roles. It is composed of a very high resolution interactive wall-display, several workstations and other peripheral displays. Our vision is that wall-displays will act as monitoring and awareness infrastructure while colleagues conduct personal work, and as a surface where they can collaborate actively when close coordination is needed.

We designed three techniques to aid workspace awareness and help identify opportunities to transition from personal to close collaboration, by displaying information about the activities of others. Our designs vary with respect to where this additional information is placed in the environment (in the focus of the shared display or on the periphery), and with how long they are displayed (transiently or permanently).

The three techniques, shown in Figure 2.8 are:

• *Awareness Bars* show the current focus of other operators' workstations on the edge of the wall display only. They are permanently displayed, but on the periphery of the wall display.

• *Focus Maps* show the history of the areas of focus of seated and walking operators. These are activated on demand when colleagues are considering transition to close collaboration. And they fade over time.

• *Step Maps* are displayed on the floor of the room. These permanently follow standing operators and show their position in front of the wall, but also show only a temporary trace of the operators' position, with history fading over time.

Finally, we provided mechanisms to move personal content on the wall-display in order to share it with colleagues when close coordination is required. This sharing can be initiated on their desktops, on the wall-display and on mobile devices.

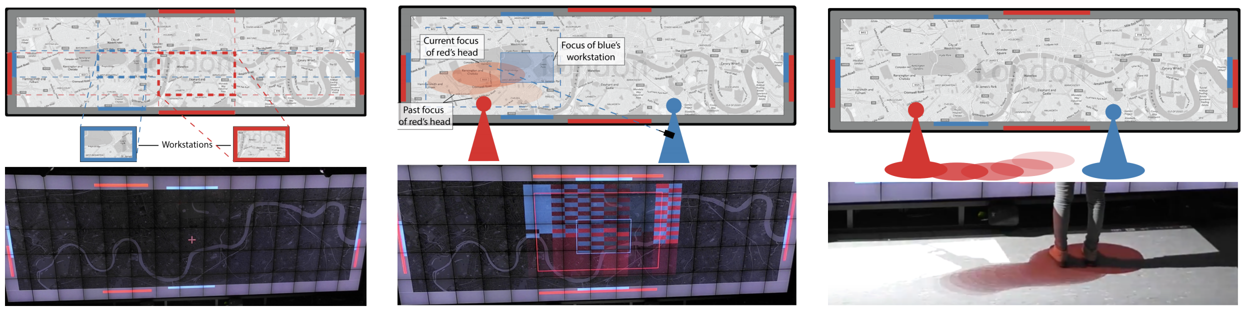


Figure 2.8: **Awareness techniques.** Top images illustrate the techniques on the wall-display, bottom images are pictures taken from the prototype. (Left) *Awareness bars*. The colored bars on the border of the wall represent the areas on which each operator is focused on on their workstation. These are always visible. The dashed lines are added for illustration purposes. (Middle) *Focus Map* on a specific area. In the illustration the blue operator explicitly asks for a focus map for a specific area (dashed rectangle) by selecting the area using a tracked smartphone (operation that can also be done from their desktop). The history of focus of both operators is displayed for this area, but fade over time. On the bottom a picture showing a temporary focus map for a specific area, with two operators' colors interleaved. (Right) *Step Map*. The red and blue operators work in front of the wall. The fading red circles on the floor indicate that the red operator was previously closer to the blue one but moved recently to the left. A close-up of a Step Map trace in the room can be seen at the bottom picture.

We recruited 8 participants to take part in a usability study, that followed a scenario inspired by the crisis incidents that motivated our designs. After being presented to the techniques, participants were introduced to a fictitious scenario, where they arrived late in the crisis and had to find out, using any technique, on which area of the map they should focus on (area of high activity by other operators). This tasks was put in place to simulate situations where colleagues explicitly request information about others' activities. They were then asked to sit on their desk, access specific data related to their role (traffic controllers) and share them with a first responder dispatcher, in order to coordinate together to which hospital to evacuate casualties and what route to follow. An experimenter impersonated the first responder. Finally, they were asked to conduct a personal task (path tracing), but were interrupted and asked if they were focusing on the same area as other operators in the room. This task was designed to test if our techniques help users be peripherally aware of the current focus of others.

Overall, participants made use of different (often unexpected) combinations of techniques for finding and verifying their answers, indicating that the exact technique to use may be a matter of preference. We observed two strategies adopted by participants. The first strategy was to use only one technique to find the answer to the task, for example used awareness bars to identify areas where other operators have focused on. The second was to one technique to find the answer (awareness bars) and another to confirm their answer (confirm areas where others are working with focus maps). Such strategies are known in decision-making as satisficing and maximizing [Sim78]. The first one consists in choosing a solution that is good enough, the second in making sure that the chosen solution is the one with the highest expected utility. We also observed that participants tend to use the display closest to them to interact. It is thus important to support flexible sharing mechanisms from all displays in the environment. It is worth noting that when discussion was necessary to make a decision (coordination task), all the participants asked their colleague (experimenter) to go in front of the wall-display.

2.4 Other work on collaborative displays [ICB+09, CBM+09, BBT+19]

While wall-displays can comfortably accommodate several analysts working at different distances, they are fairly specialized and expensive equipment. Alternatives include tabletops that are generally smaller in size, but more affordable; or projectors that are large and can be viewed by multiple people, but do not have the resolution of wall-displays and thus do not support close viewing. Although not discussed in detail, in my past work I have also explored these alternatives.

When it comes to vertical displays (such as projectors) of lower resolution, with my colleagues we explored different factors and characteristics that make existing single-user visualization applications good candidates to be used by multiple users in a shared display [ICB+09]. These factors include visualizations with minimal global changes when users interact with it (to avoid disturbing others) and free workspace organization (so that analysts can share and organize their work). Given the lower resolution of the displays considered here, they are not meant to be seen up-close. Thus we consider input with mice and keyboards from a seated position. After retrofitting an existing application for analyzing social networks [HFM07], we performed a user study observing users conducting social-network analysis tasks on the large shared display. Based on this work we proposed guidelines for supporting visualization analysis in shared interactive setups, including considerations for awareness, and the need for per-user undo actions.

When it comes to tabletops, we studied differences in collaboration when colleagues have different access mechanisms available for sharing their content [CBM+09]. We compared the use a regular filesystem access mechanism (e.g., Windows Explorer) for participants to load and share their content with each other, with that of an associative retrieval filesystem [CAK07]. The associative one retrieves all files related to a focus file, across all computers irrespective of the owner of the file. We found that colleagues are less open to sharing content when it is presented in traditional filesystem access methods. Whereas associative file systems promote sharing and make more efficient use of the workspace available on the tabletop.

More recently, we also looked at collaboration of teams of data-workers exploring together visualizations shown on a shared display, that is smaller in size than walls [BBT+19]. We are motivated by experts in different domains, that rely increasingly on simulation models of complex processes to reach insights, make decisions, and plan future projects. These models are often used to study possible trade-offs, as experts try to optimize multiple conflicting objectives in a single investigation. Understanding all the model intricacies, however, is challenging for a single domain expert. We thus propose a simple approach to support multiple experts when exploring complex model results. First, we reduce the model exploration space to only consider the Pareto front of the simulation results (i.e., the non-dominated simulations that are possible solutions to the optimization problem). Then we present to the experts the simulation dimensions (input parameters and output objectives) on a shared interactive surface, in the form of a scatterplot matrix and linked views. Alignment of understanding across different expertise is a process that is less common in sensemaking tasks, and has identified before in model exploration [CRMH12]. However, our work considers more complex models, multiple computational stages and co-located expertise. To understand how multiple experts analyze trade-offs using our approach, we carried out an observational study with real analysts, focusing on the link between expertise and insight generation during the analysis process. Our results reveal the different exploration strategies and multi-storyline approaches that domain experts adopt during their trade-off analysis. We found similar processes to those described in general sensemaking literature [KPRP07, PC05]. Our contribution here, however, is in identifying *why* these processes tend to occur, for example storytelling approaches are used to recap findings in the group. And *when* they occur, for example in large groups storytelling was periodically done throughout the exploration to achieve common ground, but for smaller groups it was done in the end of the exploration as a means to plan future actions.

2.5 Conclusions and Reflections

This chapter focuses on the use of wall-display environments for visual analysis in an effort to increase the amount of information rendered (from the machine side) and allow multiple analysts to come together (increasing human computation power). I describe different challenges, but also opportunities, that arise when designing interactive visualizations in such environments.

I start out this chapter (section 2.1) by introducing two approaches for providing interactions in situations where viewers are mobile in front of the display. Both use interfaces on mobile devices that viewers can hold, but were designed and validated using very different methods.

• Smarties [CBF14] provides flexibility to visualization designers. It is a framework for easily developing a mobile interface for wall-applications, allowing designers to customize the touch surface and widgets of synchronized mobile devices. As a toolkit, the utility of Smarties was demonstrated with example use-cases and applications.

Nevertheless, its development involved considerable iteration. Smarties was motivated by our own frustration when developing wall-display applications and the lack of input support beyond mice and keyboards. The initial versions of the framework included complex cursor behavior (such as copying, linking or mirroring cursors/pucks), fewer widgets, and did not have a touchpad. By using the initial versions of the framework on our own, we realized that complex cursor behavior was challenging to code on the application side. Whereas touchpad interactions and a variety of widgets better reflected the interfaces of actual desktop visual analysis applications, and could make easier the transfer of existing application controls to the mobile interface. The details of the iterative design process we followed, of all the alternatives we considered before reaching our final choices, as we as our inspiration behind this work, now only exist in our personal notes and versions of the system that are not well documented for sharing. I cannot help but feel that others could have benefitted from this experience, to avoid making similar mistakes, or even to inspire them to revisit our failed designs.

• SketchSliders [TBJ15], provides flexibility to analysts, that can sketch on the fly the interactive components they need during their exploration, creating personalized and customized interfaces. The methodology behind the design of SketchSliders was different. We started by understanding under which situations analysts would be interested in sketching their interface. We used a Wizard of Oz setup, where participants would sketch the interaction they desired and an experimenter would apply it to data. This helped us collect a concrete set of designs, that were then added in our prototype. Given the creative nature of sketching, controlled experiments are not appropriate in the validation of our prototypes. Instead, we conducted open-ended analysis sessions with visualization experts, in order to observe how they use SketchSliders in practice.

While our methodology and the motivation for using sketching (allowing analysts to customize their interface) is well documented in the relevant publication, the inspiration behind this work is not. The idea behind sketching widgets originated when I was working with mobile touch phones, and wanted to allow users to customize and parametrize common sequence of actions (calling specific people, looking at particular calendars, revisiting a webpage ...). At that point I was considering sketching simply as a way to visually customize the shortcuts. But as I started investigating how to generalize this approach to create shortcuts to arbitrary widgets such as sliders, I realized sketching could help parametrize the interaction (smaller sliders for coarse control, larger ones for detailed interaction). This led me to consider sketching interfaces as a possible solution in situations where (i) too many controllers are available but only few needed, and (ii) where we need to access information in different granularity. This resonated with my experience using visualization tools for multi-dimensional datasets, that are often weighted down with controllers. In particular in wall-display settings when users are mobile, the space we have available for interfaces can be ironically small (the space of a mobile device, or else we need to consider more complex gesture or voice interactions). The actual design of the final prototype was of course refined by our workshop with the visualization experts. I find it nonetheless interesting that work started on small scale devices (mobile phones) inspired work to address issues in very large wall-displays.

In the second part of the chapter section 2.2, I focus on how information is viewed and perceived in wall-display environments. Again, the work is designed and validated using different approaches.

• In Magnitude Tasks **[BI12]**, we examine three visual variables (length, area, angle) that are considered building blocks of complex visualizations. And show that our perception of them changes depending on our location in front of the wall-display. This raises questions regarding how we should encode and view visualizations on wall displays, and once again stresses the importance of physical movement, as it can help correct this distortion. The methodology followed in our work is similar to other perception experiments, starting from the seminal paper by Cleveland and McGill [Cle85] and other work in psychophysics [Wag06], that calls for a controlled experiment, with well counterbalanced conditions of varying difficulty, and repeated trials per condition.

The inspiration behind this work is also unsurprising. It was driven by our curiosity of the limits of the new visualization platform (wall-display) and our own observations that when interacting up close it was difficult to see remote content. Previous work of a similar nature conducted in other novel platforms (such as tabletops [WSFB07]) also served as inspiration. One aspect that we report on our paper, but do not have the space to elaborate on, was the fact that we run two separate studies. In the first, that we consider as a pilot, we adopted the exact same approach as previous work on tabletops, asking participants to compare two different objects and express one as the percentage of the other (magnitude estimation). Nevertheless, we found that participants tended to round their results to the closest 10%, likely because they had trouble making more accurate estimations due to the distortion. We thus changed our estimation approach, to magnitude production (asking the viewer to replicate the size of the remote object). This provided us with more nuanced results. But also raises the question if previous work that uses magnitude estimation may have also suffered from rounding approximations.

• In Hybrid Image Visualizations **[IDW+13]**, we demonstrate how to take advantage of viewing differences depending on distance. We show how to combine two visualizations that are filtered using high-pass or low-pass filters, so that one becomes visible when seen from afar and the other when seen up-close, thus increasing the amount of information that can be rendered on the wall-display. Our main contribution here was the explanation of the theory behind the approach, and the tools to create such visualizations. As such (similar to Smarties) our methodology consisted of providing many diverse examples, exploring the limits of the approach, rather than running a user study.

The inspiration behind this work was a popular post that had the famous hybrid image by Oliva combining the faces of Einstein and Monroe, and reported in our published work. While our related work section covers examples of our taxonomy for grouping visualizations that work on multiple scales, it does not reflect our process for creating it. We individually collected examples from art, nature pictures and visualizations, that were never shared outside the project, even though they were analyzed and grouped by us (a process not reported in the paper due to lack of space). It also does not include all informal pilot tests we conducted to discover the limits of visual aggregation based on distance (using different basic visual encodings such as color and size). Moreover, aspects of how we immersed ourselves in the work are hard to communicate but could inspire other researchers. For example for months several of us had large printouts of images of what we considered multi-scale visualizations pinned in our doors and office walls, that we would often look at while moving in our space. I personally felt this constant presence of examples in spaces where we moved freely helped us distill the properties that multi-scale navigations that are seen at different distances. While it is not always possible to replicate crucial properties of research questions we are working on (in our case seeing visualizations at different distances), I feel there is potential in this notion of instrumenting one's environment with research inspiration.

In the third and final part of the chapter section 2.3, I describe my work on coordination and collaboration in wall-display environments, for which again we followed different methodologies.

• We start out by examining if there are quantitative differences in collaboration between a shared wall-displays and a collaborative environment made of coordinated desktops [PBC17b]. In a simple coordination task, we showed that the wall-display was slower but leads to more consistent quality results. We pursued this work because as researchers in the field we are still looking for measurable evidence of the trade-offs in using these expensive setups in collaborative settings (compared to less expensive coordinated desktop views). And more generally whether the benefits of this infrastructure outweighs the cost of constructing and maintaining wall-display environments (in all institutions I have worked in, such environments are only used when there is dedicated engineering support put in place). This is still an ongoing investigation, although we do have increasing evidence of such benefits. Given the quantitative nature of our question, our methodology consists of a controlled experiment, where trials are of equivalent difficulty (as we need repeated measures to test learning), but not clearly identified by our participants (to avoid copying their previous solution).

• We next consider the impact of selection technique on coordination in collaborative situations, focusing on more realistic graphs [PBC17a]. We found that with basic selection, a technique with small visual footprint, participants tended to divide a non-divisible task, resulting in lower accuracy. While with a propagation-based selection, that has a larger visual footprint, participants were noticeably more accurate. In this work we use a mixed methodology, starting with a controlled experiment to compare the two techniques, under a specific task (shortest path). As propagation is a novel technique in collaborative contexts, we wanted to see how participants would appropriate it in other analysis scenarios. Thus we next conducted a more open-ended study where they conducted other topology tasks without any training.

• Finally, we consider collaboration in a specific context of use, command-and-control centers. Given the focus on a specific user group and their needs, this work followed a slightly different user-centered design methodology. Instead of starting with a specific solution in mind, we conducted observations and interviews, looking for ways to intervene and improve collaboration. This led to a side question (designing and testing different visualizations for traffic forecasting [PBC16a]) and a prototype with awareness techniques to help operators coordinate and transition from individual to group work [PBC18]. While the user needs were identified through our sessions with the expert operators, the inspiration behind the individual designs themselves does not come from our user sessions, but rather by our knowledge of HCI research in other contexts (e.g., our floor step map awareness technique was inspired by floor interaction work in the UIST conference [AKM+10]).

Overall, the work presented in this chapter has strengthened my belief that wall-displays can increase the communication bandwidth between humans and their data. When it comes to human computation power, they can bring together colleagues with diverse expertise [BBT+19, PBC18], and even in cases where expertise is not in question, they can lead to high coordination compared to other collaborative setups (such as coordinated desktops) [PBC17b]. When it comes to communicating for information from the technical side, their high pixel density can clearly accommodate more information than traditional monitors. Beyond that, they also allow viewing of information at different granularities based on viewing distance, and can even combine two different visualizations that are each seen at specific distances [IDW+13].

3 | Appropriate representations: applied visualizations & fundamental understanding

Using visual encodings to amplify cognition [CMS99] is at the center of visualization research. To help potential users process and act upon large quantities of data, visualization designers aim to provide viewers with appropriate interactive visualizations. But what constitutes an appropriate visualization? And how do we design it?

There are multiple ways a visualization designer or researcher can go about deciding what is an appropriate visualization. They can *start from the user*, understand first user data and needs (what the visualization should communicate). Then create new, or adapt existing, interactive visualizations to support them, reflecting on the process and utility of the solution. This approach is common in design-studies in our field [SMM12], and can impact the end-user domains (from medicine [TLS+14] to urban data [FPV+13]). This approach has driven much of my own work covered in this chapter, focusing on needs of business intelligence analysts [**EB11**, **EB12**, **EAB13**], genealogists [**BDF+10**] and neuroscientists [**GTPB19**].

Another is to *start from the visualization*. With an existing visualization or system in mind, researchers study what are the data, tasks, and even domains that can benefit from it, and what are its limits. This is often an approach used by researchers that have build complex interactive systems that combine visual and computational aspects. Classic examples include Jigsaw [SGL08] that was built around VAST challenges for investigative and intelligence analysis of text documents, and was later considered for more use-cases of sensemaking tasks (e.g., organizing literature, understanding business transaction data, etc. [KS12]). Or more recently Zenvisage [SKL+16] that provides a means to query for visual patterns and was considered for several application domains and scenarios. Our own EvoGraphDice tool [**CBL12a**], falls under this category. It combines a scatterplot matrix visualization with evolutionary computation to suggest interesting data views. The tool was originally built for viewing simulation input and output parameters, but has since been tested with different domain experts that have multidimensional data [**BTBL13**] (biologists, surgeons, and analysts looking at energy consumption or baking processes).

Alternatively, researchers can *start from the data*, exploring what are appropriate visualizations to use for specific data types. Examples include investigations on multidimensional data [KK96a, FL03], temporal data [AMST11], graph data [VLKS+11], text data [KK15], etc. This division is backed-up by the traditional way we teach information visualization in our institutions. Surveys such as the ones mentioned here are extremely valuable, and reflect a collection of work that focuses on more specific aspects of the data, for example how to best visualize dynamic graphs [BPF14], directional graphs [HIvWF11], small world graphs [vHvW04, ACJM03], etc. Our work on geotemporal data fits this category [**PPB20**, **PBP20**].

Very relevant, is the approach to *start from the task*, designing and comparing visualization systems or representations, in order to identify ones that best support specific tasks. Tasks can be low level, such as comparisons of basic visual elements [CM84]; to common but more high-level tasks, such as identifying correlations [Ren17] or similarity [PKF+16]; and even to more involved and complex tasks such as decision making [DBD18], visualization authoring [SLR+20], or communicating uncertainty in the data [HQC+19]. Questions about best designs for specific tasks are naturally combined with the data that is under consideration, e.g., what are the best visualizations for detecting similarity in timeseries [GTPB19] or in scatterplots [PKF+16]. It is thus not surprising that our community has long worked to create task taxonomies that range from general visual analysis tasks applicable across data types [AS04, AES05, BM13], to ones delving into specific data like graphs [LPP+06] or time oriented data [AMST11]. This chapter gives examples of my own work that consider appropriate representations for given tasks.

Researchers can also *start from fundamental questions*. These often relate to our attempt to understand what influences our perception and understanding of data generally. For example, does the size of visual marks affect how we perceive color [Sto12]? Or do cognitive biases affect visualizations [**DFP+20**]? Realistically, we go about answering these questions by narrowing down our investigations to concrete systems or visual representations and tasks (test how color perception is affected during comparison tasks when varying the size of different visual marks [Sza18], or test one specific bias in choice-making tasks using scatterplots [**DBD17a**]). The fundamental questions can also target very specific use-cases or visual representations, for example our own work on how adding contours affect multi-dimensional glyph similarity perception [**FIB+14**]. These fundamental questions may not lead to visualization designs or systems per-se, but rather to design guidelines that can inform how we create or adapt visualizations for specific contexts of use. Our work in decision making falls under this category [**DBD17a**, **DBBF19**, **DBD17b**, **DBD18**].

Finally, research can *start from the technology*. For example, consider what are appropriate ways to visualize and interact with data when dealing with new technological paradigms, such as particular querying mechanisms (e.g., approximate query processing [MFDW17]), new computation approaches (e.g., progressive computation [ZGC+17]), or understanding ML processes [SSSEA19]. Or investigate the constraints and limits imposed by the technology and how that can affect visualization and vice-versa (e.g., dealing with latency in visualization [BCN+20] or how to visualize missing data [SS19]). Another example is much of my work from the previous chapter on wall-displays, that describes our investigations of appropriate visual representations and systems for new hardware technology, as well as our more recent work on visualizations on smartwatches [**BBB+19**] mentioned briefly in this chapter.

Of course these approaches for approaching the question of how to create appropriate representations do not function as silos, but are interrelated. As mentioned, approaching the problem from the point of view of tasks is often combined with the data at hand. And even though we may start our investigation from user needs, we immediately need to consider the types of data they have at their disposal and the tasks they want to achieve. Or when constructing visual analytics systems that are applied to multiple domains, these systems may have been originally driven by the computational technology behind them.

The above list represents for me the ways we tend to often frame our research questions. Depending on how the problem is approached, appropriate validation methodologies may differ, ranging from providing use-case scenarios and informal user feedback, to empirical studies [LBI+12] that can go from user-centered design approaches to controlled laboratory experiments, or a mix of multiple methods.

This list is not necessarily complete. For example, work on visualization toolkits and making construction broadly accessible, starts with specific users in mind like visualization designers [BOH11, Fek04], or lay people such as our recent work on using images to create visualizations [**ZSBC20**]). But the design study methodology approach [SMM12] mentioned above does not apply any more, given the broad and diverse audience. Rather, more nuanced validation methodologies need to be taken into account, that consider creativity, usability and limits of our tools [RLBR18, SLR+20]. Moreover, fundamental questions can focus on broader, high-level topics. For example reflect on the evolution of visualization research topics [IIS+17], understanding how visualization are seen by members of society [PAEE19], ethical implications and responsibility of visualization designers [Cor19], and visualization literacy and teaching [ARC+17, BRBF14]. Here again we cannot necessarily apply traditional empirical methodologies.

Nevertheless, this categorization provides a means for me to organize the majority of my work, that has as a goal to create appropriate interactive visualizations that increase the communication bandwidth between humans and machines. The remaining chapter will give examples of my past work that differs in how with my colleagues we start investigating the question of what consists an appropriate visualization.

3.1 Start with the user - Business intelligence analysts and other experts

The value of a new interactive visualization or visual analysis system, often comes from the utility it provides to people who would use it in practice. As visualization designers we are often confronted with real users and their data challenges, and need to work with them to reach design solutions [SMM12]. I have had the fortune to work with several domain experts, most notably with business intelligence analysts, work conducted within the PhD of M. Elias that I co-supervised with M.-A. Aufaure.

3.1.1 *Business Intelligence analysts and their needs [EB11, EB12, EAB13]*

Our collaboration with business analysts starts in 2011, when SAP Business Objects[1] co-financed M.Elias' PhD. The fact that M.Elias spent a substantial part of her thesis in the company led to very close collaboration with business analysts and tool designers. Business Intelligence (BI) analysts collect large business datasets, that relate to processes, sales, malfunctions, client requests, and other types of business data. They analyze, organize and present their results to decision makers, that are often clients from external organizations). The value of visual presentation in the domain is key, and it usually takes the form of visualization dashboards [Few06], that combine collections of multiple visual components, such as charts, on a single view so that information can be monitored at a glance. The purpose of the dashboard can be strategic (provide quick overviews of the health of an organization), analytic (to provide a detail overview of past events), or operational (to monitor real-time data).

Dashboard Creation [EB11]. At the time our collaboration started, the creation of a BI dashboard involves multiple actors, including end-users (consumers of analysis reports based on these dashboards) and business analysts (creators of the dashboards). It is often the case that end-users intervene and provide feedback to the business analysts that adapt their customized dashboards to meet user needs. This feedback comes at different stages of the dashboard design and setup, and involves a large amount of communication between business analysts and end-users, in order to define functional specifications and a positive user experience. It is interesting to note that this distinction between dashboard creators and consumers often still exists, as evidenced by a recent study on dashboards [SCB+19], even though there are now tools such as Tableau[2] that make dashboard creation accessible to broader audiences.

In an effort to make dashboard creation accessible to end-consumers (rather than experienced analysts), we investigated how novice users construct and customize BI dashboards and how these practices differ from BI expert analysts. For this purpose, we developed in collaboration with SAP Business Objects a new platform called Exploration Views (EV). The details of this work can be found in [EB11]. The design of the EV system was based on a set of principles to allow novice visualization users to easily build and customize BI information dashboards, but also provides functionality needed by experts. These include:

• *Easy creation*. As identified in previous work [GTS10], visualization novices often have partial mental specifications for their visualization needs and tend to refine and change their designs. To ensure a user-friendly dashboard creation, the sequence of steps needs to be simple, with continuous visual feedback. Novices often have no previous knowledge of what visual templates and representations are possible for different data types. In EV this need is met by providing chart suggestions and templates to choose from (also suggested in previous work [GTS10, HvHC+08]) for common data formats.

• *Easy customization*. As novice users create dashboards, they may need to try out alternative visual templates and representations to learn what meets their needs. Thus dashboards should support iterative visualization specifications [GTS10] by being easily customizable and adaptable. In EV this is done through simple drag-and-drop operations that allow users to switch charts, select parts of a chart to make new charts, rearrange charts, or adjust the chart properties (representations, dimensions, etc.).

[1] SAP Business Objects `https://www.sap.com/products/bi-platform.html`
[2] Tableau `https://www.tableau.com/`

• *Visual Analysis support*. Finally, any dashboard is a general purpose visual analysis tool, and as such it needs to supports functions such as saving and sharing, text search mechanisms, fully linked visualizations and visual queries, and other data exploration mechanisms such as filtering.
EV was designed around these goals, in the hopes that both low-cost experimentation with visual templates and visual analytics support, can help novice users to become more accustomed to visualization creation and analysis, promoting learning.

We evaluated EV with both BI experts (7 experts) and novice visualization users (8 participants). We asked them to perform a series of tasks including dashboard creation, chart customization, and dashboard layout, as well as analytic tasks (finding trends, doing comparisons, etc.). Novices used our tool to experiment with different charts and some reported learning about new ways to represent data based on the system recommendations. Novices also tended to make use of undo/redo functionality to experiment with chart customization. This behavior was not observed in experts. Other practices were shared by both groups of users: they started with 2-3 "base charts" that they then copied and/or customized to answer all questions. And they preferred to create one chart per analytic task (starting from the base charts). Based on those and other observations we provide a set of guidelines that augment previous work on designing for visualization novices, in the context of interactive visualization systems in the form of dashboards.

Dashboard Annotation [EB12]. Dashboards are inherently visual analysis tools. As such they need to support information foraging and sensemaking [PC05]. Annotations have been traditionally used to support interpretation (sensemaking) and record insights (e.g., in systems like ManyEyes [VWvH+07]). They can help frame relevant information together, clarifying connections [KPRP07], record further opportunities for investigation [AHW10], aid communication with others [VWvH+07] and aid hand-off between analysts in asynchronous sensemaking tasks [ZGI+18].

While working with BI analysts, we studied their annotation needs. This led as to the first system that provides *context aware annotations* that attach annotations to queries and data-points rather than charts or images. These annotations support new functionality, like "annotate once, see everywhere" for visualizations (not just text [CG09] as was the case up to that point), multi-chart annotations, and annotations that are transparent across hierarchical data dimensions and aggregations.

We interviewed 8 experienced BI analysts (3-11 years), who up to then tended to annotate their dashboards outside their analysis tools, by taking screenshots and embedding them into documents together with textual explanations. They all emphasized the importance of their annotations. As one mentioned *"The data has no meaning without our analysis and contribution we add to it, so charts and dashboards would be useless without explanations"*. Based on our in-depth interviews we extracted a set of design goals for dashboard annotations, that we then applied to a dashboard prototype (details in [**EB12**]). These include:

• *Multiple Targets*. Annotations need to connect and refer to multiple data points and chart targets.

• *Chart Transparency*. Annotations need to be attached to data-points, rather than charts. Few systems up to that point (e.g., Tableau) actually allow data, rather than visualization annotations, and make sure these annotations are visible across different charts.

• *Granularity Transparency*. This is a need that comes from the nature of BI data that is often in the form of OLAP data cubes, that have many hierarchical dimensions (like time/date/year or city/province/country). Experts requested that annotations should be optionally preserved across dimension granularities. We are aware of no other system that supports this.

• *Validity & Lifetime*. As some dashboards are operational (monitor real-time data), it is possible that the values of data points change. Experts requested that annotations should be archived even if the context of the annotation (annotated data) changes, together with a visual snapshot of the data at the time of the annotation. Moreover, users should be able to define the annotation lifetime, based on a time period or data related rules (cross threshold values). We are aware of no system that supports both these aspects.

• *Sharing*. Analysts should have the option to share annotations with specific users groups, or kept them private for personal use.

Figure 3.1: **Context Aware Annotations.** Our dashboard prototype with context aware annotations. (a) The main dashboard contains icons with the number of annotations for each data-point (b). Bellow the list of all dashboard annotations in (c), is a set of recommended annotations (d) emphasizing their similarity.

Based on these findings we implemented a dashboard prototype that supports such annotations. Our annotation system is build on a common data model layer that sits on top of the underlying data sources. Our approach can be used to annotate points, but also the results of more complex queries done on the data layer (e.g., aggregated or filtered data). It can also be used to perform cross-application annotation, and information foraging outside the visualization system (see [EB12] for details). Overall, the system supports annotations that keep a record of their surrounding context (the multiple data points, queries, and charts they are attached to). They also have as special annotation properties, such as their hierarchical visibility, validity, and lifetime. Using annotation context also allows us to provide annotation recommendations [SGL09] for free, as we can identify similar contexts and suggest existing annotations to analysts. A view of our dashboard prototype (that was built upon EV [EB11]) can be seen in Figure 3.1.

After several iterations of the prototype with the original group of experts, we evaluated it with a different group of experts (6 new experts, 1 returning), to explore if they can use it and benefit from the different aspects of context aware annotations for their analysis. All participants found attaching annotations to data points (vs. entire charts) very important for verifying annotation relevance to their task irrespective of which view they used: *"at a glance I can determine the important data points, no matter where I am"*. They also noted the usefulness of recommending annotations for re-use or peer learning (*"learned from notes of others"*). Our system takes snapshots of annotated data points whose context changes, a behavior that was deemed very important: *"annotating important data points acts as a way to see their changes through time"*. In some cases annotations served as learning tools because they acted as a reference between visualizations: *"the radar chart was unfamiliar for me, so I switched it to a scatterplot and I was able to still see the annotated data points. The annotation helped me to understand the radar"*.

Apart from the relevant publication in ACM CHI [EB12], this work led to a patent (US20130124965A1) and was integrated in SAP Business Objects analysis tools.

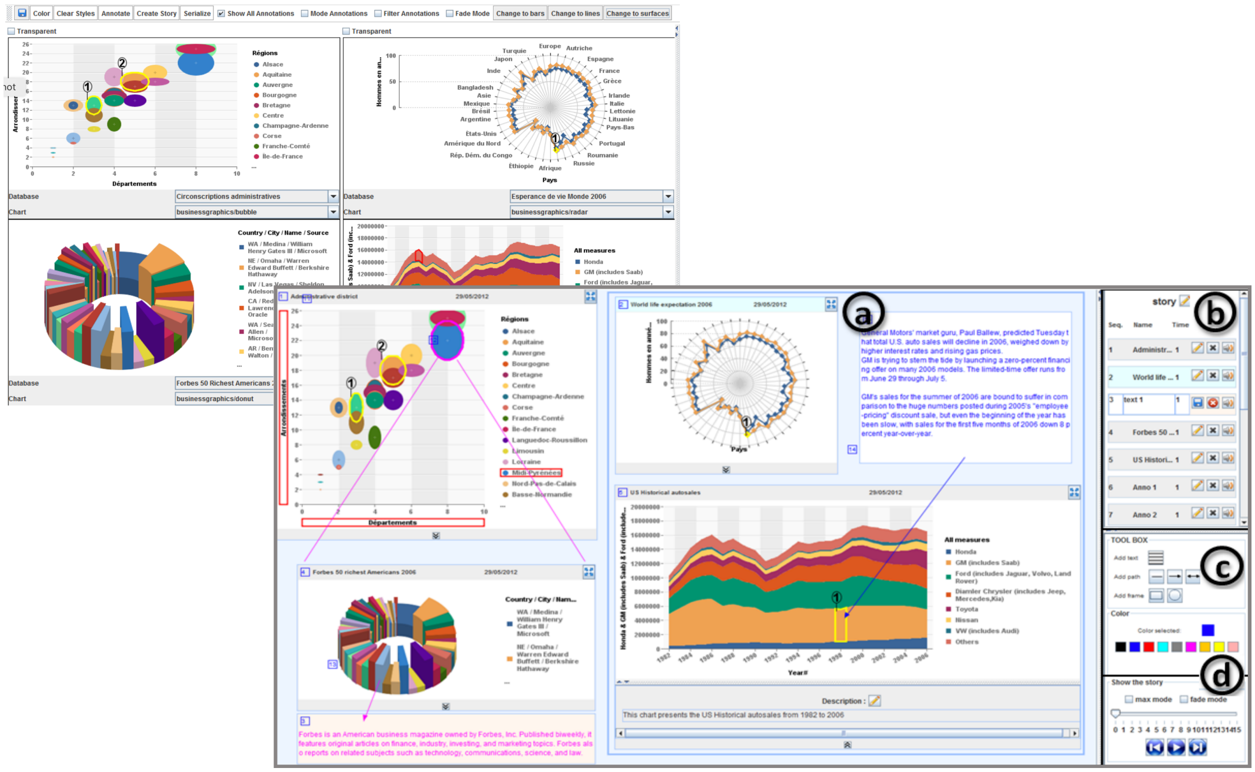


Figure 3.2: **Storytelling in BI.** Top image shows the normal analysis dashboard. Our dashboard can move into narrative mode, together with annotations, seen in the lower image. (a) Narrative board containing all story entities arranged by the story teller: information, relational (arrows), organization (groupings) and emphasis entities, and numbers indicating author reading sequence. This sequence appears in a list (b), where authors can define playback properties and add audio commentary. Below is a pallet of relational and organization entities (c), and the Playback panel (d) to control playback through the story time line.

Dashboard Storytelling [EAB13]. Data visualization has become an important asset both during sense-making analysis, but also when communicating findings to other analysts, decision makers or to a broader public [SvW08]. A story is a powerful abstraction to conceptualize patterns as part of the analytical process, but also a powerful means to present the analysis results [BCB08]. It is thus not surprising that data storytelling has become a major research thread in recent years [RHDC18]. At the time we conducted this work, the community's interest was starting [SH10, MLF+12] and there were few notable examples of systems that supported data storytelling (e.g., [EKHW07, LJ12, MLF+12]). Given that BI analysts often create dashboards that are meant for the consumption of third parties (clients or decision makers) we set out to understand their practices in presenting their analysis to others, and identify gaps and needs.

We start our investigation by interviewing 5 BI analysts (experience from 6 months to 12 years). At the time all experts communicated their analysis or read analysis from others in the form of BI reports. These contain an entire dashboard, often accompanied by several single charts and tables. Details can be additional visualizations, tables, annotations, links to the data used in the visualizations, and finally block text. Experts explained that reports are difficult to understand without detailed explanations from the creator. On the other hand, our experts also occasionally provide the interactive dashboards to their clients in case they want to search for additional details. There is thus a trade-off in curating and presenting the desired story, and allowing story readers to explore the data openly. Based on our in-depth interviews we extracted a set of design goals for dashboard storytelling support (details are in [EAB13]). These include:

• *Fluid transition from Analysis to Storytelling*. Analysis tools used to explore data and create visualizations are different from report creation tools. Our storytelling tool should be able to fluidly transition from their analysis and meta data associated with it (e.g., annotations), to report/story creation.

• *Integration*. To tell their stories, BI creators need tools that combine all materials used currently in their story creation: BI reports, interactive visualizations, ways to indicate story structure, highlighting capabilities, presentation of the story in sequence, and textual or audio explanations.

• *Narrative visual aids*. Report creators need to add focus expressions to draw attention to specific visualization data, such as highlighting, coloring, annotating and zooming. They also require ways to indicate reading sequence (e.g., vectorial references, like arrows).

• *Interactive visualizations*. Visualizations on shared reports are often non-interactive when read outside the organization. A storytelling tool should have completely interactive visualizations, although the way that readers interact with the data should be limited (by default to brushing and linking) and be controlled by the creator. This balance has been identified as a challenging aspect of storytelling [MLF+12].

• *Appropriate BI Story templates*. BI stories have specific structure not necessarily shared by other story narratives identified by Segel et al. [SH10]. Our experts identified templates of interest and highlighted the need for a new template that consists of an *annotated dashboard*.

• *Reuse*. Although BI reports and data changes from analysis to analysis, often the underlying structure of BI stories remains the same. It is thus important to be able to easily reuse the structure of stories created within the tool both for stories of evolving data and similar future stories. We are not aware of other domains where this is crucial and systems that support it.

We informed the design of our system based on these goals and an additional participatory design session we conducted with an expert BI analyst. The requirements and participatory design sessions were the basis for building our storytelling prototype seen in Figure 3.2, that was combined with our analysis and annotation dashboard prototype **[EB12]**, to allow easy transition from analysis to story creation and sharing. The storytelling view keeps the annotations and charts from the original dashboard, but also allows the creation of chart groupings and rearrangement, the definition of reading sequence for the charts, the addition of free text, of entities for emphasis, as well as relational indicators (arrows). Designers can also determine the level of interactivity for the different charts.

To validate our prototype we conducted two user feedback sessions to assess the usability and effectiveness of the system, *both from the creator's and the reader's perspective*.

In the first session two BI experts (that had taken part in the interviews) used the tool to create a story based on an existing BI report. Both experts were very enthusiastic with the prospect of having access to storytelling functionality within their analysis tool. They reaffirmed that story reading must be guided by the story creator, else the goal of the story may be lost. Both experts suggested our system should support two types of BI narrative stories: (i) *Fixed stories*, that present snapshots of datasets at specific points in time, yet are interactive (e.g., for filtering); and (ii) *Online stories*, that present dynamically evolving data, and can have the same analytic scenario regardless of data values. Thus stories may be repeated: they can have the same chart descriptions (what type of data is shown), and the same reading sequence, but different data values. Here visualizations in stories are no longer snapshots, but are updated with data changes. We have implemented this extension.

We then ran a second session to evaluate the prototype from the reader's perspective, and thus close the story communication cycle. We conducted 40 to 50 minute sessions, with 5 BI novices. Participants were asked to read a classic BI report created by one of our experts, and to also read a BI story created by an expert in the previous session that gives the same information as the report in the form of a story. All participants found that reading a story was easier as *"it showed the facts in an understandable manner"*. They felt more confident that they had understood the story (compared to lower confidence with the report) and were able to better remember the gist of the analysis message and results.

Our work **[EAB13]** received the **Brian Shackel Award** (best paper award) in INTERACT 2013.

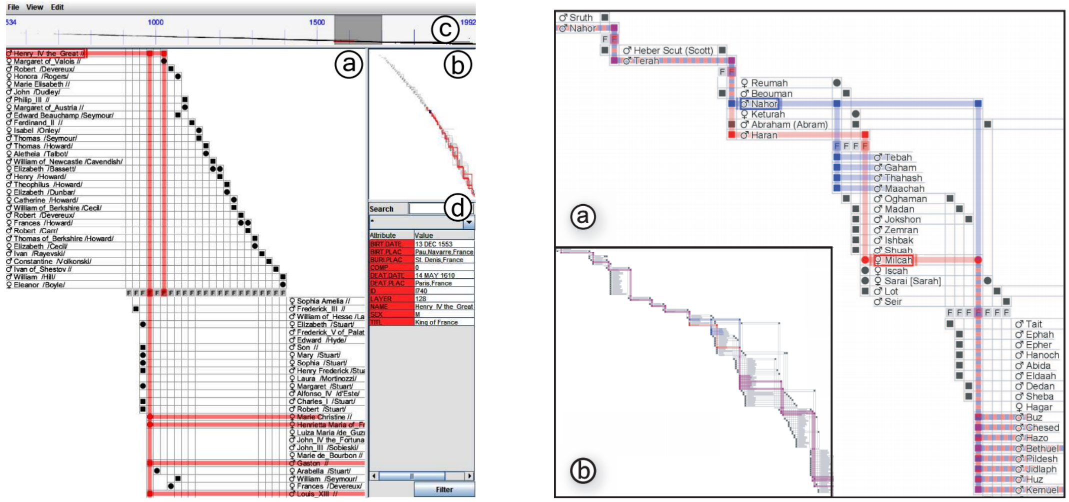


Figure 3.3: **GeneaQuilts.** Left image shows the GeneaQuilts System with part of the European Royal families. It consists of (a) the main visualization, (b) an overview, (c) a timeline and (d) a query and details panel. The Right image shows multiple selections in the Bible. (a) Selecting Milcah's bloodline in red and her husband's in blue reveals their common descendants, but also a close common ancestor (Terah) shown by the two lines blended in a dashed pattern. (b) The blending of these bloodlines in the overview.

3.1.2 *Work with other experts [BDF+10, GTPB19]*

Beyond dashboard design, I've had the pleasure to work closely with genealogists and neuroscientists. This section presents a summary of this work.

With Genealogists [BDF+10]. The study of family relationships, is a popular activity pursued by millions of people, ranging from hobbyists to professional researchers [Mil03]. They mostly use visualizations are based on node-link diagrams, which have been shown to quickly become unreadable as graph size grows [GFC05]. Considering that genealogical databases built by individuals can reach thousands of nodes, and those built by organizations tens of thousands, there is a need for a more scalable visualization solution. To this end, we introduce GeneaQuilts, a new visualization for representing large genealogies of up to several thousand individuals. The visualization takes the form of a diagonally-filled matrix (generalizing the Quilt visualization [WBS+08]), where rows are individuals and columns are nuclear families.

We conducted three extensive interviews with 8 users involved in genealogy research (3 historians investigating transmission of land and title ownership across families in France, 4 anthropologists interested in inter-marriage strategies within small tribes worldwide, and a semi-professional genealogist who investigates family ancestry of clients). Based on these, we identified unique needs in this domain, such as the need to find all paths linking two individuals and examine if they are consanguine or conjugal, intermarriages between multiple families, marriages across generations, etc.. These needs led to the refinement of GeneaQuilts and in the design of a visual analytics system around it, to support interactive genealogy exploration. The system Figure 3.3 includes an overview, a timeline, search and filtering components, degree-of-interest views of entities, and a new interaction technique called Bring & Slide that allows fluid navigation in very large genealogies. Four of our original experts came back to try our system with their data and found it to be very clean (as they were used to very complex node-link diagrams) and were able to spot new insights in their data (e.g., intermarriages they had not seen previously).

GeneaQuilts, presented in IEEE VIS/InfoVis 2010 [BDF+10], has since been downloaded hundreds of times, appeared in popular media and has been incorporated in several genealogy software platforms[3].

[3] GeneaQuilts is incorporated in: Puck `www.kintip.net`, Progeny Genealogy (referred to as Trellis Charts) `progenygenealogy.com`, in Généapro `github.com/briot/geneapro/wiki`, and in The Gramps Project `gramps-project.org`

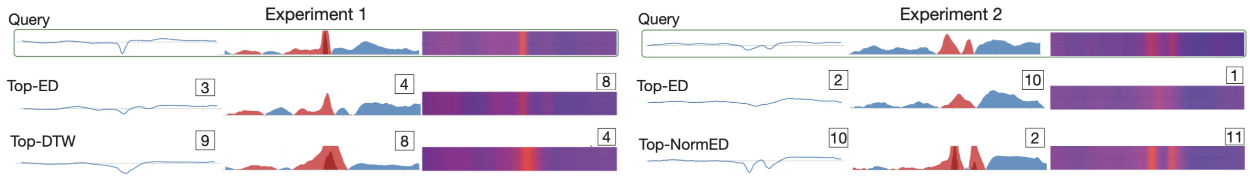


Figure 3.4: **Timeseries Similarity.** Two queries for which different visualizations resulted in different choices. Boxes show the number of participants (out of 12) who chose the specific answer. Left: Experiment considering temporal position variations. This example shows that colorfields can be more sensitive than line charts and horizon graphs to stretching deformations along the time axis (DTW). Right: Experiment considering amplitude. A strong preference for normalized measures (Euclidean Distance ED) in line charts and colorfields and a strong preference for non-normalized measures under horizon graphs. Overall, horizon graphs exaggerate flat signals and are more sensitive to deformations along the y-axis.

With Neuroscientists [GTPB19]. Timeseries are temporal sequences of data points that are of interest in different domains, including electroencephalography (EEG) signals. These are the focus of our experts from the ICM Brain and Spine Institute, that we met in two separate 1h sessions (3 and 2 experts respectively). They explained that they are looking for "epileptiform discharges", abnormal patterns that have been linked to various cognitive disruptions and reoccurrences of epileptic seizures [SFLGM15]. As opposed to epileptic seizures that produce large disturbances in the EEG signal of a patient, epileptiform discharges are especially hard to detect. According to our experts, data-mining algorithms for detecting these patterns [IES+08] result in many false positives and are not useful in practice. This is because epileptiform discharges often resemble normal background activity due to regular artifacts (pulses of the heart, the eyes, or the muscles [JDR+16]), and vary greatly across patients. Thus neuroscientists visually inspect a large set of EEG timeseries (300 sensors, with several thousand data points each).

In an attempt to aid them, we suggested they manually identify a small number of epileptiform discharges and use them as patterns to automatically detect similar subsequences. The experts could then visually verify if these are also potential discharges. To detect these similar sequences we set out to choose an automatic measure to compute timeseries similarity, such as Euclidean Distance [FRM94] or Dynamic Warping [BC94]. Each measure considers different patterns as similar, often called *invariances*, for example normalized Euclidean distance considers patterns as similar irrespective of their amplitudes, whereas DTW considers as similar patterns stretched along the time dimension. When we requested information about what types of variations or deformations in the patterns could indicate similar signals, our experts explained that some of their decisions remain subjective, and past work has shown that agreement between different experts can be low [JDR+16]. This raised an interesting question for us as visualization designers. Do visualizations actually help viewers understand what temporal patterns are similar, or are there aspects of the invariances of interest that are not communicated well in some visualizations?

We set out to investigate if different types of visualizations communicate or de-emphasize invariances in a similar way, or if visualizations need to be chosen appropriately to help experts reach consensus. The visualization literature has examined similarity perception and its relation to automatic similarity measures for line charts [EZ15, MA18, CG16], but has not yet considered alternative visual representations of timelines that are space-efficient, such as horizon graphs [SMY+05, Rei08] and colorfields [CAFG12, ACG14, NC07, SMY+05]. Motivated by how neuroscientists evaluate epileptiform patterns, we conducted two experiments (18 participants each) that study how these three visualization techniques affect similarity perception in EEG signals. We seek to understand if the timeseries results returned from automatic similarity measures are perceived in a similar manner, irrespective of the visualization; and if what people perceive as similar with each visualization aligns with different automatic measures and their invariances.

Our findings indicate that horizon graphs align with similarity measures that allow local variations in temporal position or speed (i.e., DTW) more than the two other techniques. On the other hand, horizon graphs do not align with measures that are insensitive to amplitude scaling (i.e., like normalized Euclidean distance), but the inverse seems to be the case for line charts and colorfields. Overall, our work indicates that the choice of visualization affects what temporal patterns we consider as similar, i.e., the notion of similarity in timeseries is not visualization independent.

This work [**GTPB19**] was part of A. Gogolou's thesis (co-supervised with T. Palpanas and T. Tsandilas) and was followed by work on optimizing similarity search performance, discussed in the last chapter.

3.2 Start with the question - Decision Making and Perception Studies

When A. Dimara, co-supervised with P. Dragicevic, started her PhD on decision making using visualizations, the replication crisis had raised questions about research practices [JLP12], and whether confirmation bias affects how we researchers frame our work and make decisions about what is relevant. This drove our interest to study how visualization can help us make better decisions, and more generally how we make decisions using visualizations. This was the main question that drove our research, that we tackled by breaking the question down into easier to manage interrogations. This section mainly focuses on my decision making work, and only briefly mentions my other work that started from specific questions.

*3.2.1 Decision making [**DBD17a**, **DBBF19**, **DFP^{+}20**, **DBD17b**, **DBD18**]*

The Attraction Effect [DBD17a]. In visualization research we tend to consider as effective visualizations that help viewers accurately access information [CMS99, ZBK15]. Decision making research cautions that this may not be the case when it comes to making decisions. It is now well known that full access to information does not necessarily yield good decisions [Kah11]. When dealing with complex decisions, humans often resort to heuristics, *"simple procedures that help find adequate, though often imperfect, answers to difficult questions"* [Kah11]. Heuristic strategies have evolved because they can be very effective [Gig08]. Unfortunately, they can also have imperfections that manifest themselves as cognitive biases [Kah11]. With E. Dimara and P. Dragicevic we set out to understand if cognitive biases, detected in other contexts, can affect people when they attempt to make decisions using visualization.

In an attempt to answer this question, we started looking at candidate biases that may appear when we use visualizations. After going through literature in psychology, economics and sociology, we decided to focus on the attraction effect (also called decoy effect or asymmetric dominance effect). The attraction effect is a cognitive bias that appears when people are faced with three options, two that are *uncomparable* (there is no best option, as they represent trade-offs, such as price/quality) and a third option that is called the *decoy*. The decoy is similar, but slightly inferior, to one of the two uncomparable options (often called *target*). When this third option is introduced, it tends to shift peoples' preference towards to uncomparable option that resembles the decoy. This shift in preference is irrational because it violates a basic axiom of rational choice theory, the principle of *regularity*, according to which the preference for an alternative cannot be increased by adding a new alternative to the choice set [HPP82].

We focus on the attraction effect for several reasons. First, it is one of the most studied cognitive biases in fields such as psychology, consumer research and behavioral economics. And second, these studies generally employ very small sets of alternatives (typically three) and numerical presentation formats, so it is still unknown whether the bias generalizes to data visualizations.

To investigate the existence of the attraction effect in visualizations, we conducted two crowdsourced experiments. In the first between-subjects study (305 participants), we replicated an existing study from psychology that only had three alternatives in tabular format [MHH13], where participants had to choose a gym based on its cleanliness and variety of equipment. But we introduced a new condition where alternatives were instead presented in a scatterplot visualization. Even though past work has not tested visualizations, it has used scatterplots when reporting their results to illustrate the alternatives used in attraction effect experiments [HC95, HPP82, OP95, Sim89]. Scatterplots can also scale to more than three data points, which is important for our next study. Trails of our first study can be seen in Figure 3.5-left.

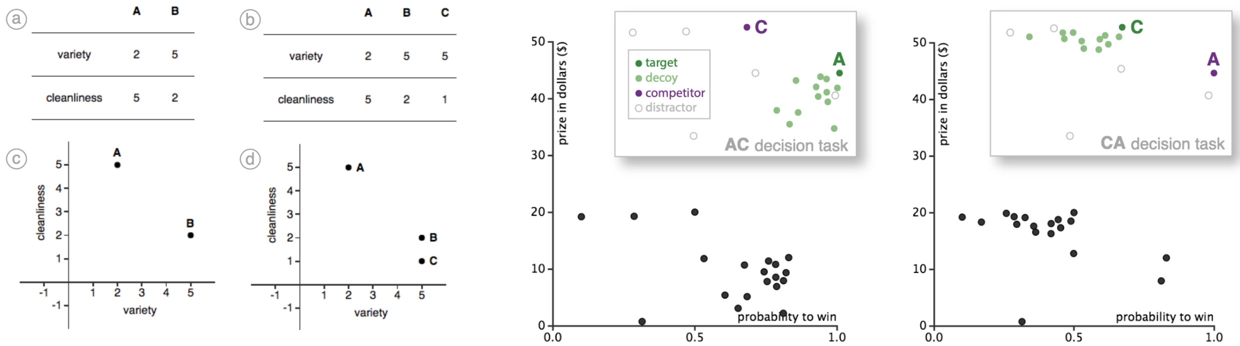


Figure 3.5: **The Attraction Effect experiments**. The left image shows examples of experimental stimuli in the first study, for the table (a,b) and the scatterplot (c,d) conditions. The left decision task (a,c) has no decoy, while the right decision task (b,d) has a decoy on B. The right image shows experimental stimuli from the second study, for the two matched decision tasks AC and CA (black-and-white background images), and explanatory annotations (box overlays).

Results for the tabular condition show a shift in preferences (consistent with the attraction effect) that is nonetheless lower than the original study (replication studies often have smaller effect sizes [O^+15]). When it comes to scatterplots, we also have evidence for an attraction effect, indicating it does exist when using visualizations. This observed shift in preference after adding a third irrelevant option gives credence to the idea that people may make irrational decisions even when they use visualizations as decision making aids. Thus we decided to explore the effect further, using scatterplots with larger sets of alternatives.

In the second study Figure 3.5-right we were inspired by a within-subject study (that generally have higher statistical power) that again dealt with three alternatives [Wed91]. In it, participants had to select a lottery ticket with a probability and a chance to win: two being the uncomparable alternatives and one being the decoy. We adapted the study to add more decoy objects, the majority of which are dominated by (inferior to) one of the two uncomparable alternatives (target). Our results (from 72 participants) showed again a clear shift in preference depending on where the decoy objects were placed, even though these decoys are rationally not relevant to the choice (the best choices are always the uncomparable alternatives). We were thus able to generalize the attraction effect procedure to more than three alternatives, and verify that the effect can persist when participants are presented with more realistic scatterplot visualizations.

Our work on the attraction effect [**DBD17a**] is the first to provide evidence of the existence of a cognitive bias when using visualizations to make decisions, and received an **Honorable mention** (top 4 papers) in IEEE VIS / InfoVis 2016.

Mitigating the Attraction Effect [DBBF19]. While we were trying to generalize the attraction effect to more than three alternatives, we ran a pilot study (published as a technical report [**DBD16**]), where the two uncomparable data points were highlighted. In that study we were unable to detect a decoy effect. This got us thinking that maybe highlighting the two choices is a way to mitigate the bias and a few years later we set out to investigate this further.

Cognitive bias mitigation methods, also known as *debiasing*, are challenging. They tend to focus primarily on educating the decision maker, (e.g., through statistics training [FKN86]), have shown limited or only temporary success [Poh16], and are often ineffective [Fis82, Ark91, Kah03, SSS02]. Another approach is to debias the environment instead of the decision maker [KB93]. Previous studies that use this approach altered the design of textual information [KB93, HP13], but did not consider the use of visualizations.

Given our pilot findings we start out to study if highlighting the Pareto front, i.e., the uncomparable choices, would prevent participants from shifting their choices based on the decoys present in the scatterplot. In a first crowdsource study (207 participants), we replicated our attraction effect study [**DBD17a**], highlighting the uncomparable points (non dominated alternatives). A trial can be seen in Figure 3.6-Left.

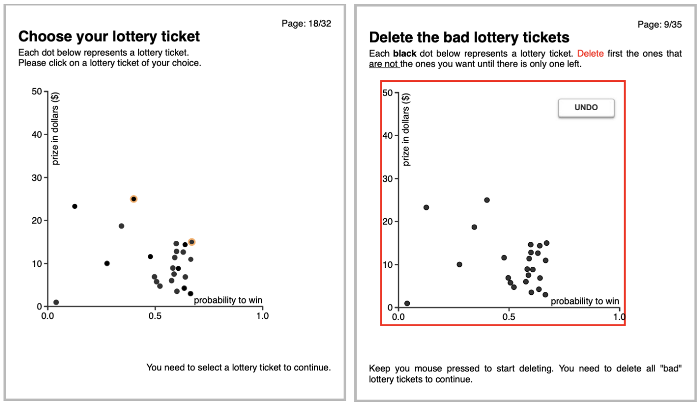


Figure 3.6: On the Left, the stimulus of the PARETO experiment. Non-dominated (uncomparable) datapoints have a colored outline. On the Right, the stimulus of the DELETION experiment. When the mouse is pressed, a red outline indicates that the cursor will delete datapoints. Participants deleted all datapoints but one - their final choice.

We found suggestive evidence that showing the Pareto front indeed weakens the bias, but does not eliminate it (when compared to a condition where the Pareto front is not highlighted). The highlighting also reduces task time. However, we examined only one design and others remain to be tested (e.g., indicate the Pareto front more strongly, with a stronger visual cue such as a line, or fading out the decoy objects).

As we are considering mitigation strategies, we looked at literature on models of how humans chose between items of different attributes, that represent trade-offs. One model includes strategies in which attribute values are considered (*compensatory* strategies). A common example is the "weighted additive" strategy the advocates that the decision maker weights all attributes by importance and chooses the one with the highest weighted sum [PBJ93]. This strategy is often supported in visualization systems [CL04, GLG+13, PSTW+17] by allowing analysts to weight attributes and rank options based on a combined score. Another model (called *non-compensatory*) includes strategies where the decision maker can reject an alternative with a bad value for one attribute, even if it has perfect values for the other attributes [Wri75]. A known non-compensatory strategies is the "satisficing" one [Sim56], where the decision maker chooses the first that satisfies some thresholds. If no such alternative exists, the decision maker relaxes the thresholds and repeats the process, or chooses a random alternative [PBJ93]. Another, less common example in this category is the "elimination by aspects" strategy [Tve72] in which the decision maker rejects all alternatives that do not satisfy a given threshold and repeats until only one alternative is left [Plo93].

We focus on this last strategy of elimination, allowing participants to explicitly delete data points until only their choice is left. A typical attraction effect choice task is divided into two subtasks: the decision maker is expected to first recognize the dominant points by rejecting the decoy(s), and, second to choose between the two trade-off choices [CG15]. We suspected in our previous work **[DBD17a]** that it is the first part (dominance recognition/comparison) that causes the bias. We thus attempt to differentiate these two tasks (dominance recognition and choice) with interaction - ask participants to delete undesirable choices. A trial of this experiment is seen in Figure 3.6-Right. In our second study (203 participants), participants saw two conditions, one where they had to delete points they did not want as their choice, and one where they could click on their choice. We observed a very strong drop in the attraction effect when using deletion, proving it is a very effective debiasing strategy.

Apart from significantly reducing the attraction effect, this work illustrated that visualization and interaction can help reduce cognitive biases in decision making processes.

Task-based Taxonomy of Cognitive Biases [DFP+20]. At the end of her PhD, E. Dimara started organizing the literature for her thesis, thinking of what would be a good way to present it to a visualization audience. This is literature that comes from many domains (economics, psychology, sociology, ...) and the biases mentioned are not always empirically proven. We realized that this research is not really accessible to visualization researchers ("as a visualization designer, when I create a new visualization for a specific purpose, what are the cognitive biases I should worry about?"). Existing bias taxonomies are organized by cognitive theories (reasons behind the biases) and as such are hard to associate with visualization tasks.

We thus set out to construct a more actionable taxonomy for visualization designers and researchers. In our taxonomy we first, aim to help bridge the gap between cognitive psychology and visualization research by providing a broad review of cognitive biases, targeted to information visualization researchers. Second, we define a taxonomy of cognitive biases classified by *user task* where the bias appears, instead of by proposals for psychological explanations for why biases occur. Finally, we provide starting points for finding studies that have investigated the bias, and we highlight open research problems.

After a standard bibliographic search we gathered an initial list of biases, and searched for the most representative paper that *empirically tested* each of the biases in our list. We then categorized the cognitive biases, using a bottom-up grouping method similar to card sorting. And finally, we reviewed each bias from a visualization perspective by 1) searching for existing relevant visualization work (if any) and 2) brainstorming future opportunities for visualization research.

We ended with a classification of biases that have been experimentally observed, grouped in seven categories depending on the tasks they appeared in. These were:

• *ESTIMATION.* This category includes biases that appear when people need to estimate likelihoods. Examples include the Base Rate Fallacy (ignoring the base rate probability of the general population) that visualization researchers have already looked at [MDF12, KBGH15].

• *DECISION*. This category includes biases that appear when people need to conduct any task involving the selection of one over several alternative options. The Attraction Effect that we studied **[DBD17a, DBBF19]** is one such bias.

• *HYPOTHESIS ASSESSMENT*. Biases in this category appear when people are tasked to investigate whether one or more hypotheses are true or false. The Confirmation Bias, where people tend to favor reasoning or information that confirms a preferred hypothesis, is one such example [Mah77].

• *RECALL*. This category includes biases observed when participants were asked to recall or recognize previous material. An example is the Serial-Positioning effect, where people best recall first (primacy) and last (recency) items in a series [MJ62].

• *CAUSAL ATTRIBUTION*. Biases in this category appear when people explain the causes of behavior and events [Kel73]. An example is the Self-serving bias, that suggests that people tend to attribute success to their own abilities and efforts, but ascribe failure to external factors [CS99].

• *OPINION REPORTING*. This category includes biases observed when participants were asked to answer questions regarding their beliefs or opinions on political, moral, or social issues. An example is the Bandwagon effect, where peoples' reported beliefs can change according to the majority opinion [NCG93].

• Finally, the *OTHER* category includes biases that are not necessarily tied to a task.

The taxonomy lays out the problem space, facilitates hypothesis generation, and hopefully will guide future studies that will ultimately help visualization designers anticipate, and possibly alleviate, cognitive biases. The details of the taxonomy creation process, details of the categories, examples of opportunities, and a full list of biases and references for them are available in our TVCG paper **[DFP+20]**.

Are decision making tasks different to analytic tasks? [DBD17b] While working on cognitive biases we started considering more generally the notion of decision making tasks when validating visual analytic systems. When evaluating a system, what determines if the participants conduct a decision or an analysis task is the *narrative* the experimenter provides. For example, if we ask participants to select a house to buy, this simulates a decision. When we ask them to find the cheapest house, this is an analytic task. Given the idea that there is a decision strategy that is "satisficing" [Sim56] (e.g., chose an item that is good enough), we expect that the nature of the participants' answers could change depending on the narrative. It more generally made as consider the question, how does the narrative we provide to our participants when we evaluate our systems affect their results? This question is particularly challenging when conducting studies in crowdsourcing platforms (as was the case in all our cognitive bias research).

Crowdsourcing platforms have become very popular in visualization research as they can access a large and diverse pool of participants, and allow for rapid evaluations [HB10, KCS08, MDF12, BBIF12, BRBF14].

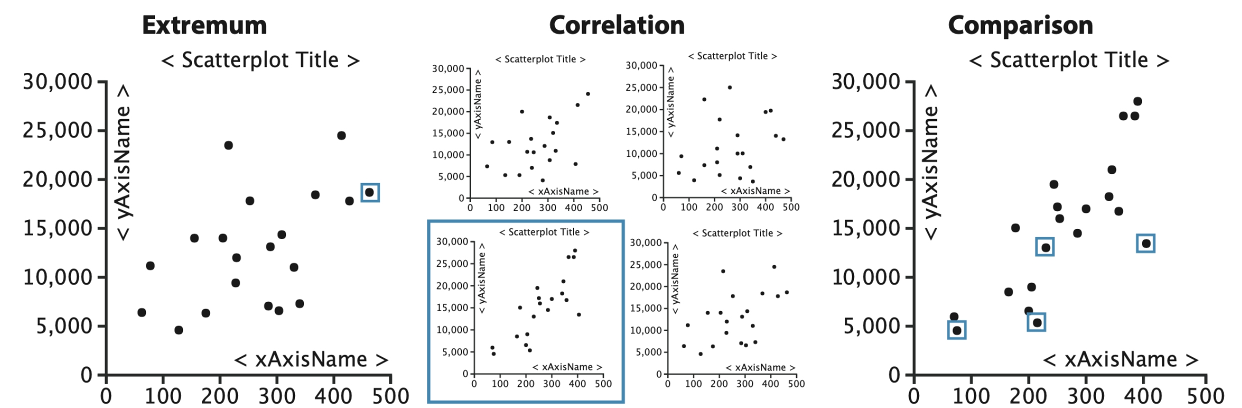


Figure 3.7: **Framing Narratives.** Stimuli used in the Extremum, Correlation and Comparison tasks. Correct answers are shown in blue. Axes were labeled (**X, Y**) for Abstract task instructions, and (**size (m2), price ($)**) in all other conditions. The title was **Diagram Z : Datapoints** in Abstract task instructions, and was **Diagram Z : Houses** in Semantic instructions (all tasks) and in the Decision making narrative. In all other conditions the title was **Agency Z : Houses**. Z was an integer (1, 2, 3, or 4) identifying the scatterplot.

However, engaging crowdworkers and obtaining high-quality responses is challenging [Hul11, ECD14]. In particular, task instructions need to be carefully crafted in such remote studies, where the instructor cannot help or motivate participants. We thus want to understand how task instructions and the narratives we provide to participants affect the quality of responses when evaluating visualization tools. Researchers and practitioners have already use narratives in the context of data analysis and communication, in order to improve data understanding and engagement of the users [HD11, SH14]. Nevertheless, the effects of adding these narratives is unclear, in particular in crowdsourcing settings where incentives vary across people, and attention and motivation are hard to control for. For example, what would be the difference between instructing participants to identify the data point with the minimum value, and between instructing them to imagine they are trying to find the cheapest available house. Both questions are equivalent at the task level and consist of finding an extremum. The second version is possibly more salient and engaging, and with a context that is easy to understand, characteristics linked to good crowdsourcing performance [Hul11, MW10]. At the same time, the first version is more succinct and less demanding in terms of time and patience, aspects that have also been emphasized in crowdsourcing guidelines [ECD14].

To determine if the type of narrative instructions affect performance, we chose three tasks that we presented to participants with different types of narratives. The tasks (seen in Figure 3.7) were:

• *Extremum*, where participants had to find the data point with highest value according to the X dimension.

• *Correlation*, where they had to chose the scatterplot with the highest correlation (among four).

• *Comparison*, where participants had to compare data points across their two dimensions simultaneously. The task consisted of finding a data point without any "competitor". A competitor was defined as a data point that has both larger X and smaller Y (i.e., an uncomparable alternative on the Pareto front). The task had four possible correct answers.

To test the influence of narratives, we studied the effect of providing very different task instructions. The examples seen here are for the Comparison task. The narratives move from:

• *abstract* task instructions that provide no contextual information for the dataset (e.g., “Select a data point that has no competitor”);

• to adding *minimum semantics* to the dataset (e.g., “Select a house that has no competitor”);

and to further adding a backstory narrative that justifies the purpose of the task. For the backstory narratives, we compared two popular types of narratives from the visualization literature:

• *analytic* narratives involving answering investigative questions about data (e.g., ask them to put themselves in the situation of a real estate analyst and to find answers to analytical questions (e.g., "Given what you read, select a house that is a good deal.");

• and *decision* making narratives involving making personal choices based on data (e.g., put themselves in the situation of a house buyer and, given some criteria and constraints, to make choices (e.g., "Given what you read, which house would you buy?").

We ran a between subjects design crowdsourced study (405 participants), where participants each saw the three tasks with one type of narrative. Our findings indicate that adding minimum data semantics can provide subjective benefits (such as confidence, perceived easiness, and enjoyment). However, we found no evidence that it increases accuracy. We even found some evidence that our longer backstory narratives could hurt accuracy. Finally, in the comparison task, that represents a trade-off and thus has elements of real-life decision making, we found that the decision making narrative was less accurate than the analytic narrative. Most likely, the decision making framing caused participants to focus more on subjective preferences and less on giving a correct answer. This is of interest to researchers conducting evaluation of visualizations for decision making. Our findings imply that decision making tasks are more error-prone than equivalent analytic tasks, and that evaluating a decision-support system with standard analytic questions may not reflect a realistic use of the system and may overestimate its performance.

Details of the wordings for the other tasks, as well as experimental hypothesis and detailed results can be found in our ACM CHI 2017 paper [**DBD17b**].

Visualizations for Decision Making Tasks [DBD18]. Our previous work on narratives [**DBD17b**] suggests that decision tasks may be more error-prone than equivalent analytic tasks. Thus when researchers evaluate their tools, finding good performance with elementary analytic tasks does not necessarily guarantee good performance in decision making tasks. Since many decision tasks have no clear ground truth, evaluating visualizations for their ability to support decisions is difficult, and there is a lack of methodological guidance on how to do so. So we set out to explore conceptual and methodological issues in evaluating visualizations for their ability to support decisions. We focused on *multi-attribute choice tasks*, that consists of finding the best among a set of alternatives that have several attributes. One example is buying a house, where each available house is defined by its price and a number of features such as area, number of bedrooms, orientation, etc.. There is no unique way of defining a "good" alternative, and the best definition depends on the context. "Goodness" can be defined in objective terms (e.g., Pareto dominance) or in subjective terms (e.g., personal satisfaction with the choice).

Since in a multi-attribute choice task all alternatives are known in advance, and defined across a set of attributes, all information can be provided as a data table [OL03] where rows are alternatives and columns are attributes. Several visualizations exist to visualize such multidimensional datasets, that we cover in detail in our IEEE VIS/InfoVis paper (see [**DBD18**] for details).

For our study we focus on three common visualization techniques, that fall under the Lossless geometric projection categorization by Keim and Kriegel's taxonomy [KK96b]: *Scatterplot Matrix, Parallel Coordinates* and *Tabular* visualization (Figure 3.8). These are common components of existing visualization tools that have been presented with scenarios related to multi-attribute choice tasks (e.g.,[EDF08, AS94, CL04, GLG+13]). There is also extensive work on evaluating these tools in *analytic tasks*, such as value retrieval [WS92, BC08, CCH+14], range tasks [WS92, ROF12], finding extrema [YMSJ05, BC08, CCH+14], finding outliers [WS92], and identification of patterns [WS92], correlations [YMSJ05], and clusters [YMSJ05]. In other words, a number of evaluations have employed analytic tasks to study these visualizations. Nevertheless, they have not been compared under decision making tasks.

In our paper we expand on how such a comparative study should be constructed, emphasizing the need to: (i) include all features that are considered standard for each visualization (e.g., provide ordering of attributes not just for Tabular or Parallel Coordinates, but also Scatterplot Matrix); (ii) keep the visualizations as comparable as possible through a consistent visual design (e.g., colors, labels, space allocated), a consistent interaction design (e.g., allow for range selection in Tabular visualizations and Scatterplot

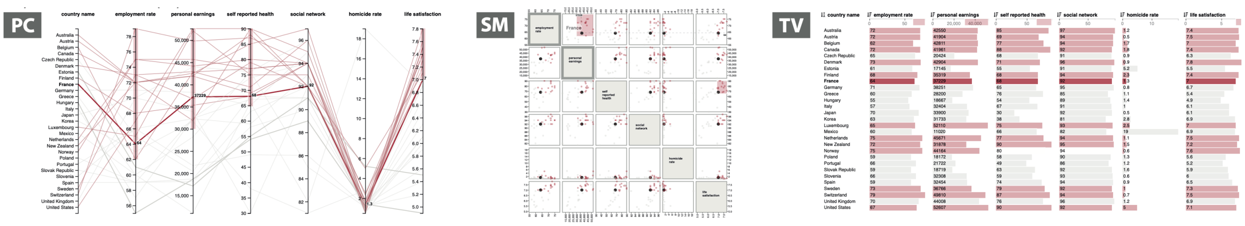


Figure 3.8: **Visualization Techniques for Decision Making.** The visualizations we evaluated: Parallel Coordinates (PC), Scatterplot Matrix (SM) and Tabular Visualization (TV).

Matrix), and by having all interactions present the same amount of information across visualizations. We stress the importance of creation of datasets that are of equivalent difficulty for the trials in each technique. And we introduce a procedure by which participants ranked their a-priori preferences for all attributes, in order to create a decision score that is tailored to each participant. Finally, as another measure of decision quality we look at participants' attachment to their decision (we ask them to imagine that an automatic recommender system could suggest another choice taking into account their preferences, and then inquire whether they would switch to this choice or not).

To verify that our 21 participants were able to understand the visualizations and use them effectively, we first evaluated the visualizations on analytic tasks: a *value retrieval*, a *range task* and a *correlation task*. As these three techniques have not all been compared together for all these analytic tasks, we summarize briefly the results. All three techniques yielded close-to-perfect accuracy. There were however large differences in completion times: Scatterplot Matrix was slowest for value retrieval and range tasks, but by far the fastest in correlation tasks. The lower performance of Scatterplot Matrix in the two low-level analytic tasks can be explained by the difficulty of dealing with two axes concurrently. On the other hand, its efficiency for correlation tasks is not surprising, as scatterplots are known to convey correlation effectively [LMvW10, KH16a]. Furthermore, Scatterplot Matrix shows all pairwise correlations simultaneously, while both Parallel Coordinates and Tabular required manual attribute reordering to examine them in sequence. Though Parallel Coordinates is often considered a good choice for conveying correlations [HYFC14a, HLKW12, LMvW10], it was outperformed by Tabular both on time and accuracy. Overall, tabular visualizations seem to be a compelling choice, despite the low attention they have received in the literature on multidimensional visualization.

The second part of our evaluation involved actual decision-making tasks. We found the techniques to be comparable across metrics, with a slight speed advantage for Tabular. Participants also preferred Tabular over Parallel Coordinates overall. Participants reported being more attached to choices made with Scatterplot Matrix on average. The reasons for this are currently unclear, although one explanation could be that Scatterplot Matrix supports overview tasks (confirmed by our results with the correlation task), which made participants more confident that they did not miss a particularly interesting alternative.

Evaluating visualizations for their ability to support decision making is challenging. The quality of a decision is hard to capture with objective measures, as decisions often involve personal preferences which are themselves hard to capture reliably. Our new metrics for decision quality showed a large variability in responses compared to the analytic tasks. This is likely due to the fact that our multi-attribute choice tasks involve personal preferences and are inherently subjective. In addition, participants may not be able to perfectly express (or be aware of) their criteria preferences, which likely adds further noise to our accuracy metrics. As a result, many of our metrics are not sensitive enough to capture differences across conditions that likely exist [Coh94]. Additional work is needed on establishing more sensitive metrics of choice quality. It seems though that the time metric can become a useful tie breaker when participants achieve sufficient decision accuracy across techniques.

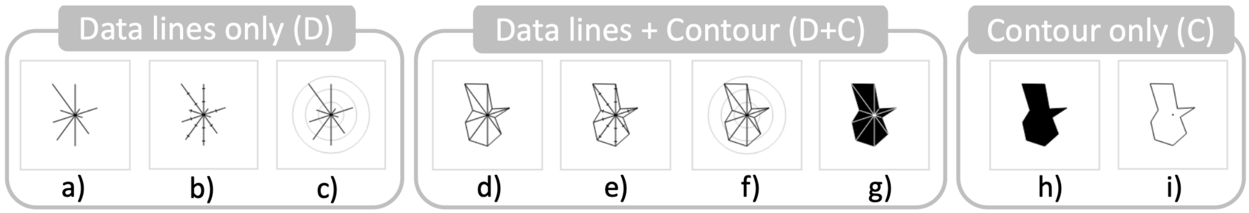


Figure 3.9: **Star Glyph Variations.** Three types of star glyph variations we tested in our contour experiments. Data lines only (D): only the data lines encode the data; Data lines+Contour (D+C): data lines are connected at the endpoints to create a closed shape; Contour only (C): only the contour line is drawn. Additional variations are tickmarks (in b, e), gridlines (in c, f), and fill style (in g, h).

3.2.2 *Other specific questions [FIB+14, FIBK17, IBDF11, DBJ+11, BBIF12, BBB+19]*

Beyond the work I have been involved in regarding how we make decisions using visualizations, I have contributed on several other pieces of work that also start from a very specific question. These center in how we visually perceive information under different conditions. A common theme across all my work that starts from such specific questions is the methodology. These questions most of the time can be turned into a hypothesis, that can be in-turn tested in one or more controlled experiments, varying only a small number of variables between conditions in order to isolate the impact of these variables.

Multi-dimensional Glyphs. As I mentioned in my work with neuroscientists on timeseries similarity (subsection 3.1.2), the notion of similarity is an important task as it allows viewers to detect patterns, group information and identify outliers. With colleagues we investigated aspects that affect the perception of similarity for star glyph visualizations [FIB+14].

A Star glyph is a small, data graphic that represents a multi-dimensional data point, often used in small-multiple settings, on maps, or as overlays on other types of data graphics. In these settings, visual comparison of the star glyphs, to group similar ones or find outliers, is important. There are many variations of star glyphs, that differ in the amount of reference structures they use, the existence "rays," or whether or not the individual rays are connected to form a contour for the glyph [War02]. For example there is a version of the star glyph with unconnected rays, sometimes called *whisker or fan plot*, or a connected version also called *star plot* [War04]. We hypothesized that contours would affect the detection of similarity, as previous work has showed that a closed contour influences the perception of a coherent shape [EZ93]. Possible star glyph variations that we tested are seen in Figure 3.9.

To test this hypothesis, we conducted three experiments. In the first, we explored if contours influenced how visualization experts and trained novices chose glyphs with similar data values. We found that contrary to what we expected, star glyphs without contours make the detection of data similarity easier. Given these results, we conducted a second study to understand the intuitive notion of similarity when it comes to star glyphs. This second study confirmed that star glyphs without contours most intuitively supported the detection of data similarity. In a final experiment, we tested the effect of star glyph reference structures (like tickmarks and gridlines) on similarity. Our results show that adding reference structures does improve the correctness of similarity judgments for star glyphs with contours only, but does not seem to influence the perception of similarity for the standard star glyph. Based on this, we have evidence that the simple star glyph without contours performs best under several criteria, reinforcing its practice and popularity in the literature. Contours seem to enhance the detection of other types of similarity, (e.g., shape similarity) and are distracting when data similarity has to be judged.

This work on glyphs led us to realize that beyond star glyphs, there are many different glyph variations that have been introduced in the literature to better fit certain data types, or to solve specific tasks more effectively. There are in fact nearly endless possibilities on how to map data dimensions to visual glyph encodings [Mun15]. This flexibility allows designers to envision new glyph representations for specific contexts, but at the same time it is overwhelming. Understanding when and which types of designs work best or are preferred by viewers, can guide designers and practitioners. Many user studies in the literature have investigated different data glyph designs and their variations (in particular Chernoff faces [Che73]). Nevertheless, up to our work, there was no systematic overview of these studies. With my colleagues we present an overview of past studies on glyphs, by systematically sampling the literature (64 papers from the visualization literature as well as work from statistics and psychology) [**FIBK17**]. We list their designs, questions, data, and tasks. We also discuss the types of glyphs and their design characteristics analyzed by researchers in the past, and a synthesis of the study results. Based on our meta analysis of all results we further contribute a set of design implications and a discussion on open research directions.

Dual-Scale Charts [IBDF11]. Line charts, are arguably among the most common data representations. One of their problems in practice is that they become difficult to read when the amount of data goes beyond the available display resolution. In particular, there are cases when the density of data points (or their degree of interest) is not uniform, e.g., in historical timelines with dense event clusters and large empty spaces in-between. One way to overcome visual resolution limitations is to use more than one scale (number of data units per display unit) in the same chart. This results in charts where there are some regions with high magnification and others with low magnification. There are several ways to visually integrate these different scales. A popular approach is cut-out charts, where part of a high scale (context) chart, is seen in a cut-out rendered in a lower scale for detail (focus). Another is superimposed charts (Figure 3.10), where both focus and context regions are drawn along the full width of the chart and share the same y-axis. Yet another, consists of changing the resolution along a single axis, i.e., applying a non-occluding step function [Car99]. With my colleagues we set out to investigate if indeed the best approach is to split the chart as it is commonly held, or if other variations fair well under elementary graphical perception tasks such as comparing lengths and distances [**IBDF11**].

After discussing the design space of dual-scale charts, we compare five chart variations using a traditional magnitude estimation task for the visual variables position, distance, angle. These visual variables are the most highly ranked among Clevelands tasks [Cle85] and we hypothesized them to be most impacted by changes in scale. In our experiment design, we varied the location of the item that participants had to estimate (in the high magnification area/focus, the low magnification area/context, and across both areas). An experiment with 15 participants showed that cut-out charts which include collocated full context and focus are the best alternative, and that superimposed charts in which focus and context overlap should indeed be avoided.

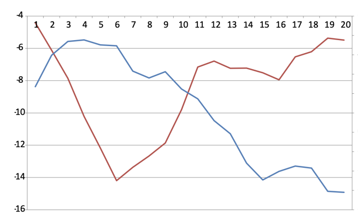


Figure 3.10: Superimposed chart (the red line is plotted according to the left and the blue according to the right y-axis).

Animated Transitions [DBJ+11]. Animated transitions are often used to smoothly convey the transformation between visual states in visualization design, to switch between data dimensions [EDF08], visual representations [HR07, YFDH01], or when navigating in time [**CDBF10**][Gap06]. Apart from the aesthetic appeal, research suggests that it also helps users to understand the underlying data [BB99, HR07, TMB02].

Cartoon animators sometimes use a "slow in" or "slow out" effect [CU93a, JT81], causing more frames to be dedicated to the beginning or end of the animation. Essentially, slow-in and slow-out *distort* time throughout the animation. Computer applications have adopted this idea [KB84], and so have many graphical toolkits (e.g., [BGM04, BOH11]). There are several arguments for using slow-in/slow-out, the most important being the claim that the slow-in/slow-out pacing helps users to anticipate the beginning and ending of the animation. However, no perceptual studies have been performed to confirm this informal design rule. We set out to investigate this question [**DBJ+11**].

We study the question under an object tracking task: tracking a single point within a moving point-cloud. We vary the type of temporal distortion: constant speed transitions, slow-in/slow-out, fast-in/fast-out, and an adaptive technique that slows down the visually complex parts of the animation. In our paper we explain in detail the different types of these timing profiles. Our 12 participants tested the different timing profiles in one real and one artificially generated dataset, where we controlled the movement profiles and number of distractor objects. Our results showed that slow-in/slow-out indeed outperforms other techniques, but we also saw other subtle differences depending on the type of visual transition.

Sketchiness for Uncertainty [BBIF12]. Often the quality of the data we collect (and want to visualize) is uncertain and researchers in visualization have long studied how to best represent this uncertainty [HQC+19]. With a colleague that has worked in the past with utility companies, we observed that this uncertainty is often not in numerical form, but rather it is qualitative. For example, past work on utility maps discusses types of uncertainty in their maps of assets as being (from least certain to more certain): schematic, assumed, indicative, third party survey, and internal survey [BD09]. Such ordinal data can be visualized using Bertin's [Ber10] visual variables texture, value, or size. When it comes to visualizing uncertainty, a set of visual variables are considered more 'intuitive' for this domain; examples include blur, dashing rendering of lines, and color saturation. These variables may bear direct perceptual resemblance to what the uncertainty indicates and, thus, may provide an easier reading of uncertainty [CR00, Mac92, SMI99]. As we are not aware of any study that actually tests if indeed these particular encodings are more intuitive, and how good they are at communicating uncertainty, we investigated further [**BBIF12**].

Apart from the visual variables of blur, sharpness of focus, and color saturation, we considered an additional one *sketchiness* as a visual variable. Sketchiness has already been used to portray uncertainty, in the domain of archeology visualization [SMI99, PGG+09] and for the visualization of 3D shape concepts in CAD [NKD06]. We thus want to explore if sketchiness as a visual variable can be used to depict uncertainty information in line marks such as for graphs, hierarchies and route maps. To generate different levels of uncertainty, we provide an empirically-based method that results in strokes that resemble hand-drawn strokes of various levels of proficiency (ranging from child to adult strokes), where the amount of deviations in the line corresponds to the level of uncertainty in the data.

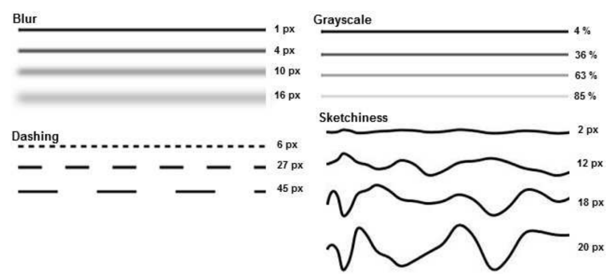


Figure 3.11: The established levels per uncertainty visualization technique.

To test our hypothesis we ran two crowdsource studies that compare in a qualitative and quantitative way the four uncertainty visualization techniques (blur, dashing, color saturation, sketchiness). Overall, our results show that sketchiness is as intuitive as blur when it comes to viewers spontaneously associating it with uncertainty. People can comfortable recognize up to 3-4 different levels, that we found is very similar to other visual variables (Figure 3.11). We thus think sketchiness is a viable alternative for visualizing uncertainty in lines (and has indeed been used by others [WII+12b]). Nevertheless, it is interesting to note that in terms of preference people subjectively prefer dashing style over blur, grayscale and sketchiness for depicting uncertainty for line marks. Some of their comments indicate that sketchiness (and to a lesser degree blur) changes the line geometry, and so tend to consider squiggles as being related to the actual underlying spatial features. Thus as a variable sketchiness should be avoided for spatial contexts such as maps to indicate uncertainty, but rather be used in abstract contexts such as hierarchies and diagrams.

Glanceability in Smartwatches [BBB+19]. This final work starts from a specific question, but is largely motivated by the new available technology of smartwatches. Field studies of smartwatch use show that most interactions with them involve quick glances or peeks at the smartwatch [PBML16], which were often shorter than 5 seconds. These quick glances limit the amount of information a viewer can take in. Given this short viewing window, with my colleagues, we set out to understand if smartwatches are appropriate for viewing visualizations, and what visualizations are better suited for them [**BBB+19**].

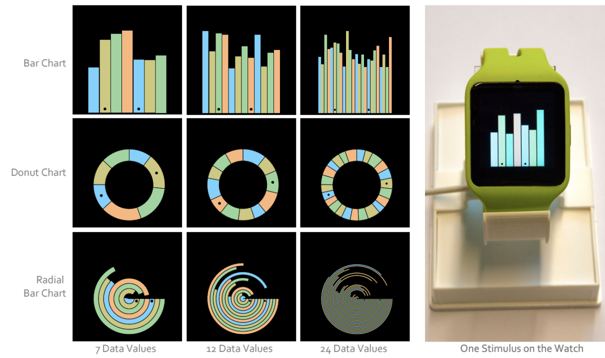


Figure 3.12: Example images of the stimuli used for Bar, Donut, and Radial visualizations, with 7, 12, and 24 data items. On the right, one stimulus shown on the smartwatch.

We focus on three visualizations that we saw often in smartwatch faces: bar charts, donut charts, and radial bar chart (Figure 3.12). We first empirically derive metrics on how people position and orient a smartwatch when reading information, by capturing participants' hand and head positions while reading a watch [**BBB$^+$18**]. Then we study how "glanceable" the three visualizations are. To do so, we ran lab studies where participants performed a simple data comparison task (comparing the size of two marks). The visualizations were shown using a staircase procedure that varied progressively the duration participants had at their disposal [KP10]. We tested this task for three data sizes 7, 12, and 24 data values, with 18 participants each. The difference between studies was that in the first study, the compared marks had a controlled size difference of 25%, while we used randomized data for the size of the targets in the second study.

For both studies, participants performed the task on average in $<$300 ms for the bar chart, $<$ 220 ms for the donut chart, and in $<$1780 ms for the radial bar chart. Glanceable time thresholds in the second study were on average 1.14–1.35$\times$ higher than in the first study. Our results show that bar and donut charts should be preferred on smartwatch displays when quick data comparisons are necessary. Nevertheless, all glanceable thresholds (even for donut) were less than the 5 seconds that field studies have identified as the time spent glancing at a smartwatch. This indicates participants can reliably do quick comparison tasks on smartwatches using visualizations.

3.3 Start with the tool - EvoGraphDice

In 2010 I got involved in a national project that brought together researchers from complex systems, visualization and learning. There we started a discussion about how to best explore large multi-dimensional spaces, such as these represented by simulations of complex systems (such as aircraft air quality, wine or backing production). With colleagues from visualization (N. Boukhelifa) and Evolutionary Computation and AI (E. Lutton) we started thinking about combining ScatterDice [EDF08] existing multi-dimensional tool we had access to, with learning from users' actions, in order to guide them when exploring large multi-dimensional spaces. This resulted in a series of published work [**CBL12a**, **BTBL13**, **TBBL13**, **BBTL15**, **BBT$^+$19**] around EvoGraphDice, the tool we developed. Figure 3.13 shows a screen shot of the tool.

3.3.1 *EvoGraphDice [CBL12a, BTBL13, TBBL13, BBTL15, BBT$^+$19]*

Scatterplot Matrices are a often used to visualize multi-dimensional data (as discussed in section 3.2). Nevertheless, it still suffers when it comes to the number of possible dimensions it can show in the matrix. Nevertheless, visually exploring a large space of alternative views on the data (individual scatterplots) to find interesting patterns is challenging. One way to aid users navigate the a space is the "grand tour" method [Asi85] which provides a complete view of the search space through a sequence of projections showing various viewpoints of the data. However, the time required to inspect these views may be prohibitive [Hub85]. A related approach is "projection pursuit" [Fri87], where the aim is to only visit interesting views (projections that deviate more from a normal distribution). The criteria for deciding whether a projection is interesting have mostly been defined prior to user exploration, using objective measures such as the quality metrics [BTK11].

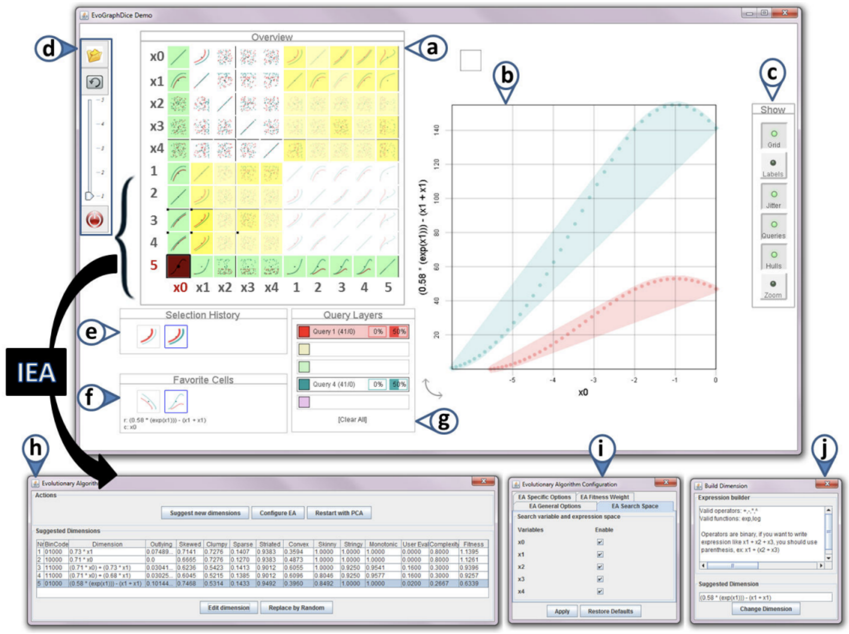


Figure 3.13: **EvoGraphDice** showing an exploration session of a synthetic dataset. An overview scatterplot matrix (a) showing the original data set of 5 dimensions ($x0..x4$) and the new dimensions ($1..5$) suggested by the evolutionary algorithm (colored yellow). The main plot view (b) with two selection queries active, in blue and red. The tool bar on the right (c) changes attributes for the main plot view. On the left (d) is a tool bar for controlling the evolutionary generation, with (top to bottom) "favorite" toggle button, "evolve" button for generating new dimensions, a slider for analysts to give a subjective rating of cells, and a restart (PCA) button. Tool with the history of selections (e) and the favorite cells (f), as well as a selection query window (g). Finally, on the lower part is the main control window for IEA (h) that opens on demand, with a window to limit the search space of dimensions (i) and a dimension editor (j).

To aid this exploration, we build a visual analysis tool to explore multidimensional datasets [**CBL12a**]. The system proposes interesting views based on both: objective measures, such as visual patterns in the two-dimensional projections of the data (like scagnostics [WW08]); and subjective measures corresponding to user satisfaction with the presented view. These subjective measures are not known prior to user exploration. We combined the ScatterDice tool [EDF08], with low dimension projection to handle data multi-dimensionality. Projections can be linear and non-linear combinations of dimensions for an axis of the projection. User exploration is guided by an Interactive Evolutionary Algorithm (IEA) which can both generate new views and adapt to user interest. At the start of the exploration when no user rankings are available we use PCA to propose new dimensions. The tool can be seen in Figure 3.13.

Evolutionary Algorithms (EA) are stochastic optimization heuristics that copy, in an abstract manner, the principles of natural evolution that let a population of individuals be adapted to its environment [Gol89]. They have the major advantage over other optimization techniques of making only few assumptions on the function to be optimized. In short, an EA considers populations of potential solutions (in our case projections of the data) and individuals that reproduce are the best ones with respect to the problem to be solved. Whether an individual is a good solution is judged in our system both by objective metrics, and subjective user feedback. Evolutionary optimization techniques are particularly efficient to address complex problems (irregular, discontinuous) where classical deterministic methods fail [Ban97, PLM08], but they can also deal with varying environments [JB05], or non computable and subjective quantities [Tak08]. Our relevant publications in both EA and visualization venues explain in more detail the visualization and algorithmic aspects of the system [**CBL12a**, **BTBL13**].

EvoGraphDice Evaluations. From the beginning of our work on the system we were considering how to go about validating the tool. EvoGraphDice is at it's core an Interactive Visual Machine Learning system (IVML) where, a human operator and a machine collaborate to achieve a task (also known as human-in-the-loop [BKSS14] or mixed initiative systems [Hor99]). It is by no means the only one, there are examples of such tools for classifying or clustering a set of data points [Ame12, BLBC12], for finding interesting data projections [BKSS14, CBL12b], or for designing creative art works [Lut06, SPCL+13]. The underlying assumption is that the human-machine co-operation yields better results than a fully automated or manual system. But how does one go about validating such a system?

We first look at the *usability* and *utility* of the tool for different types of experts that need to analyze multi-dimensional data **[BTBL13]**. We reached out to our research institutions and networks and recruited five domain experts which explored their data using our tool (from domains like scientific simulation, medicine and geography). Due to the open-ended style of exploration using EvoGraphDice, and the subjective nature of the goals of each expert, we chose a qualitative observational study methodology [Car08]. Our goal was to determine first if our tool is understandable and can it be learnt. And second if it was of value to the experts, in particular (i) if they could confirm known insight in their data, and (ii) if they were able to evolve new views with combined dimensions that contain new insight, allow them to generate a new hypothesis, and if so how easy or difficult is it to reach those findings.

Almost all participants were able to easily confirm prior knowledge about their datasets. One expert found this task challenging because of the lack of data aggregation that her type of analysis requires. Overall, participants were able to confirmed known insights in their data, such as correlation, clusters or outliers. If we include hypothesis formation as part of insight generation (as is done in other insight based evaluations [SND05]), four of our participants generated new insight in the form of distinct observations about the data and formulated better hypothesis, and one formed new hypothesis. They highlighted the following benefits of the tool: try out alternative scenarios by editing dimensions, think laterally, quantify a qualitative hypothesis, formulate a new hypothesis or refine an existing one. We observed that the frequency of evolving new dimensions or limiting the search space of dimensions varied across experts, depending on whether they had an a-priori hypothesis. The looser the initial hypothesis, the more often they tried to change the search space; and the more focused the hypothesis the more generations they inspected. These two strategies of *exploration* and *exploitation* are supported by EAs [Ban97] where on the one hand the user wants to visit new regions of the search space and on the other hand they want to explore solutions (combined dimensions) close to one region of the search space.

We then look at the *algorithmic* behavior of the tool **[TBBL13, BBTL15]**. To do so, we conducted a more controlled experiment to observe how users leverage the system and how the system evolves to match their needs. The goal of the user study was to collect data about user interactions (ratings, evolutions, favorites) and the fitness function that they system is trying to optimize. In particular, we wanted to understand user strategies in solving an exploratory task, and the algorithmic convergence, focusing on the learning behavior of the algorithm across generations and its ability to adapt to user focus. The task was designed as a game. We synthesized a 5D dataset with two curvilinear dependencies between two variables ($x0$ and $x1$) and random data for the rest of the dimensions (Figure 3.13). Participants were asked to evolve a scatterplot where it is possible to separate the two curves in with a straight line (equivalent to separating the two corresponding convex hulls). Of our twelve participants ten successfully separated the two curves, while the remaining participants evolved views very close to a correct solution within the allocated time of 20 min.

We examined the collected data in terms of: user strategy analysis to understand the different approaches users took to solve the task, convergence analysis to assess the algorithm's ability to steer the exploration toward a focused area of the search space, and diversity analysis to assess the richness and variability of solutions provided by the algorithm. When it comes to participant strategies, we show that they centered around three dominant scagnostics (skinny, convex and sparse) that are relevant for the game task. And that the stability of the exploration strategy may be an important factor for determining the outcome of the task and the speed of convergence. Successful game sessions had a more consistent strategy when compared to the unsuccessful ones, and they converged more quickly on average. When it comes to convergence, again user strategy seems to play a role. On average the surrogate function the system is optimizing follows the order of user rating of scatterplots fairly consistently, even though users seem to take different rating strategies: coarse - tending to lump evaluation scores to fewer levels, or fine-tuned - covering all possible rate levels, or a combination of both). Our results suggest that when users taking a more consistent approach (either fine-tuned or coarse) the system seems to converge more quickly. Finally, our diversity analysis shows that, in terms of the visual pattern, the system provides more diverse solutions at the beginning of the exploration session, before slowly converging to a more focused search space for most sessions. These effects correspond to the exploration component (random search) and the exploitation component (focus) of the genetic engine.

Overall, we found this type of dual evaluation provided us with a wide perspective of our tool. Looking at the system as a general visual analysis tool gave us evidence of its utility. And evaluating its convergence to user strategies provided us with evidence that it can evolve based on user needs. Moreover, it highlighted interdependencies in the combined human-machine system: consistent searching and rating strategies from the human side led to faster conversion and solutions. We expect that as users use the tool more they will adopt these strategies (following the notion of co-adaptation [Mac00], where both humans and machines adapt to each other). In a recent publication, we list a number of such systems and the types of evaluations they employ **[BBL18]** and found that the majority usually evaluate one aspect of the tool, focusing on the quality of the user interaction with the system (human-centered evaluations), or the robustness of the algorithms that are deployed (algorithm-centered evaluations), and only in a few cases detailed attention is drawn to the quality of human-machine co-operation and learning. We are still investigating what are best ways evaluate such systems **[BBL18]** an interest we share with other visualization researchers as we saw in a workshop on the topic in IEEE VIS 2019[4].

3.3.2 *Another tool - GraphDice [BCD+10]*

The EvoGraphDice tool was not the only one inspired by ScatterDice[EDF08]. In 2010 when I was visiting the AVIZ team that built the tool we had the inspiration to see if it would be possible to extend it to use in multivariate graphs. This led to GraphDice **[BCD+10]** that is the first tool to use a plot matrix to navigate multivariate graphs. Node-link diagrams can be seen as extensions of scatterplots where data points are connected with links, thus the use of node-link diagrams in the system is consistent with the ScatterDice "dice rolling" paradigm. Because in multivariate data, attributes can be categorical (e.g., in a social network it can be Gender or Country), this results in multiple nodes overlapping on one or more dimension axes. To provide a clearer visual indication of this overlap, we use jitter [AS07], with overlapping nodes being placed around a circle. The design of links is more challenging. For edges with a numerical or an ordered categorical attribute, we automatically create an associated interval node attribute and aggregates the edge values at each node. For example, in a co-authoring network, edges are articles with a publication year. GraphDice automatically creates an interval node attribute for the "year" attribute named "e:year" Links are drawn as curves with their endpoints positioned according to the value of their "year'"attribute. Thus GraphDice allows the exploration of node and edge attributes in a unified way.

Early on in the design of the tool, we felt it would be of interest for Social Network analysis [WF94], as they typically associate data to both vertices and edges. For example, if vertices are persons, they can have a name, a birth date, a position in a company, and many more attributes. Similarly, relations can also have attributes, such as date of friendship connection. To this end we early on also added the option to apply social network analysis (SNA) algorithms to the visualized network, producing additional attributes such as degree, centrality, and clustering coefficients.

At the end of the system design, we conducted a full day workshop with a historian specializing in

[4]EVIVA-ML Valuation of Interactive VisuAl Machine Learning systems `https://eviva-ml.github.io/`

SNA. The participant used GraphDice to explore a historical migration dataset and was able to use the tool effectively with less than 15 min of training. Our historian commented on how easy it is to manipulate GraphDice compared to her other tools (like UCINET [BEF09] or Pajek [dNMB05] available at that time). She pointed out it is very well suited for discovering patterns and learning new datasets, as it provides visual representations of a network from many points of view. She added that even if she doesn't want to learn statistics (a requirement in most current tools) but can still see results. GraphDice [**BCD+10**] is the tool that we extended into EvoGraphDice.

3.4 Start with the data and task - GeoTemporal tasks [**PPB20**, **PBP20**]

In this last section I describe briefly work that started from discussions between V. Peña Araya and E. Pietrga, when we were considering geotermporal data and what are the best visualizations for them. Before being able to answer, we had to decide what tasks we were interesting in.

Geotemporal Correlation [PPB20]. We were inspired to look at this question by the Hans Rosling's 2006 TED [Ros06] talk on country demographics. To understand the story and the insights that the speaker reveals, the viewer has to look in an animated scatterplot, at the life expectancy and the fertility rate together. The spatial dimension also plays an important role in this story. Rosling refers to individual countries, but also different groups of countries multiple times. This example highlights the importance of multivariate geotemporal data visualization, where insights involve two variables that are related thematically, that are situated both spatially and temporally [ATC+15].

The problem of designing an appropriate visual representation in this context is challenging, as multiple data of different nature must be combined, each having specific characteristics: the thematic variables (life expectancy, fertility rate), the spatial properties of those entities (countries, continents), and the evolution of the thematic variables over time (years). Prior studies have compared geotemporal visualization techniques for a single variable that evolves over space and time [GMH+06, SdWvW14, LD11, LD12]. Others have looked at two variables on a map (bivariate maps), but at a specific point in time [Elm13, GLQ18, NASK18]; or at how to visualize the correlation between two variables [RB10, Ren17, KH16b, YHR+19], including visualizations that can be used to depict temporal evolution [HYFC14b], but not in a geospatial context. Thus choosing which visualization technique is effective at communicating correlation between two thematic variables, that evolve over both space and time, remain unclear.

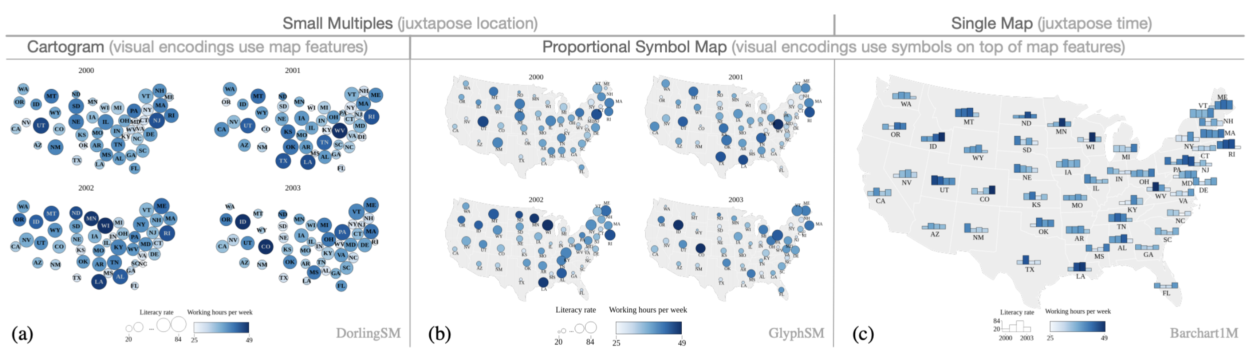


Figure 3.14: The three visualizations compared in our first study. (a) Dorling cartograms as small multiples, (b) proportional symbols (circles) on maps as small multiples, and (c) proportional symbols (bar charts) on a single map. In this example, each map shows the values of two artificially-created variables over four years. In each case, both variables have an overall positive correlation (Pearson correlation coefficient ≥ 0.75) and no monotonic evolution.

In our work we identify different possible strategies to combine thematic, spatial and temporal data into a visualization. We can decide to combination thematic variables in a way that juxtaposes all time steps for a given location (with symbols on a single map), or juxtaposes all locations for a given time step (small multiples). We also need to decide how to encode the thematic variables themselves, using symbols on the map or using features of the map itself. This categorization gave rise to three promising techniques that can be seen in Figure 3.14.

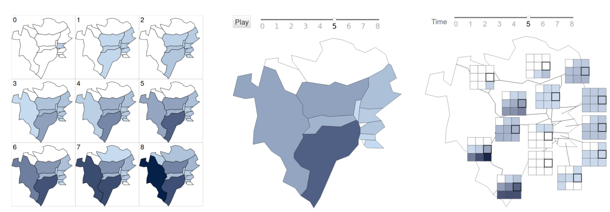
We designed a study to determine which visualization better communicates whether two variables are correlated over time (or not), and if they are, what is the pattern to their evolution. Given the nature of the techniques that either juxtapose time or space, we expect them to fare differently depending on the number of time steps and the number of geographical entities. We consider 9 time steps and 48 locations. And our tasks varied in granularity for both dimensions: time (all time steps, a subrange of steps, one step only) and space (all locations, locations in a subregion, one location only).

We ran a study with 18 participants. Our results confirm the intuition that the techniques perform differently depending on the spatial and temporal granularity. Small multiple variations perform better for a single point in time, whereas a symbol map performs better when considering all time steps. The situation is less clear when considering a time range: if there are only a few locations, a single map with bar charts is better, if there are many locations, small multiples are better. We did not find any evidence that proportional symbol maps are better than Dorling diagrams. Extending this study to more thematic variables is a clear next step.

Our IEEE VIS/InfoVis 2020 paper provides details about the dataset generation and tasks [**PPB20**].

Geotemporal Propagation [PBP20]. A phenomenon that has been largely unexplored in geotemporal visualization is *propagation*. Propagations are the topic of investigation of several domains, such as disease spread in epidemiology, viruses in cybersecurity, keyword tags becoming viral on social media. In some of these examples it is the network topology that matters, but in others it is the geography that is key in understanding the nature of propagation. An example is the Ebola virus epidemic in West Africa from 2013 to 2016, that caused more than 11,300 deaths and had major socio-economical disruptions [Org19]. By analyzing the propagation patterns, it was found that the virus tended to disperse more frequently among geographically-close regions, and mainly within countries [DCR17]. This is an important insight for health officials looking at strategies to attenuate the spread.

Propagation has similarities to other spatio-temporal movement patterns [DWL08], for example, spatial autocorrelation has been found between influenza and commuting paths [CPB14]. However, the origin of propagation makes it also unique. It is not induced by the motion of entities in space, but by the *replication* of entities (virus infection, meme or hashtag reuse). This replication and the context in which it develops makes propagation different from other movement patterns in two respects. First, if we consider human epidemics, diseases propagate continuously, but todays' transport connectivity also allows for geographical hops [MMD+13]. This differs from movement trajectories of individual entities (e.g., migrating birds) or vector fields (e.g., water flow) that are continuous. Second, the replication of entities allows propagation to have multiple synchronous strong peaks in different and distant geographical places [VBS+06]. This is not the case when looking at individual entity movement, as their discrete nature constrains the number and strength of synchronous peaks.



(a) Small-multiple Maps (b) Animated Map (c) Glyphs over Map

Figure 3.15: Propagation visualization strategies evaluated. The illustration shows a simplified dataset with only a small region and 9 time-steps: (a) small-multiple maps; (b) a single animated map; (c) a single map with glyphs overlaid over each region - each cell in a glyph encodes the value of one time-step for the corresponding region.

We set out to understand what geovisualization techniques best reveal propagation patterns. Again we divided the techniques in ones that juxtapose space or time, and identified three main visualization strategies: small-multiple maps, animated maps, and maps with glyphs (Figure 3.15). We compared them using five tasks derived from both a general taxonomy of movement patterns [DWL08] and from the literature on propagation analysis. In these tasks, 18 participants had to characterize one of the following dimensions of propagation phenomena: their speed and virus arrival at specific locations, their direction, their geographical scope, as well as the temporal location of peaks, and the presence or absence of spatial jumps in the pattern. The data was generated using a widespread propagation model [New02], that we tweaked to ensure the task had the desired characteristics (such as hops). We started out with several hypothesis. For example we expected that animation would be best at tasks that require comparing the map across multiple time frames, such as detecting direction and sudden changes (such as hops). Whereas a single map with glyphs would be better for tasks focused on a single area (such as detecting the arrival of the virus). Overall, our results show that small-multiple maps perform the best overall, but that both animation and maps with glyphs outperform them in specific tasks (identification of direction and arrival time respectively). As a next step we plan to consider visualizations that abstract parts of the propagation patterns (e.g., aggregate major directions of movement) to see if these communicate better the different aspects of propagation.

Our ACM CHI 2020 paper provides details about the dataset generation and tasks [**PBP20**].

3.5 Conclusions and Reflections

This chapter provides a panorama of my past work investigating how to build appropriate visual representations, from different starting points.

The first part of the chapter (section 3.1) focuses on work that started from the needs of specific domain users, such as business intelligence analysts [**EB11**, **EB12**, **EAB13**], genealogists [**BDF+10**] and neuroscientists [**GTPB19**]. Here we followed a user-centered design methodology: starting with understanding the context and the challenges that the domain experts face, continued with design sessions for possible solutions, and ended with their feedback on the final designs. Their needs motivated my colleagues and I to come up with innovative solutions, such as context aware annotations for dashboards, that attach annotations to data points or queries irrespective of chart and aggregation [**EB12**].

The frequency and intensity of end-user involvement varied across these projects, for several reasons. For example, when starting work on storytelling for business intelligence in 2012, we had no preconceptions of what the users needed, and the topic of storytelling in the visualization community was just starting. As such, the inspiration of the final tool came mainly from our interviews and participatory design workshop with them, necessitating several iterations. In our other work with these analysts, we were greatly inspired by the technology available to them, and previous research from the field. For example, the fact that their tool had mechanisms to suggest charts drove our research questions for our work on dashboard creation, as did the data-layer available in the dashboard that we could use to create annotations that were attached to data points and queries. In these later cases, expert feedback evolved and refined our designs, but was not the initial inspiration of our work. Similarly, for our work with historians, our initial inspiration was past work (Quilts [WBS+08]) that we extended based on the genealogists' needs (e.g., making sure to stress temporal aspects, genealogy cycles and overlaps).

Technology being the inspiration behind user-driven research, is not a bad thing. On the contrary, it can prove beneficial. After understanding user practices, in user-centered design we often conduct participatory design sessions to get experts' input on potential design ideas. But it can be challenging for end-users to come-up with innovative solutions, as they cannot always envision designs that are far from what they are comfortable with. And techniques such as paper prototyping to test ideas can aid with testing interactions, but do not necessarily communicate the visual complexity of entire datasets. I have found it very useful in such situations to have toy prototypes (of the technologies that inspired the work) as a way to give end-users inspiration, to ground our dialogue with them, and to show them things that are possible, so that they can reflect on and build-upon with their own ideas. In all projects with expert users, access to them can be challenging. My experience has been that as researcher we need a "champion" from the user side, someone (analyst, historian, neuroscientist) that is part of the experts' organization, is intrigued by visualization, and becomes the crucial contact point that ensured continuous access to other experts. We were very fortunate to have such champions in all projects. This is a role that previous work on design study papers [SMM12] does not seem to distinguish from other end-user roles (gatekeeper, front-line-analysts, etc.).

In the case of business analysts, it was also extremely helpful to have M. Elias conducting her PhD in the mist of their organization. This allowed her to establish a level of trust with the experts. And although the development of prototypes on top of their existing working platform was technically challenging, this helped experts see the immediate value of our work and be open to long-term collaboration, participate in studies, workshops, etc. The domain of HCI has long accepted the value of ethnographic approaches in design (e.g., rapid ethnography [Mil00] in field studies, observations and designers embedded in end-user teams), but these are less adopted in visualization research. The seminal design-study paper by Sedlmair et al. [SMM12] argues in-fact against traditional ethnographic approaches (such as fly-on-the-wall observations) as they are very time consuming and often not revealing of analysis goals. I feel that there is potential in having visualization researchers embedded in the end-users communities, even periodically, to improve rapport, trust and common understanding. This position is now shared by other researchers, as seen in the recent publication on design by immersion [HBH+20].

The second part of the chapter (section 3.2) describes my past research that starts from fundamental and theoretical questions, including our investigation of whether cognitive biases exist when making decisions using visualizations [**DBD17a**, **DBBF19**, **DFP+20**, **DBD17b**, **DBD18**]. A traditional methodology applied when facing such questions is to start with concrete questions that can help isolate a phenomenon to study. For example in our work on biases it was to chose and focus on one bias (the attraction effect) and progressively built our study (replication first to see if it exists in visualization, extension to more realistic visualization contexts, mitigation of the effect).

The possible inspiration behind this type of work is also very broad. In my case some comes from attempting to prove (or disprove) commonly held beliefs. Our work on dual-scale charts [**IBDF11**] was motivated by recommendations starting from Cleveland [Cle85] that dual-scale charts need to have clear visual breaks across scales, and the popular belief they are not appropriate representations (e.g., [Few08]). Or in our work comparing different temporal profiles for animation transitions [**DBJ+11**], we set out to investigate if "slow in" or "slow out" animation effects help viewers track movement better (as it has been advocated in the design community [CU93b, TJ95]). In our work we found indeed evidence to support these beliefs. But it has not been the case for other research work. For example a series of studies on pie charts [SK16, KS16] have suggested they are not read based on slice angle (as was commonly held), but rather based on arc length and less often on arc area. Or work that has shown that some types of chartjunk (that is generally held as being undesirable [Tuf86]) can help make charts more memorable [BMG+10] and engaging [HKF15]. All this work and findings are the result of directly questioning assumptions we hold as researchers and visualization practitioners. I feel this is a sign of a very healthy community that does not feel threatened to self-reflect and put under the microscope long-held beliefs.

The work on cognitive biases and decision making in particular, started when E. Dimara set out to investigate how groups make decisions. Starting reading the relevant literature brought us face to face with the more fundamental question, how do we make decisions when using visualizations (even alone, before considering complex group decision making). It is tempting to assume that data understanding leads to good decision making, but digging into literature from economics, sociology and marketing showed us the way we make decisions is not always rational (cognitive biases). We thus embarked on a journey to understand how cognitive biases may affect our decision making when using visualizations, and what we could do about it. This was challenging. E. Dimara had to dig into obscure and hard to digest research in several domains (economics, psychology, sociology, marketing), analyze it, and came up with a perspective that is actionable for the vis community (that ended in our taxonomy paper). But we feel it is worth it, as it made previously unknown and scattered findings from other domains accessible to the visualize community [DFP+20]. It was a proud moment when a senior researcher said that some of their recent papers were directly inspired by that taxonomy. As a discipline visualization is extremely broad, touching topics that are central to other domains (design, creativity, humanities, machine learning, etc.) and we need people that wear two hats, that understand visualization but are also embedded in these domains, and can translate the findings and challenges to us. Many efforts in our community are done to help us open to these fields (often in the form of workshops like ones bringing Vision research, Data Science, Humanities, ML) but surely more opportunities exist.

The third part of the chapter (section 3.3) summarizes how we started out to adapt an existing visualization tool, Scatterdice [EDF08], to use in different contexts. With GraphDice [BCD+10] we wanted to investigate if the tool's success with multi-dimensional data, could be transferred to multi-variate graphs. While that was our initial inspiration, our work went further into examining the potential of the tool for more specific data and tasks (social network analysis).

Similarly, EvoGraphDice started our with the idea that we could combine the tool with machine learning (in our case interactive evolutionary algorithms) to aid the exploration of very high-dimensional data, providing interesting dimensional views. This work also started by an inspiration for a tool, but led a large sequence of work on applying the use of a tool that combines human and algorithmic computation to different context and domains (and finding the limits of the approach) [CBL12a, BTBL13, TBBL13, BBTL15, BBT+19], including work on collaboration mentioned in the previous chapter [BBT+19].

While we were trying to find appropriate ways to validate our tool, we were confronted with a more high-level question of what are appropriate methods to evaluate mixed-initiative systems (or Interactive Visual Machine Learning systems). The evaluation here is particularly challenging, as the machine is learning from user exploration, but the user's goals may also be evolving [Mac00], and there is generally no ground truth to be reached. Apart from combining algorithmic (convergence) and user-centered evaluations (such as insight evaluations with experts [SND05]), there may be other possibilities. In our recent work on the evaluation of such systems [BBL18], we put forward the idea of also looking at metrics related to creativity of the solutions reached by the experts - same as art, there is creativity in scientific thinking characterized by lateral thinking and surprising findings. Finding what are the appropriate methods and measures to use to evaluate interactive visual machine learning systems is a topic that requires further investigation and is of interest to the visualization community, as the attendance in our recent workshop on the topic[5] indicated.

Finally, during the design, development and the evaluation of of EvoGraphDice we were fortunate to closely collaborate with an expert in the ML side (Evolutionary Algorithms), that was also an end-user of our tool: she used it to visualize the evolution of generations of her optimization algorithms under different input parameters. In this instance it was the machine learning expert that was immersed in the designers' environment, wearing two hats and enabling the communication across our two domains. This resembles the concept of participatory design, where end-users are incorporated in the design team, but there are also differences. Our expert was not only an end-user providing design feedback, but also responsible for part of the system (ML). Given the immense interest in explainable AI and plethora of work that brings together visualization and ML experts, this types of very fruitful relationships will continue to grow. It is worth considering ways to more systematically structure them, and to consider their appropriateness as both design but also as system validation approaches.

[5] EVIVA-ML workshop on the EValuation of Interactive VisuAl Machine Learning systems `https://eviva-ml.github.io/`

The last part of the chapter (section 3.4) considers examples of my work that start from specific types of data and tasks. In particular, with my colleagues we looked at geotemporal data and tasks that have not been considered before. In the first study, we looked at how to visualize correlation between two thematic variables (e.g., life expectancy and fertility rate) across space and time [**PPB20**]. In the second, we examined the unique characteristics of propagation movements (e.g., a virus) and compared different visualizations in how well they convey them.

The inspiration behind this work was an animation by G. Dudas[6] that accompanied their Nature article on the topic [DCR17]. Their visualization illustrates aggregated patterns of movement, showing the results of their analysis in a data-storytelling manner. But made us realize that we do not in-fact know what are the best visualizations to analyze propagation phenomena and their unique movement characteristics. This led us to consider generally spatio-temporal tasks and it was at this point we were surprised by the lack of general guidelines for spatio-temporal correlation with two (or more) thematic variables. We thus took a step back to deal with the more studied topic (correlation) before moving into propagation.

The goal of both studies is to find the visualization that best supports the tasks, as such we decided on a comparative study for our validation methodology. To ensure that conditions were comparable across visualizations we controlled several factors (such as the correlation coefficient and monotonic evolution in the first, or the propagation duration and number of peaks and hops in the second). We thus needed to control aspects of the data, while at the same time ensuring the data was close to real-life datasets used in the analysis of geotemporal phenomena. This can be a difficult balance to strike. In our case, for the correlation study we imitated characteristics of real datasets (values for thematic variables and correlations). For the propagation study we relied on common simulation models, that we then tweaked to control the presence of desired characteristics. I feel this approach gives a good balance: it ensures that the data share characteristics with the real-world phenomena, while at the same time provides researchers with some degree of control. Nevertheless, the process of can be time consuming. In our case it required a lot of piloting and tweaking of simulation parameters to reach the final outcome. My co-authors and I feel it is very important to share this information, both to save future researchers time and to aid replication of our results and make future studies comparable. We have thus made all our material (code, dataset, analysis) available, a practice that I am happy to see our community adopting at large.

[6] Animation of Ebola virus spread `https://youtu.be/j4Ut4krp8GQ`

4 Perspectives, Future Directions and Closing Remarks

The previous chapters provided a collection of my past work that is centered on wall-display platforms and opportunities and challenges surrounding them (chapter 2); and on work I conducted attempting to create appropriate visual representations or gain deeper understanding about how they influence our perception and decision making (chapter 3). I next describe my plans for future work.

4.1 Visualization using Collaborative Technologies

The different sections in chapter 2 showed challenges and opportunities when designing interactive visualizations in wall-display environments, when it comes to interaction, representation and collaboration. Nevertheless, many opportunities remain.

When it comes to viewing data, our work identified a potential challenge in collaboration since people see information differently from different positions [**BI12**]. The perception of visual information can be affected by other factors unique to collaborative settings, such as the occlusion (or partial occlusion) of areas of the display due to the positioning of one's colleagues around the space. While existing work has discussed using the space that users occlude to show personalized information [KRMD15], it is not clear if this occlusion affects viewers' understanding of information and collaborative analysis more generally. The question on data understanding becomes more complex since wall-displays are rarely used in isolation. Often they are part of a more complex multi-surface environment [BLHN+12] (that may include tables, mobile devices, etc.). Previous work on tabletops [WSFB07] investigated how varying screen orientation from a horizontal to up-right position influenced the accurate perception of elementary graphical elements. Our findings on wall-displays are different (we found an increase in error with horizontal placement that was not observed before). These perceptual differences across displays show that we cannot generalize findings from one display to another, nor can we predict how information will perceived when spread out across displays in complex environments. I plan to continue my work seeking to understand how we view information in such environments.

Beyond traditional collaborative environments (walls, tables), we see an increasing number of augmented reality headsets, that could further the capabilities of these more traditional collaborative devices, by displaying information in any part of the environment (including on the shared devices. This is the topic of the recent bool on immersive analytics [MSD+18], where I co-authored a chapter using such devices in collaborative settings [**BCBM18**]. We discuss situations where new immersive technologies can be used to support analytical reasoning and decision making in general, and in collaborative contexts in particular. While there is a recent research trend to aid and study use of Virtual or Augmented Reality when analyzing information visualization (e.g.,[CCD+17, WWS20], when it comes to collaboration most work focuses on analytics systems (for a recent survey see [FP19]) and rarely studies the impact of the technology on the collaboration itself (e.g., [CDK+17]). With PhD student R. James I am currently investigating how to best combine these personal immersive displays, with more traditional collaborative displays (such as walls and tables). As these technologies vary widely in cost, size, resolution, and ways of interacting with them, we will explore which technology is more appropriate for given tasks and user needs. More particularly, we plan to study when physically collocated setups, such as wall-displays, are required for collaborative analysis, and when less expensive Augmented Reality headsets can be considered as low-resolution alternatives. And where a mix of these technologies can come in play, and be used concurrently in collaborative situations.

4.2 Study Interaction in Visualization systems

Still considering perception, there is little work that explores the interplay between interaction and visual perception. Among them Jota et al. [JNJ+10] studied the impact of viewing angles on pointing performance on a wall and found that the visual size of an object affected performance more than its actual size. When users actively manipulate information, this affects their understanding of it. And Saket et al. [SSRE18] found that a magnitude production study (where participants compare visual variables, but give interactively their response) produces similar ranking results to classic magnitude estimation studies [CM84]. In the context of wall-displays interaction could be a means to mitigate perception limitations. If we can reliably identify visual distortion (as we did in our work **[BI12]**) we can suggest interaction techniques that can help to alleviating specific perception limitations. For example we can envision cases where we bring close to them remote content that we know is perceived incorrectly, or that is currently occluded by colleagues (along the lines of work where content hidden by physical objects [KSMS12] or users hands [VB10] is displaced to different locations). Understanding the impact of interaction in visual analysis more generally, is a broader topic that I want to pursue. With my colleagues we have already seen for example that the choice of interaction techniques can affect coordination and quality of simple graph topology tasks on wall-displays **[PBC17a]**. This is a topic that merits further investigation in complex collaborative situations. Recent work that attempts to pinpoint the notion of interaction in visualization [DP20], also highlights that while the benefits of interaction are generally acknowledged, it is rarely the focus of our research efforts.

More generally, the space of interface design for visual analysis tools remains fairly unexplored. It is only in recent years that novel interaction paradigms are studied in visualization literature. Examples include the work by Jansen et al. [JDF12b] that compares tangible slider controllers with traditional virtual ones in situations when controlling a wall-display. And the work by Saket et al. [SHPE20] that studies the strategies adopted by participants when using direct manipulation [Shn83] to alter the graphical encodings in a visualization (an approach that is adopted in a non-systematic way by several systems, e.g., [RLP10, KC14, KJEY11]). Our own work on SketchSliders **[TBJ15]** is complementary, instead of customizing visual encodings, we consider the customization of the tools that analysts will need in their exploration. I feel this approach merits further consideration, beyond the case of wall-displays. Already the comments open up possible directions of future research, studying the possibility that one interface does not fit all analysis needs. Our participants often felt that the decisions they made about their tools helped them structure their analysis, adjusting and copying these tools allowed them to keep track of alternative explorations, and stated that their tool choices and hand-written annotations captured subtle aspects of exploration that were important to capture and store with their analysis steps. Studying the implications of creating tools on-the-fly for visual analysis is a direction I would like to continue working on.

4.3 Appropriate Visual Representations

My work on creating appropriate visual representations is very broad, and each individual piece of work has opened up avenues for me to pursue in the future. Our work with domain experts continues. For example, with our neuroscientist colleagues we have established the need for them to get access to similarity results quickly, even though they are dealing with a large amount of data, and thus can suffer from computation delays. There are back-end solutions for providing approximate and progressive query results in timeseries data [ZIP16] as it becomes available. Such systems are generally appreciated by users due to their quick feedback [BEF17, ZGC+17]. Nevertheless, how to present these to users is an open question. For example, users can be mislead into believing false patterns in early progressive results [MFDW17, TKBH17] with early progressive results. It is thus important to communicate the progress of ongoing computations [ASSS18, SASS16], including the uncertainty and convergence of results [ASSS18] and guarantees on time and error bounds [FP16]. This is not a trivial topic, as in many cases the underlying data is evolving (e.g., dynamic data) or too large to determine their distribution. How to visually explain such guarantees is also a challenge, as it largely depends on the type of data and the guarantees presented. Moreover, it is not clear if users can interpret them correctly (especially if they are probabilistic in nature and based on partial data), if users trust the guarantees, can reason and analyze data under this uncertainty, and make decisions using them. The last part of A. Gogolou's PhD explores how to (i) provide these guarantees in the context of timeseries (work under review), and (ii) visualize their uncertainty and convergence. As previous researchers, I am interested to see how our experts interpret and make use of these indications in practice during their visual analysis process.

To deal with the large amount of data most experts deal with today, we often combine human and computer analytical capabilities [Hee19, JLC18, BL09b]. We followed this approach in our work that combines visualization and evolutionary computation **[CBL12a, BTBL13]**, to help steered the system towards the exploration of areas that the human finds interesting. These system are often called interactive visual machine learning (IVML) systems. Here the role of the human is not only to interpret and understand the underlying models or decisions (as is the case in Explainable AI). But to also actively act on, and react to, these models. This raises questions related to trust and usability, but also questions with respect to the evaluation of the various facets of the IVML system, both as separate components, and as an entity that includes both human and machine intelligence [SSZ+16]. This is a challenging topic to tackle. First, both the the machine learning component, but also the human, learns and evolves [Mac00] as they progress in their exploration. Thus the desired outcome of the process may be continuously evolving and thus hard to capture and evaluate. Moreover, these complex systems face uncertainty both from the automatic inference [AWV+19] and from the analysts trying to reason under uncertainty [HQC+19]. These characteristics make the evaluation of IVML systems a challenging, but very exciting topic of research. With colleagues from INRA we have started to tackle some aspects of it **[BBTL15]**, considering evaluations both from the algorithmic and human side. Nevertheless, we still don't have a good grasp of what are good measures and methodologies for evaluating such systems, considering possible metrics and taxonomies to categorize the types of systems and how to best evaluate them. This is more generally an open discussion in our community, and in we recently organized a workshop in IEEE VIS on the topic[1] with the goal to elaborate on a concrete research agenda.

And when it comes to more fundamental questions, there are still many questions that remain about how we make decisions using visualization systems. Our taxonomy paper **[DFP+20]** identified many open problems to study, like how our memory of visualizations may be biased, how does the use of automation in conjunction with visualization (such as the use of recommendations based on machine learning) can affect our decision making process, can we use visualizations to mitigate these biases (as we did with the attraction effect **[DBBF19]**), etc. Recent work in the visualization community has started considering possible biases during data understanding and analysis.

More importantly, it raises fundamental questions about how we conduct our own research. For example, confirmation bias is known to exist in research. But is it possible others affect our results and conclusions? For example, when we present users with three systems or techniques to rank, do we introduce attraction effects or compromise biases, leading participants to favor particular systems and techniques? I would like to continue work on this topic, and beyond specific biases attempt to conduct a meta-analysis of work in our domain that may have been affected by such biases.

I plan to start with a smaller set of biases that relate to choice tasks (such as the attraction effect and

[1]EVIVA-ML Valuation of Interactive VisuAl Machine Learning systems `https://eviva-ml.github.io/`

the compromise bias), as in our domain the validation of newly introduced techniques often considers comparisons and subjective rankings. But in the long run I plan to systematically expand to other biases and other research domains that may also be affected.

4.4 Inspiration and Design Process in Visualization

In the following years I plan to explore two high-level questions about how we conduct research in interactive visualization. The first is, as mentioned above, the study of possible biases that affect us as human-computer interaction and visualization researchers. The second is to try to analyze the inspiration behind our work as a community, and explore if there are ways to capture it and share it with others.

Thinking past my own work it has come apparent that sources of inspiration vary. Much of it can come from *domain experts* that reach out to us with their problems (as was the case of my work on storytelling for Business Analysts [**EAB13**]). Other comes from research *colleagues in other fields* that reach out to us with new problems or advances in their own fields. This was the case behind the work I did with colleagues from evolutionary computation (the development of EvoGraphDice [**CBL12a**]) and from data mining (the work we are currently conducting with neuroscientists on progressive similarity search guarantees). In the visualization field we have a long relationship with colleagues from Machine Learning, Vision science and the Humanities that have inspired and driven much work over the years (as evidenced by the many workshops and other events cross-pollinating these domains with visualization, like Vis×Vision, Vis×AI, VIS4DH - Visualization for the Digital Humanities, VizSec -Visualization for Cyber Security).

I have also been inspired by previews work and applied it to *new use contexts*, like our work on multivariate graph visualization [**BCD+10**], and on GeneaQuilts with genealogists [**BDF+10**]. This was sometimes done in unexpected ways, for example when my work on mobile shortcuts inspired SketchSliders [**TBJ15**]. Other researchers have very likely been inspired by previous works and their trade-offs in order to create *hybrid approaches*, like the work on NodeTrix (node link + matrix) graph visualization [HFM07].

Another inspiration may be *our own needs* as researchers, like our work developing the Smarties toolkit [**CBF14**], and trying to deal with limits of technology at our disposal (likely this is the reason behind work investigating missing data visualization, visualization under latency or uncertainty, etc.). Or of thinking of the larger societal *impact* of visualization (like the recent work on visualization ethics [Cor19]).

Other inspiration can be more serendipitous. *Art* is one. Several years back an XKCD comic[2] sparked a series of work on timeline visualizations [TM12, LWW+13]. Or it can be the artistic approach itself. Recently, the Dear Data book [LP16] by information designers Lupi and Posavec inspired several work on custom-created and personalized visualizations, like work on sketched and personalized visualizations [XHRC+18, KIHR+19], as well as our own recent work on using pictures as elements for creating visualizations [**ZSBC20**]). Another inspiration is popular media, for example our work on hybrid-image visualizations [**IDW+13**] was originally motivated by an article in a science magazine.

This is by no means a complete list. And while several papers clearly acknowledge the inspiration behind them, this is not always systematically done when writing a research article, as space may be at a premium, or because as researchers we need to present a coherent story of our work. Moreover, our papers rarely capture the iteration of our ideas and techniques. Domains interested in designing solutions for wicked problems [KR70] have developed methodologies for capturing their process of design rationales [Lee97], i.e., their reasons, justifications, alternatives considered, and trade-offs evaluated. To a lesser degree, HCI has a tradition of following iterative design process and HCI papers often refer to previous iterations of tools and of pilot results. I feel that in much of the visualization literature this information is lost, but could be important for other researchers in the field (to teach or reuse). The research replication crisis has prompted efforts in our research community (among other domains) to adopt open science practices that can aid replication and reproduction (pre-registration of studies, sharing of code, data and analysis scripts and results). It may be time that as HCI and visualization researchers we need to consider

[2] XKCD Movie Narrative Charts `https://xkcd.com/657/`

if we should more systematically also share our inspirations and design steps and outcomes. In the following years I plan to investigate how different members of our research community view and keep track of information related to inspiration and iteration of of the work, if they can see value in sharing and communicating it to others, as well as what are methodologies we could adapt from other domains (like design rationale) that would work with the practices and constraints of our field.

4.5 Closing Remarks

The list of my past and future work consists of a diverse body of work. Some have common characteristics and an underline theme (use of collaborative technology, understanding of decision making), and others vary in topic and methodology and may seem disconnected. While there are indeed topics that are constantly motivating my research (e.g., use of new technology in visualization), my work inspiration and focus is often serendipitous: I work on problems that I find interesting and I can find interest in many topics. This is in part due to the freedom provided by an academic career, but it is also in large part due to my training. I've found the fact that I was trained in HCI to be an extremely valuable attribute when conducting visualization research, as I can approach different research questions applying very different methodologies. I am extremely grateful to the University of Toronto for its exceptional HCI curriculum.

My HCI training has provided me with a big suite of methodological tools that I can rely on when looking at a new problem, ranging from techniques to break down problems (design of controlled experiments and isolating factors), to understanding user needs (user observations, interviews, contextual inquiries), to design (participatory design, iterative prototyping), to evaluate and validate (informal feedback, observational studies, controlled lab experiments), and to analyze the results (both with qualitative and quantitative methods). This is in part the reason that I have been able to look at very different research topics and approach them using very different methodologies that I adapt to the questions at hand. I was of course not an expert in all methodologies when finishing my PhD degree, but the HCI training and methodological diversity of the HCI community, helped me be open to many different ways of conducting research, and the drawbacks and benefits of each. This type of training is something I believe we need to encourage as a field.

This view of research from both a visualization and HCI perspective has also served as inspiration for solutions proposed in my work. By keeping informed as best I can of the literature across the two communities, I have been influenced in the types of solutions I propose, that often combine inspiration from the two domains (e.g., use sketching for UI design in a visualization). Finally, it has blessed me with my many amazing colleagues, with their different points of view and expertise. The work presented here has only been possible because of my students, and colleagues from industry and academia, who have worked with me across years and countries. I cannot thank them enough.

And while a very broad set of interests makes writing a manuscripts such as this one challenging, I feel it is the reason why visualization and HCI researchers (myself included) love our job. I was recently asked what I like about being a researcher, and my response was that what I like is being an HCI and Visualization researcher in particular. Our research community is extremely fortunate to have access to a staggering variety of topics to work on. Data exists in every aspect of life and we can choose within our profession to learn more, become immersed in, and have an impact on an infinite number of other domains, from neuroscience, genealogy and economics, to vision science and cognition. Sometimes it takes a retrospective career document to remind us of the beauty of our domain.

Bibliography

[AB18] A. Adadi and M. Berrada. Peeking inside the black-box: A survey on explainable artificial intelligence (xai). *IEEE Access*, 6:52138–52160, 2018.

[ABK+07] Sören Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. Dbpedia: A nucleus for a web of open data. In *Proceedings of the 6th International The Semantic Web and 2Nd Asian Conference on Asian Semantic Web Conference*, ISWC'07/ASWC'07, pages 722–735, Berlin, Heidelberg, 2007. Springer-Verlag.

[ACG14] Danielle Albers, Michael Correll, and Michael Gleicher. Task-driven evaluation of aggregation in time series visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 551–560, New York, NY, USA, 2014. ACM.

[ACJM03] David Auber, Yves Chiricota, Fabien Jourdan, and Guy Melançon. Multiscale visualization of small world networks. In *Proceedings of the Ninth Annual IEEE Conference on Information Visualization*, INFOVIS'03, pages 75–81, USA, 2003. IEEE Computer Society.

[AES05] Robert Amar, James Eagan, and John Stasko. Low-level components of analytic activity in information visualization. In *Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization*, INFOVIS '05, page 15, USA, 2005. IEEE Computer Society.

[AEYN11] Christopher Andrews, Alex Endert, Beth Yost, and Chris North. Information visualization on large, high-resolution displays: Issues, challenges, and opportunities. *Information Visualization*, 10(4):341–355, October 2011.

[AHW10] Simon Attfield, Sukhvinder Hara, and B.L. Wong. Sensemaking in visual analytics: Processes and challenges. In *EuroVAST 2010: The 1st European Symposium on Visual Analytics Science and Technology*, 2010.

[AJI10] Fouad Alallah, Dean Jin, and Pourang Irani. Oa-graphs: Orientation agnostic graphs for improving the legibility of charts on horizontal displays. In *ACM International Conference on Interactive Tabletops and Surfaces*, ITS '10, pages 211–220, New York, NY, USA, 2010. ACM.

[AKM+10] Thomas Augsten, Konstantin Kaefer, René Meusel, Caroline Fetzer, Dorian Kanitz, Thomas Stoff, Torsten Becker, Christian Holz, and Patrick Baudisch. Multitoe: High-precision interaction with back-projected floors based on high-resolution multi-touch input. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology*, UIST '10, pages 209–218, New York, NY, USA, 2010. Association for Computing Machinery.

[Ame12] *Regroup: Interactive Machine Learning for On-demand Group Creation in Social Networks*, CHI '12, New York, NY, USA, 2012. ACM.

[AMJ+12] S. Afzal, R. Maciejewski, Yun Jang, N. Elmqvist, and D.S. Ebert. Spatial Text Visualization Using Automatic Typographic Maps. *Transactions on Visualization and Computer Graphics*, 18(12):2556–2564, 2012.

[AMST11] Wolfgang Aigner, Silvia Miksch, Heidrun Schumann, and Christian Tominski. *Visualization of Time-Oriented Data*. Springer Publishing Company, Incorporated, 2011.

[Ans73] F.J. Anscombe. Graphs in statistical analysis. *The American Statistician*, 1(27):17–21, 1973.

[ARC+17] Basak Alper, Nathalie Henry Riche, Fanny Chevalier, Jeremy Boy, and Metin Sezgin. Visualization literacy at elementary school. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 5485–5497, New York, NY, USA, 2017. Association for Computing Machinery.

[Ark91] Hal R Arkes. Costs and benefits of judgment errors: Implications for debiasing. *Psychological bulletin*, 110(3):486, 1991.

[ARV+12] Alec Azad, Jaime Ruiz, Daniel Vogel, Mark Hancock, and Edward Lank. Territoriality and behaviour on and around large vertical publicly-shared displays. In *Proceedings of the Designing Interactive Systems Conference*, DIS '12, pages 468–477, New York, NY, USA, 2012. Association for Computing Machinery.

[AS94] Christopher Ahlberg and Ben Shneiderman. Visual information seeking: Tight coupling of dynamic query filters with starfield displays. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 313–317. ACM, 1994.

[AS04] Robert Amar and John Stasko. A knowledge task-based framework for design and evaluation of information visualizations. In *Proceedings of the IEEE Symposium on Information Visualization*, INFOVIS '04, pages 143–150, Washington, DC, USA, 2004. IEEE Computer Society.

[AS07] Aleks Aris and Ben Shneiderman. Designing semantic substrates for visual network exploration. *Information Visualization*, 6(4):281–300, 2007.

[Asi85] Daniel Asimov. The grand tour: a tool for viewing multidimensional data. *SIAM J. Sci. Stat. Comput.*, 6(1):128–143, January 1985.

[ASSS18] Marco Angelini, Giuseppe Santucci, Heidrun Schumann, and Hans-Jörg Schulz. A review and characterization of progressive visual analytics. *Informatics*, 5:31, 2018.

[ATC+15] Li An, Ming-Hsiang Tsou, Stephen ES Crook, Yongwan Chun, Brian Spitzberg, J Mark Gawron, and Dipak K Gupta. Space–time analysis: Concepts, quantitative methods, and future directions. *Annals of the Association of American Geographers*, 105(5):891–914, 2015.

[AWS92] Christopher Ahlberg, Christopher Williamson, and Ben Shneiderman. Dynamic queries for information exploration: an implementation and evaluation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '92, pages 619–626, New York, NY, USA, 1992. ACM.

[AWV+19] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, and et al. Guidelines for human-ai interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, New York, NY, USA, 2019. Association for Computing Machinery.

[Ban97] Wolfgang Banzhaf. *Handbook of Evolutionary Computation*, chapter Interactive Evolution. Oxford University Press, 1997.

[BB99] Benjamin B. Bederson and Angela Boltman. Does animation help users build mental maps of spatial information? In *Proceedings of the IEEE Symposium on Information Visualization*, pages 28–35, 1999.

[BB05a] Anastasia Bezerianos and Ravin Balakrishnan. Canvas portals: View and space management on large displays. *IEEE Computer Graphics and Applications*, 25:34–43, 2005.

[BB05b] Anastasia Bezerianos and Ravin Balakrishnan. The vacuum: facilitating the manipulation of distant objects. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '05, pages 361–370, New York, NY, USA, 2005. ACM.

[BBB+18] Tanja Blascheck, Anastasia Bezerianos, Lonni Besançon, Bongshin Lee, and Petra Isenberg. Preparing for Perceptual Studies: Position and Orientation of Wrist-worn Smartwatches for Reading Tasks. In *Proceedings of the Workshop on Data Visualization on Mobile Devices held at ACM CHI*, Montréal, Canada, April 2018.

[BBB+19] T. Blascheck, L. Besan con, A. Bezerianos, B. Lee, and P. Isenberg. Glanceable visualization: Studies of data comparison performance on smartwatches. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):630–640, Jan 2019.

[BBC] BBC. The route masters: Running london's roads: Gridlock. https://vimeo.com/137839650. Accessed: 2017-01-17.

[BBIF12] N. Boukhelifa, A. Bezerianos, T. Isenberg, and J. Fekete. Evaluating sketchiness as a visual variable for the depiction of qualitative uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2769–2778, Dec 2012.

[BBL18] Nadia Boukhelifa, Anastasia Bezerianos, and Evelyne Lutton. Evaluation of Interactive Machine Learning Systems. In J. Zhou and F. Chen, editors, *Human and Machine LearningVisible, Explainable, Trustworthy and Transparent*, pages 341–360. Springer, July 2018.

[BBT+19] Nadia Boukhelifa, Anastasia Bezerianos, Ioan Cristian Trelea, Nathalie Méjean Perrot, and Evelyne Lutton. An exploratory study on visual exploration of model simulations by multiple types of experts. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, New York, NY, USA, 2019. Association for Computing Machinery.

[BBTL15] Nadia Boukhelifa, Anastasia Bezerianos, Waldo Cancino Ticona, and Evelyne Lutton. Evolutionary visual exploration: Evaluation of an iec framework for guided visual search. *Evolutionary Computation, preprint*, 2015.

[BC94] Donald J. Berndt and James Clifford. Using dynamic time warping to find patterns in time series. In *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*, AAAIWS'94, pages 359–370. AAAI Press, 1994.

[BC08] Jeanette Bautista and Giuseppe Carenini. An empirical evaluation of interactive visualizations for preferential choice. In *Proceedings of the working conference on Advanced visual interfaces*, pages 207–214. ACM, 2008.

[BCB08] E. A. Bier, S. K. Card, and J. W. Bodnar. Entity-based collaboration tools for intelligence analysis. In *2008 IEEE Symposium on Visual Analytics Science and Technology*, pages 99–106, Oct 2008.

[BCBM18] Mark Billinghurst, Maxime Cordeil, Anastasia Bezerianos, and Todd Margolis. *Collaborative Immersive Analytics*, pages 221–257. Springer International Publishing, Cham, 2018.

[BCC+05] L. Bavoil, S. P. Callahan, P. J. Crossno, J. Freire, C. E. Scheidegger, C. T. Silva, and H. T. Vo. Vistrails: enabling interactive multiple-view visualizations. In *VIS 05. IEEE Visualization, 2005.*, pages 135–142, Oct 2005.

[BCD+10] Anastasia Bezerianos, Fanny Chevalier, Pierre Dragicevic, Niklas Elmqvist, and Jean-Daniel Fekete. GraphDice: A System for Exploring Multivariate Social Networks. *Computer Graphics Forum - Eurographics/IEEE-VGTC Symposium on Visualization 2010 (EuroVis 2010)*, 10(3):863–872, 2010.

[BCN+20] L. Battle, R. J. Crouser, A. Nakeshimana, A. Montoly, R. Chang, and M. Stonebraker. The role of latency and task complexity in predicting visual search behavior. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):1246–1255, Jan 2020.

[BD09] Nadia Boukhelifa and David John Duke. Uncertainty Visualization – Why Might it Fail? In *CHI Extended Abstracts*, pages 4051–4056, New York, 2009. ACM.

[BDB06] Anastasia Bezerianos, Pierre Dragicevic, and Ravin Balakrishnan. Mnemonic rendering: an image-based approach for exposing hidden changes in dynamic displays. In *Proceedings of the 19th annual ACM symposium on User interface software and technology*, UIST '06, pages 159–168, New York, NY, USA, 2006. ACM.

[BDBFG07] S. Battiato, G. Di Blasi, G. M. Farinella, and G. Gallo. Digital Mosaic Frameworks – An Overview. *Computer Graphics Forum*, 26(4):794–812, 2007.

[BDF+10] Anastasia Bezerianos, Pierre Dragicevic, Jean-Daniel Fekete, Juhee Bae, and Ben Watson. GeneaQuilts: A System for Exploring Large Genealogies. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1073–1081, October 2010.

[BEF09] S. P. Borgatti, M. G. Everett, and L. C. Freeman. UCINET 6 for Windows: Software for social network analysis, 2009.

[BEF17] Sriram Karthik Badam, Niklas Elmqvist, and Jean-Daniel Fekete. Steering the craft: Ui elements and visualizations for supporting progressive visual analytics. *Comput. Graph. Forum*, 36(3):491–502, June 2017.

[Ber10] Jacques Bertin. *Semiology of Graphics: Diagrams, Networks, Maps*. ESRI Press, Redlands, California, 2010.

[Bez07] Anastasia Bezerianos. Using alternative views for layout, comparison and context switching tasks in wall displays. In *Proceedings of the 19th Australasian conference on Computer-Human Interaction: Entertaining User Interfaces*, OZCHI '07, pages 303–310, New York, NY, USA, 2007. ACM.

[BGM04] Benjamin B. Bederson, Jesse Grosjean, and Jon Meyer. Toolkit design for interactive structured graphics. *IEEE Transactions on Software Engineering*, 30(8):535–546, 2004.

[BI12] Anastasia Bezerianos and Petra Isenberg. Perception of visual variables on tiled wall-sized displays for information visualization applications. *IEEE Transactions on Visualization and Computer Graphics - InfoVis*, 18(12):2516–2525, 2012.

[BKSS14] M. Behrisch, F. Korkmaz, L. Shao, and T. Schreck. Feedback-driven interactive exploration of large multidimensional data supported by visual classifier. In *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 43–52, Oct 2014.

[BL09a] Enrico Bertini and Denis Lalanne. Surveying the complementary role of automatic data analysis and visualization in knowledge discovery. In *Proceedings of the ACM SIGKDD Workshop on Visual Analytics and Knowledge Discovery: Integrating Automated Analysis with Interactive Exploration*, VAKD '09, pages 12–20, New York, NY, USA, 2009. Association for Computing Machinery.

[BL09b] Enrico Bertini and Denis Lalanne. Surveying the complementary role of automatic data analysis and visualization in knowledge discovery. In *Proceedings of the ACM SIGKDD Workshop on Visual Analytics and Knowledge Discovery: Integrating Automated Analysis with Interactive Exploration*, pages 12–20. ACM, 2009.

[BLBC12] E. Brown, J. Liu, C. Brodley, and R. Chang. Dis-function: Learning distance functions interactively. In *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 83–92, Oct 2012.

[BLC+11] Jeffrey Browne, Bongshin Lee, Sheelagh Carpendale, Nathalie Riche, and Timothy Sherwood. Data analysis on interactive whiteboards through sketch-based interaction. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '11, pages 154–157, New York, NY, USA, 2011. ACM.

[BLHN+12] M. Beaudouin-Lafon, S. Huot, M. Nancel, W. Mackay, E. Pietriga, R. Primet, J. Wagner, O. Chapuis, C. Pillias, J.R. Eagan, T. Gjerlufsen, and C. Klokmose. Multisurface interaction in the wild room. *Computer*, 45(4):48–56, 2012.

[BLIC19] M. Brehmer, B. Lee, P. Isenberg, and E. K. Choe. Visualizing ranges over time on mobile phones: A task-based crowdsourced evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):619–629, Jan 2019.

[BM13] M. Brehmer and T. Munzner. A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2376–2385, Dec 2013.

[BMG+10] Scott Bateman, Regan L. Mandryk, Carl Gutwin, Aaron Genest, David McDine, and Christopher Brooks. Useful junk? the effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2573–2582, New York, NY, USA, 2010. Association for Computing Machinery.

[BNB07] Robert Ball, Chris North, and Doug A. Bowman. Move to improve: Promoting physical navigation to increase user performance with large displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 191–200, New York, NY, USA, 2007. Association for Computing Machinery.

[BOH11] M. Bostock, V. Ogievetsky, and J. Heer. D3 data-driven documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2301–2309, Dec 2011.

[BOP+14] Christophe Bortolaso, Matthew Oskamp, Greg Phillips, Carl Gutwin, and T.C. Nicholas Graham. The effect of view techniques on collaboration and awareness in tabletop map-based tasks. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*, ITS '14, pages 79–88, New York, NY, USA, 2014. ACM.

[BPF14] Benjamin Bach, Emmanuel Pietriga, and Jean-Daniel Fekete. Graphdiaries: Animated transitions andtemporal navigation for dynamic networks. *IEEE Transactions on Visualization and Computer Graphics*, 20(5):740–754, May 2014.

[BRBF14] J. Boy, R. A. Rensink, E. Bertini, and J.-D. Fekete. A principled way of assessing visualization literacy. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1963–1972, Dec 2014.

[BSM04] Ragnar Bade, Stefan Schlechtweg, and Silvia Miksch. Connecting time-oriented data and information to a coherent interactive visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pages 105–112, New York, NY, USA, 2004. Association for Computing Machinery.

[BSP+93] Eric A. Bier, Maureen C. Stone, Ken Pier, William Buxton, and Tony D. DeRose. Toolglass and magic lenses: The see-through interface. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '93, pages 73–80, New York, NY, USA, 1993. ACM.

[BTBL13] Nadia Boukhelifa, Waldo Cancino Ticona, Anastasia Bezerianos, and Evelyne Lutton. Evolutionary visual exploration: Evaluation with expert users. *Comput. Graph. Forum*, 32(3):31–40, 2013.

[BTK11] Enrico Bertini, Andrada Tatu, and Daniel Keim. Quality metrics in high-dimensional data visualization: An overview and systematization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2203–2212, December 2011.

[CAFG12] Michael Correll, Danielle Albers, Steven Franconeri, and Michael Gleicher. Comparing averages in time series data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 1095–1104, New York, NY, USA, 2012. ACM.

[CAK07] A. Collins, T. Apted, and J. Kay. Tabletop file system access: Associative and hierarchical approaches. In *Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'07)*, pages 113–120, Oct 2007.

[Car99] M. Sheelagh T. Carpendale. *A Framework for Elastic Presentation Space*. PhD thesis, Simon Fraser University, March 1999.

[Car03] Stuart Card. The human-computer interaction handbook. chapter Information Visualization, pages 544–582. L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 2003.

[Car08] Sheelagh Carpendale. Evaluating information visualizations. In Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, and Chris North, editors, *Information Visualization*, pages 19–45. Springer-Verlag, Berlin, Heidelberg, 2008.

[CASM16] Edwin Chan, Craig Anslow, Teddy Seyed, and Frank Maurer. *Envisioning the Emergency Operations Centre of the Future*, pages 349–372. Springer, Cham, 2016.

[CBF14] **Olivier Chapuis, Anastasia Bezerianos, and Stelios Frantzeskakis. Smarties: An input system for wall display development. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 2763–2772, New York, NY, USA, 2014. ACM.**

[CBL12a] W. Cancino, N. Boukhelifa, and E. Lutton. Evographdice: Interactive evolution for visual analytics. In *2012 IEEE Congress on Evolutionary Computation*, pages 1–8, June 2012.

[CBL12b] W. Cancino, N. Boukhelifa, and E. Lutton. Evographdice: Interactive evolution for visual analytics. In *IEEE Congress on Evolutionary Computation*, pages 1–8. IEEE, 2012.

[CBM+09] **Anthony Collins, Anastasia Bezerianos, Gregor McEwan, Markus Rittenbruch, Rainer Wasinger, and Judy Kay. Understanding file access mechanisms for embedded ubicomp collaboration interfaces. In *Proceedings of the 11th international conference on Ubiquitous computing*, Ubicomp '09, pages 135–144, New York, NY, USA, 2009. ACM.**

[CCD+17] Maxime Cordeil, Andrew Cunningham, Tim Dwyer, Bruce H. Thomas, and Kim Marriott. Imaxes: Immersive axes as embodied affordances for interactive multivariate data visualisation. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, pages 71–83, New York, NY, USA, 2017. Association for Computing Machinery.

[CCH+14] Cristina Conati, Giuseppe Carenini, Enamul Hoque, Ben Steichen, and Dereck Toker. Evaluating the impact of user characteristics and different layouts on an interactive visualization for decision making. In *Computer Graphics Forum*, volume 33, pages 371–380. Wiley Online Library, 2014.

[CDBF10] **Fanny Chevalier, Pierre Dragicevic, Anastasia Bezerianos, and Jean-Daniel Fekete. Using text animated transitions to support navigation in document histories. In *Proceedings of the 28th international conference on Human factors in computing systems*, CHI '10, pages 683–692, New York, NY, USA, 2010. ACM.**

[CDK+17] M. Cordeil, T. Dwyer, K. Klein, B. Laha, K. Marriott, and B. H. Thomas. Immersive collaborative analysis of network connectivity: Cave-style or head-mounted display? *IEEE Transactions on Visualization and Computer Graphics*, 23(1):441–450, Jan 2017.

[CG09] Wen-Huang Cheng and David Gotz. Context-based page unit recommendation for web-based sensemaking tasks. In *Proceedings of the 14th International Conference on Intelligent User Interfaces*, IUI '09, pages 107–116, New York, NY, USA, 2009. Association for Computing Machinery.

[CG15] Paolo Crosetto and Alexia Gaudeul. Testing the strength and robustness of the attraction effect in consumer decision making. 2015.

[CG16] Michael Correll and Michael Gleicher. The semantics of sketch: Flexibility in visual query systems for time series data. In *2016 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 131–140, Oct 2016.

[Che73] Herman Chernoff. The Use of Faces to Represent Points in K-Dimensional Space Graphically. *Journal of the American Statistical Association*, pages 361–368, 1973.

[CHP04] Roland Chrobok, Sigurdur F. Hafstein, and Andreas Pottmeier. Olsim: A new generation of traffic information systems. *Forschung und wissenschaftliches Rechnen, GWDG-Berichte*, Nr. 63:11–25, 2004.

[CL04] Giuseppe Carenini and John Loyd. Valuecharts: analyzing linear models expressing preferences and evaluations. In *Proceedings of the working conference on Advanced visual interfaces*, pages 150–157. ACM, 2004.

[Cle85] William S. Cleveland. *The Elements of Graphing Data*. Wadsworth Advanced Book Program, 1985.

[CM84] William S. Cleveland and Robert McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, 79(387):531–554, 1984.

[CMS99] Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman, editors. *Readings in information visualization: using vision to think*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999.

[CMvdP10] William O. Chao, Tamara Munzner, and Michael van de Panne. Napkinvis: Rapid pen-centric authoring of improvisational visualizations. *Posters Compendium InfoVis*, 2010.

[Coh94] Jacob Cohen. The Earth is round (p<.05). *American psychologist*, 49(12):997, 1994.

[Con93] Sharolyn Converse. Shared mental models in expert team decision making. In Jr. N. J. Castellan, editor, *Individual and group decision making: Current issues*, pages 221–246. Hillsdale, 1993.

[Cor19] Michael Correll. Ethical dimensions of visualization research. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, New York, NY, USA, 2019. Association for Computing Machinery.

[CPB14] Segolene Charaudeau, Khashayar Pakdaman, and Pierre-Yves Boëlle. Commuter mobility and the spread of infectious diseases: Application to influenza in france. *PLOS ONE*, 9(1):1–9, 01 2014.

[CR68] F W Campbell and J G Robson. Application of Fourier Analysis to the Visibility of Gratings. *The Journal of Physiology*, 197(3):551–566, 1968.

[CR00] Andrej Cedilnik and Penny Rheingans. Procedural Annotation of Uncertainty Information. In *Proc. IEEE Visualization*, pages 77–84, Los Alamitos, 2000. IEEE Computer Society.

[CRMH12] Jason Chuang, Daniel Ramage, Christopher Manning, and Jeffrey Heer. Interpretation and trust: Designing model-driven visualizations for text analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 443–452, New York, NY, USA, 2012. Association for Computing Machinery.

[CS99] W Keith Campbell and Constantine Sedikides. Self-threat magnifies the self-serving bias. *Review of General Psychology*, 3(1):23–43, 1999.

[CU93a] Bay-Wei Chang and David Ungar. Animation: From cartoons to the user interface. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, pages 45–55, 1993.

[CU93b] Bay-Wei Chang and David Ungar. Animation: From cartoons to the user interface. In *Proceedings of the 6th Annual ACM Symposium on User Interface Software and Technology*, UIST '93, pages 45–55, New York, NY, USA, 1993. Association for Computing Machinery.

[CVW09] C. Collins, F. B. Viegas, and M. Wattenberg. Parallel tag clouds to explore and analyze faceted text corpora. In *2009 IEEE Symposium on Visual Analytics Science and Technology*, pages 91–98, Oct 2009.

[DB92] Paul Dourish and Victoria Bellotti. Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work*, CSCW '92, pages 107–114, New York, NY, USA, 1992. Association for Computing Machinery.

[DBBF19] E. Dimara, G. Bailly, A. Bezerianos, and S. Franconeri. Mitigating the attraction effect with visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):850–860, Jan 2019.

[DBD14] Evanthia Dimara, Anastasia Bezerianos, and Pierre Dragicevic. Accounting for availability biases in information visualization. *DECISIVe: Workshop on Dealing with Cognitive Biases in Visualisations / Ellis, Geoffrey (ed.). In IEEE VIS 2014*, 2014.

[DBD16] Evanthia Dimara, Anastasia Bezerianos, and Pierre Dragicevic. Testing the Attraction Effect on Two Information Visualization Datasets. Research Report RR-8895, Inria Saclay Ile de France, March 2016.

[DBD17a] E. Dimara, A. Bezerianos, and P. Dragicevic. The attraction effect in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):471–480, Jan 2017.

[DBD17b] Evanthia Dimara, Anastasia Bezerianos, and Pierre Dragicevic. Narratives in crowdsourced evaluation of visualizations: A double-edged sword? In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 5475–5484, New York, NY, USA, 2017. Association for Computing Machinery.

[DBD18] E. Dimara, A. Bezerianos, and P. Dragicevic. Conceptual and methodological issues in evaluating multidimensional visualizations for decision support. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):749–759, Jan 2018.

[DBJ+11] Pierre Dragicevic, Anastasia Bezerianos, Waqas Javed, Niklas Elmqvist, and Jean-Daniel Fekete. Temporal Distortion for Animated Transitions. In *CHI 11: International Conference on Human Factors in Computing Systems*, Vancouver, Canada, May 2011. ACM.

[DCR17] Gytis Dudas, Luiz Max Carvalho, ..., and Andrew Rambaut. Virus genomes reveal factors that spread and sustained the ebola epidemic. *Nature*, 544(7650):309, 2017.

[DD19] V. Dibia and C. Demiralp. Data2vis: Automatic generation of data visualizations using sequence-to-sequence recurrent neural networks. *IEEE Computer Graphics and Applications*, 39(5):33–46, Sep. 2019.

[DFAB03] Alan Dix, Janet Finlay, Gregory D. Abowd, and Russell Beale. *Human-Computer Interaction*. Pearson Education, 2003.

[DFP+20] E. Dimara, S. Franconeri, C. Plaisant, A. Bezerianos, and P. Dragicevic. A task-based taxonomy of cognitive biases for information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 26(2):1413–1432, Feb 2020.

[DHMM13] T. Dwyer, N. Henry Riche, K. Marriott, and C. Mears. Edge compression techniques for visualization of dense directed graphs. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2596–2605, Dec 2013.

[dNMB05] Wouter de Nooy, Andrej Mrvar, and Vladimir Batagelj. *Exploratory Social Network Analysis with Pajek*. Structural Analysis in the Social Sciences. Cambridge University Press, March 2005.

[DP20] E. Dimara and C. Perin. What is interaction for data visualization? *IEEE Transactions on Visualization and Computer Graphics*, 26(1):119–129, Jan 2020.

[DWL08] Somayeh Dodge, Robert Weibel, and Anna-Katharina Lautenschütz. Towards a taxonomy of movement patterns. *Information visualization*, 7(3-4):240–252, 2008.

[EAB13] Micheline Elias, Marie-Aude Aufaure, and Anastasia Bezerianos. Storytelling in visual analytics tools for business intelligence. In *INTERACT (3)*, pages 280–297, 2013.

[EALN11] Alex Endert, Christopher Andrews, Yueh Hua Lee, and Chris North. Visual encodings that support physical navigation on large displays. In *Proceedings of Graphics Interface 2011*, GI '11, pages 103–110, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 2011. Canadian Human-Computer Communications Society.

[EB11] Micheline Elias and Anastasia Bezerianos. Exploration views: understanding dashboard creation and customization for visualization novices. In *Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction - Volume Part IV*, INTERACT'11, pages 274–291, Berlin, Heidelberg, 2011. Springer-Verlag.

[EB12] Micheline Elias and Anastasia Bezerianos. Annotating bi visualization dashboards: Needs & challenges. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 1641–1650, New York, NY, USA, 2012. ACM.

[ECD14] Serge Egelman, Ed H Chi, and Steven Dow. Crowdsourcing in hci research. In *Ways of Knowing in HCI*, pages 267–289. Springer, 2014.

[EDF08] Niklas Elmqvist, Pierre Dragicevic, and Jean-Daniel Fekete. Rolling the Dice: Multidimensional Visual Exploration using Scatterplot Matrix Navigation. *Visualization and Computer Graphics*, 14(6):1–8, August 2008.

[EKHW07] R. Eccles, T. Kapler, R. Harper, and W. Wright. Stories in geotime. In *2007 IEEE Symposium on Visual Analytics Science and Technology*, pages 19–26, Oct 2007.

[Elm13] Martin E Elmer. Symbol considerations for bivariate thematic maps. *In Proceedings of the 26th International Cartography Conference (ICC)*, 2013.

[EZ93] James Elder and Steven Zucker. The effect of contour closure on the rapid discrimination of two-dimensional shapes. *Vision Res*, 1993.

[EZ15] Philipp Eichmann and Emanuel Zgraggen. Evaluating subjective accuracy in time series pattern-matching using human-annotated rankings. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*, IUI '15, pages 28–37, New York, NY, USA, 2015. ACM.

[FBDM+15] Audrey Fertier, Anne-Marie Barthe-Delanoë, Johan Manceau, Sébastien Truptil, and Frédérick Bénaben. Towards a support to stakeholders' collaboration during a loire river major flooding. In Narjès Bellamine Ben Saoud, Carole Adam, and Chihab Hanachi, editors, *Information Systems for Crisis Response and Management in Mediterranean Countries*, pages 183–193, Cham, 2015. Springer International Publishing.

[Fek04] J.-D. Fekete. The infovis toolkit. In *IEEE Symposium on Information Visualization*, pages 167–174, Oct 2004.

[Few06] Stephen Few. *Information Dashboard Design: The Effective Visual Communication of Data*. O'Reilly Media, Inc., 2006.

[Few08] Stephen Few. Dual-Scaled Axes in Graphs—Are They Ever the Best Solution? In *Visual Business Intelligence Newsletter*. Perceptual Edge, March 2008.

[FIB+14] Johannes Fuchs, Petra Isenberg, Anastasia Bezerianos, Fabian Fischer, and Enrico Bertini. The influence of contour on similarity perception of star glyphs. *IEEE Trans. Vis. Comput. Graph.*, 20(12):2251–2260, 2014.

[FIBK17] J. Fuchs, P. Isenberg, A. Bezerianos, and D. Keim. A systematic review of experimental studies on data glyphs. *IEEE Transactions on Visualization and Computer Graphics*, 23(7):1863–1879, July 2017.

[Fis82] B. Fischhoff. Debiasing/kahneman, d., slovic, p. and tversky, a. In Daniel Kahneman, Paul Slovic, and Amos Tversky, editors, *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press, 1982.

[FKN86] Geoffrey T Fong, David H Krantz, and Richard E Nisbett. The effects of statistical training on thinking about everyday problems. *Cognitive psychology*, 18(3):253–292, 1986.

[FL03] M. C. Ferreira de Oliveira and H. Levkowitz. From visual data exploration to visual data mining: a survey. *IEEE Transactions on Visualization and Computer Graphics*, 9(3):378–394, July 2003.

[FP99] Jean-Daniel Fekete and Catherine Plaisant. Excentric labeling: Dynamic neighborhood labeling for data visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, pages 512–519, New York, NY, USA, 1999. ACM.

[FP02] Jean-Daniel Fekete and Catherine Plaisant. Interactive information visualization of a million items. In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis'02)*, INFOVIS '02, pages 117–, Washington, DC, USA, 2002. IEEE Computer Society.

[FP16] Jean-Daniel Fekete and Romain Primet. Progressive analytics: A computation paradigm for exploratory data analysis. *CoRR*, abs/1607.05162, 2016.

[FP19] Adrien Fonnet and Yannick Prié. Survey of Immersive Analytics. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–22, 2019.

[FPV+13] N. Ferreira, J. Poco, H. T. Vo, J. Freire, and C. T. Silva. Visual exploration of big spatio-temporal urban data: A study of new york city taxi trips. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2149–2158, Dec 2013.

[Fri87] J. H. Friedman. Exploratory projection pursuit. *J. Am. Statistical Assoc*, 82(397):249–266, 1987.

[FRM94] Christos Faloutsos, M. Ranganathan, and Yannis Manolopoulos. Fast subsequence matching in time-series databases. In *Proceedings of the 1994 ACM SIGMOD International Conference on Management of Data*, SIGMOD '94, pages 419–429, New York, NY, USA, 1994. ACM.

[FSME14] S. Frey, F. Sadlo, K. Ma, and T. Ertl. Interactive progressive visualization with space-time error control. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2397–2406, Dec 2014.

[Gap06] GapMinder, 2006. http://www.gapminder.org/.

[GFC05] Mohammad Ghoniem, Jean-Daniel Fekete, and Philippe Castagliola. On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis. *Information Visualization*, 4(2):114–135, 2005.

[GG95] Carl Gutwin and Saul Greenberg. Support for group awareness in real-time desktop conferences. In *Proceedings of The Second New Zealand Computer Science Research Students' Conference*, pages 18–21. University of Calgary, 1995.

[GG98a] Stephen J. Guastello and Denise D. Guastello. Origins of coordination and team effectiveness: A perspective from game theory and nonlinear dynamics. *Journal of Applied Psychology*, pages 423–437, 1998.

[GG98b] Carl Gutwin and Saul Greenberg. Design for individuals, design for groups: Tradeoffs between power and workspace awareness. In *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work*, CSCW '98, pages 207–216, New York, NY, USA, 1998. ACM.

[GG02] Carl Gutwin and Saul Greenberg. A descriptive framework of workspace awareness for real-time groupware. *Comput. Supported Coop. Work*, 11(3):411–446, November 2002.

[Gig08] Gerd Gigerenzer. Why heuristics work. *Perspectives on psychological science*, 3(1):20–29, 2008.

[GLG+13] Samuel Gratzl, Alexander Lex, Nils Gehlenborg, Hanspeter Pfister, and Marc Streit. Lineup: Visual analysis of multi-attribute rankings. *IEEE transactions on visualization and computer graphics*, 19(12):2277–2286, 2013.

[GLQ18] Peichao Gao, Zhilin Li, and Zhe Qin. Usability of value-by-alpha maps compared to area cartograms and proportional symbol maps. *Journal of Spatial Science*, pages 1–17, March 2018.

[GMH+06] Amy L Griffin, Alan M MacEachren, Frank Hardisty, Erik Steiner, and Bonan Li. A comparison of animated maps with static small-multiple maps for visually identifying space-time clusters. *Annals Assoc. of American Geographers*, 96(4):740–753, 2006.

[Gol89] D. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition, 1989.

[Gol99] E. Bruce Goldstein. *Sensation and Perception*. Brooks/Cole Publishing, Pacific Grove, USA, 5th edition, 1999.

[GTPB19] A. Gogolou, T. Tsandilas, T. Palpanas, and A. Bezerianos. Comparing similarity perception in time series visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):523–533, Jan 2019.

[GTS10] L. Grammel, M. Tory, and M. Storey. How information visualization novices construct visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):943–952, Nov 2010.

[HB05] J. Heer and D. Boyd. Vizster: visualizing online social networks. In *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005.*, pages 32–39, Oct 2005.

[HB10] Jeffrey Heer and Michael Bostock. Crowdsourcing graphical perception: Using mechanical turk to assess visualization design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 203–212, New York, NY, USA, 2010. Association for Computing Machinery.

[HBED18] Tom Horak, Sriram Karthik Badam, Niklas Elmqvist, and Raimund Dachselt. When david meets goliath: Combining smartwatches with a large vertical display for visual data exploration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, New York, NY, USA, 2018. Association for Computing Machinery.

[HBF08] Nathalie Henry, Anastasia Bezerianos, and Jean-Daniel Fekete. Improving the readability of clustered social networks using node duplication. *IEEE Transactions on Visualization and Computer Graphics*, 14:1317–1324, November 2008.

[HBH+20] K. W. Hall, A. J. Bradley, U. Hinrichs, S. Huron, J. Wood, C. Collins, and S. Carpendale. Design by immersion: A transdisciplinary approach to problem-driven visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):109–118, Jan 2020.

[HC95] Timothy B. Heath and Subimal Chatterjee. Asymmetric decoy effects on lower-quality versus higher-quality brands: Meta-analytic and experimental evidence. *Journal of Consumer Research*, 22(3):268–284, 1995.

[HD11] Jessica Hullman and Nick Diakopoulos. Visualization rhetoric: Framing effects in narrative visualization. *TVCG*, 17(12):2231–2240, 2011.

[HDA13] Jessica Hullman, Nicholas Diakopoulos, and Eytan Adar. Contextifier: Automatic generation of annotated stock visualizations. In *CHI 2013*, Conference on Human Factors in Computing Systems - Proceedings, pages 2707–2716, 5 2013.

[Hee19] Jeffrey Heer. Agency plus automation: Designing artificial intelligence into interactive systems. *Proceedings of the National Academy of Sciences*, 116(6):1844–1850, 2019.

[HF09] Christian Holz and Steven Feiner. Relaxed selection techniques for querying time-series graphs. In *Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology*, UIST '09, pages 213–222, New York, NY, USA, 2009. ACM.

[HFM07] N. Henry, J.-D. Fekete, and M. J. McGuffin. Nodetrix: a hybrid visualization of social networks. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1302–1309, Nov 2007.

[HIvWF11] Danny Holten, Petra Isenberg, Jarke J. van Wijk, and Jean-Daniel Fekete. An extended evaluation of the readability of tapered, animated, and textured directed-edge representations in node-link graphs. In *Proceedings of the 2011 IEEE Pacific Visualization Symposium*, PACIFICVIS '11, pages 195–202, USA, 2011. IEEE Computer Society.

[HKF15] Steve Haroz, Robert Kosara, and Steven L. Franconeri. Isotype visualization: Working memory, performance, and engagement with pictographs. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 1191–1200, New York, NY, USA, 2015. Association for Computing Machinery.

[HKR+05] Kirstie Hawkey, Melanie Kellar, Derek Reilly, Tara Whalen, and Kori M. Inkpen. The proximity factor: Impact of distance on co-located collaboration. In *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work*, GROUP '05, pages 31–40, New York, NY, USA, 2005. Association for Computing Machinery.

[HLKW12] Julian Heinrich, Yuan Luo, Arthur E. Kirkpatrick, and Daniel Weiskopf. Evaluation of a bundling technique for parallel coordinates. *Proceedings of International Conference on Information Visualization Theory and Applications*, pages 594–602, 2012.

[HN98] Jerry L. Hintze and Ray D. Nelson. Violin plots: A box plot-density trace synergism. *The American Statistician*, 52(2):181–184, May 1998.

[Hor99] Eric Horvitz. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, pages 159–166, New York, NY, USA, 1999. Association for Computing Machinery.

[HP13] Ivan Hernandez and Jesse Lee Preston. Disfluency disrupts the confirmation bias. *Journal of Experimental Social Psychology*, 49(1):178–182, 2013.

[HPP82] Joel Huber, John W Payne, and Christopher Puto. Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, 9(1):90–98, 1982.

[HQC+19] J. Hullman, X. Qiao, M. Correll, A. Kale, and M. Kay. In pursuit of error: A survey of uncertainty visualization evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):903–913, Jan 2019.

[HR07] Jeffrey Heer and George Robertson. Animated transitions in statistical data graphics. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1240–1247, 2007.

[Hub85] P. J. Huber. Projection Pursuit. *The Annals of Statistics*, 13(2):435–475, 1985.

[Hul11] Jessica R Hullman. Not all hits are created equal: Controlling for reasoning and learning processes in mturk. In *Proc. of the ACM CHI Workshop on Crowdsourcing and Human Computation*. Citeseer, 2011.

[HvHC+08] Jeffrey Heer, Frank van Ham, Sheelagh Carpendale, Chris Weaver, and Petra Isenberg. *Creation and Collaboration: Engaging New Audiences for Information Visualization*, pages 92–133. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[HVW07] Jeffrey Heer, Fernanda B. Viégas, and Martin Wattenberg. Voyagers and voyeurs: Supporting asynchronous collaborative information visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 1029–1038, New York, NY, USA, 2007. Association for Computing Machinery.

[HvW09] Danny Holten and Jarke J. van Wijk. Force-directed edge bundling for graph visualization. In *Proceedings of the 11th Eurographics / IEEE - VGTC Conference on Visualization*, EuroVis '09, pages 983–998, Chichester, GBR, 2009. The Eurographs Association & John Wiley & Sons, Ltd.

[HYFC14a] Lane Harrison, Fumeng Yang, Steven Franconeri, and Remco Chang. Ranking visualizations of correlation using weber's law. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1943–1952, 2014.

[HYFC14b] Lane Harrison, Fumeng Yang, Steven Franconeri, and Remco Chang. Ranking Visualizations of Correlation Using Weber's Law. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1943–1952, December 2014.

[IBDF11] Petra Isenberg, Anastasia Bezerianos, Pierre Dragicevic, and Jean-Daniel Fekete. A Study on Dual-Scale Data Charts. *Transactions on Visualization and Computer Graphics*, 17(12):2469–2487, 2011.

[ICB+09] Petra Isenberg, Sheelagh Carpendale, Anastasia Bezerianos, Nathalie Henry, and Jean-Daniel Fekete. Coconuttrix: collaborative retrofitting for information visualization. *IEEE Comput. Graph. Appl.*, 29:44–57, September 2009.

[ICH+13] Alexandra Ion, Y.-L. Betty Chang, Michael Haller, Mark Hancock, and Stacey D. Scott. Canyon: Providing location awareness of multiple moving objects in a detail view on large displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 3149–3158, New York, NY, USA, 2013. ACM.

[IDW+13] Petra Isenberg, Pierre Dragicevic, Wesley Willett, Anastasia Bezerianos, and Jean-Daniel Fekete. Hybrid-Image Visualization for Large Viewing Environments. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2346–2355, 2013.

[IES+08] K.P. Indiradevi, Elizabeth Elias, P.S. Sathidevi, S. Dinesh Nayak, and K. Radhakrishnan. A multi-level wavelet approach for automatic detection of epileptic spikes in the electroencephalogram. *Computers in Biology and Medicine*, 38(7):805 – 816, 2008.

[IES+11] Petra Isenberg, Niklas Elmqvist, Jean Scholtz, Daniel Cernea, Kwan-Liu Ma, and Hans Hagen. Collaborative visualization: Definition, challenges, and research agenda. *Information Visualization*, 10(4):310–326, October 2011.

[IFP+12] P. Isenberg, D. Fisher, S. A. Paul, M. R. Morris, K. Inkpen, and M. Czerwinski. Co-located collaborative visual analytics around a tabletop display. *IEEE Transactions on Visualization and Computer Graphics*, 18(5):689–702, May 2012.

[IIH+13] Petra Isenberg, Tobias Isenberg, Tobias Hesselmann, Bongshin Lee, Ulrich von Zadow, and Anthony Tang. Data visualization on interactive surfaces: A research agenda. *IEEE Comput. Graph. Appl.*, 33(2):16–24, March 2013.

[IIS+17] P. Isenberg, T. Isenberg, M. Sedlmair, J. Chen, and T. Möller. Visualization as seen through its research paper keywords. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):771–780, Jan 2017.

[JB05] Yaochu Jin and Jürgen Branke. Evolutionary optimization in uncertain environments-a survey. *IEEE Trans. Evolutionary Computation*, 9(3):303–317, 2005.

[JDF12a] Yvonne Jansen, Pierre Dragicevic, and Jean-Daniel Fekete. Tangible remote controllers for wall-size displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 2865–2874, New York, NY, USA, 2012. ACM.

[JDF12b] Yvonne Jansen, Pierre Dragicevic, and Jean-Daniel Fekete. Tangible remote controllers for wall-size displays. In *the 2012 ACM annual conference*, page 2865, New York, New York, USA, 2012. ACM Press.

[JDF13] Yvonne Jansen, Pierre Dragicevic, and Jean-Daniel Fekete. Evaluating the efficiency of physical visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 2593–2602, New York, NY, USA, 2013. Association for Computing Machinery.

[JDR+16] J. Jing, J. Dauwels, T. Rakthanmanon, E. Keogh, S.S. Cash, and M.B. Westover. Rapid annotation of interictal epileptiform discharges via template matching under dynamic time warping. *Journal of Neuroscience Methods*, 274:179 – 190, 2016.

[JE12] Waqas Javed and Niklas Elmqvist. Exploring the design space of composite visualization. In *Proceedings of the 2012 IEEE Pacific Visualization Symposium*, PacificVis '12, pages 1–8, Washington, DC, USA, 2012. IEEE.

[JH14] Mikkel R. Jakobsen and Kasper Hornbæk. Up close and personal: Collaborative work on a high-resolution multitouch wall display. *ACM Trans. Comput.-Hum. Interact.*, 21(2), February 2014.

[JLC18] Liu Jiang, Shixia Liu, and Changjian Chen. Recent research advances on interactive machine learning. *Journal of Visualization*, pages 1–17, 2018.

[JLMV06] A. Johnson, J. Leigh, P. Morin, and P. Van Keken. Geowall: Stereoscopic visualization for geoscience research and education. *IEEE Computer Graphics and Applications*, 26(6):10–14, Nov 2006.

[JLP12] Leslie K. John, George Loewenstein, and Drazen Prelec. Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, 23(5):524–532, 2012. PMID: 22508865.

[JNJ+10] Ricardo Jota, Miguel A. Nacenta, Joaquim A. Jorge, Sheelagh Carpendale, and Saul Greenberg. A comparison of ray pointing techniques for very large displays. In *Proceedings of Graphics Interface 2010*, GI '10, pages 269–276, Toronto, Ont., Canada, Canada, 2010. Canadian Information Processing Society.

[JSH19] Yvonne Jansen, Jonas Schjerlund, and Kasper Hornbæk. Effects of locomotion and visual overview on spatial memory when interacting with wall displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, New York, NY, USA, 2019. Association for Computing Machinery.

[JT81] Ollie Johnston and Frank Thomas. *The Illusion of Life: Disney Animation*. Disney Editions, 1981.

[Kah03] Daniel Kahneman. A perspective on judgment and choice: mapping bounded rationality. *American psychologist*, 58(9):697, 2003.

[Kah11] Daniel Kahneman. *Thinking, fast and slow*. Macmillan, 2011.

[KB84] Doris H. U. Kochanek and Richard H. Bartels. Interpolating splines with local tension, continuity, and bias control. In *Computer Graphics (SIGGRAPH '84 Proceedings)*, volume 18, pages 33–41, July 1984.

[KB93] Joshua Klayman and Kaye Brown. Debias the environment instead of the judge: An alternative approach to reducing error in diagnostic (and other) judgment. *Cognition*, 49(1):97–122, 1993.

[KBGH15] Azam Khan, Simon Breslav, Michael Glueck, and Kasper Hornbæk. Benefits of visualization in the mammography problem. *International Journal of Human-Computer Studies*, 83:94–113, 2015.

[KC14] B. Kondo and C. Collins. Dimpvis: Exploring time-varying information visualizations by direct manipulation. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2003–2012, Dec 2014.

[KCS08] Aniket Kittur, Ed H Chi, and Bongwon Suh. Crowdsourcing user studies with mechanical turk. In *Proc. of the SIGCHI conference on human factors in computing systems*, pages 453–456. ACM, 2008.

[KCSG04] Russell Kruger, Sheelagh Carpendale, Stacey D. Scott, and Saul Greenberg. Roles of orientation in tabletop collaboration: Comprehension, coordination and communication. *Comput. Supported Coop. Work*, 13(5-6):501–537, December 2004.

[Kei02] Daniel A. Keim. Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):1–8, January 2002.

[Kel73] Harold H Kelley. The processes of causal attribution. *American psychologist*, 28(2):107, 1973.

[KH16a] Matthew Kay and Jeffrey Heer. Beyond weber's law: A second look at ranking visualizations of correlation. *IEEE transactions on visualization and computer graphics*, 22(1):469–478, 2016.

[KH16b] Matthew Kay and Jeffrey Heer. Beyond weber's law: A second look at ranking visualizations of correlation. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):469–478, January 2016.

[KIHR+19] Nam Wook Kim, Hyejin Im, Nathalie Henry Riche, Alicia Wang, Krzysztof Gajos, and Hanspeter Pfister. Dataselfie: Empowering people to design personalized visuals to represent their data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, New York, NY, USA, 2019. Association for Computing Machinery.

[KJEY11] Bum chul Kwon, Waqas Javed, Niklas Elmqvist, and Ji Soo Yi. Direct manipulation through surrogate objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 627–636, New York, NY, USA, 2011. Association for Computing Machinery.

[KJM+07] S. Y. Kim, Y. Jang, A. Mellema, D. S. Ebert, and T. Collinss. Visual analytics on mobile devices for emergency response. In *2007 IEEE Symposium on Visual Analytics Science and Technology*, pages 35–42, Oct 2007.

[KK96a] D. A. Keim and H. . Kriegel. Visualization techniques for mining large databases: a comparison. *IEEE Transactions on Knowledge and Data Engineering*, 8(6):923–938, Dec 1996.

[KK96b] Daniel A Keim and H-P Kriegel. Visualization techniques for mining large databases: A comparison. *IEEE Transactions on knowledge and data engineering*, 8(6):923–938, 1996.

[KK15] K. Kucher and A. Kerren. Text visualization techniques: Taxonomy, visual survey, and community insights. In *2015 IEEE Pacific Visualization Symposium (PacificVis)*, pages 117–121, April 2015.

[KKTD17] U. Kister, K. Klamka, C. Tominski, and R. Dachselt. Grasp: Combining spatially-aware mobile devices and a display wall for graph visualization and interaction. *Comput. Graph. Forum*, 36(3):503–514, June 2017.

[KL06] Robert Kincaid and Heidi Lam. Line graph explorer: Scalable display of line graphs using focus+context. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '06, pages 404–411, New York, NY, USA, 2006. Association for Computing Machinery.

[KMS+08] Daniel A. Keim, Florian Mansmann, Jörn Schneidewind, Jim Thomas, and Hartmut Ziegler. *Visual Analytics: Scope and Challenges*, pages 76–90. Springer-Verlag, Berlin, Heidelberg, 2008.

[KP10] Frederick A. A. Kingdom and Nicolaas Prins. *Psychophysics: A Practical Introduction*. Elsevier, 1st edition, 2010.

[KPRP07] G. Klein, J. K. Phillips, E. L. Rall, and D. A. Peluso. A data-frame theory of sensemaking. In In R. R. Hoffman, editor, *Expertise out of context: Proceedings of the Sixth International Conference on Naturalistic Decision Making*, pages 113–155. Lawrence Erlbaum Associates Publishers, 2007.

[KR70] W. Kunz and H.W.J. Rittel. *Issues as Elements of Information Systems*. California. University. Center for Planning and Development Research. Working paper, no. 131. Institute of Urban and Regional Development, University of California, 1970.

[KRMD15] Ulrike Kister, Patrick Reipschläger, Fabrice Matulic, and Raimund Dachselt. Bodylenses: Embodied magic lenses and personal territories for wall displays. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*, ITS '15, pages 117–126, New York, NY, USA, 2015. Association for Computing Machinery.

[KS12] Y. Kang and J. Stasko. Examining the use of a visual analytics system for sensemaking tasks: Case studies with domain experts. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2869–2878, Dec 2012.

[KS16] Robert Kosara and Drew Skau. Judgment error in pie chart variations. In *Proceedings of the Eurographics / IEEE VGTC Conference on Visualization: Short Papers*, EuroVis '16, pages 91–95, Goslar, DEU, 2016. Eurographics Association.

[KSMS12] Mohammadreza Khalilbeigi, Patrik Schmittat, Max Mühlhäuser, and Jürgen Steimle. Occlusion-aware interaction techniques for tabletop systems. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '12, pages 2531–2536, New York, NY, USA, 2012. ACM.

[LB16] Jonas Landgren and Fredrik Bergstrand. Work practice in situation rooms – an ethnographic study of emergency response work in governmental organizations. In Paloma Díaz, Narjès Bellamine Ben Saoud, Julie Dugdale, and Chihab Hanachi, editors, *Information Systems for Crisis Response and Management in Mediterranean Countries*, pages 157–171, Cham, 2016. Springer International Publishing.

[LBI+12] Heidi Lam, Enrico Bertini, Petra Isenberg, Catherine Plaisant, and Sheelagh Carpendale. Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Visualization and Computer Graphics*, 18(9):1520–1536, September 2012.

[LCBLL16] Can Liu, Olivier Chapuis, Michel Beaudouin-Lafon, and Eric Lecolinet. Shared interaction on a wall-sized display in a data manipulation task. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 2075–2086, New York, NY, USA, 2016. Association for Computing Machinery.

[LD11] Mark A Livingston and Jonathan W Decker. Evaluation of trend localization with multivariate visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2053–2062, December 2011.

[LD12] Mark A Livingston and Jonathan W Decker. Evaluation of multivariate visualizations: a case study of refinements and user experience. In *Visualization and Data Analysis 2012*, volume 8294, page 82940G. SPIE - International Society for Optics and Photonics, January 2012.

[Lee97] J. Lee. Design rationale systems: understanding the issues. *IEEE Expert*, 12(3):78–85, May 1997.

[LJ12] Patrik Lundblad and Mikael Jern. *Visual Storytelling in Education Applied to Spatial-Temporal Multivariate Statistics Data*, pages 175–193. Springer, 01 2012.

[LKD19] R. Langner, U. Kister, and R. Dachselt. Multiple coordinated views at large displays for multiple users: Empirical findings on user behavior, movements, and distances. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):608–618, Jan 2019.

[LKH10] Ove Daae Lampe, Johannes Kehrer, and Helwig Hauser. Visual analysis of multivariate movement data using interactive difference views. In *Proceedings of Vision, Modeling, and Visualization*, VMV '10, pages 315–322. EG, 2010.

[LKS13] Bongshin Lee, Rubaiat Habib Kazi, and Greg Smith. Sketchstory: Telling more engaging stories with data through freeform sketching. *IEEE Trans. Vis. Comput. Graph.*, 19(12):2416–2425, 2013.

[LMvW10] Jing Li, Jean-Bernard Martens, and Jarke J van Wijk. Judging correlation from scatterplots and parallel coordinate plots. *Information Visualization*, 9(1):13–30, 2010.

[LP16] Giorgia Lupi and Stefanie Posavec. *Dear Data*. Chronicle Books, 2016.

[LPP+06] Bongshin Lee, Catherine Plaisant, Cynthia Sims Parr, Jean-Daniel Fekete, and Nathalie Henry. Task taxonomy for graph visualization. In *Proceedings of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization*, BELIV '06, pages 1–5, New York, NY, USA, 2006. Association for Computing Machinery.

[Lut06] E. Lutton. Evolution of fractal shapes for artists and designers. *IJAIT, International Journal of Artificial Intelligence Tools*, 15(4):651–672, 2006. Special Issue on AI in Music and Art.

[LWW+13] S. Liu, Y. Wu, E. Wei, M. Liu, and Y. Liu. Storyflow: Tracking the evolution of stories. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2436–2445, Dec 2013.

[MA18] Miro Mannino and Azza Abouzied. Expressive time series querying with hand-drawn scale-free sketches. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, New York, NY, USA, 2018. Association for Computing Machinery.

[Mac92] Alan M. MacEachren. Visualizing Uncertain Information. *Cartographic Perspectives*, 13(Fall):12–19, 1992.

[Mac00] W. E. Mackay. Responding to cognitive overhead: co-adaptation between users and technology. *Intellectica*, 30(1):177–193, 2000.

[Mah77] Michael J Mahoney. Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive therapy and research*, 1(2):161–175, 1977.

[McG84] J.E. McGrath. *Groups: Interaction and Performance*. Prentice-Hall, 1984.

[MDF12] Luana Micallef, Pierre Dragicevic, and Jean-Daniel Fekete. Assessing the effect of visualizations on bayesian reasoning through crowdsourcing. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2536–2545, 2012.

[MF17] Justin Matejka and George Fitzmaurice. Same stats, different graphs: Generating datasets with varied appearance and identical statistics through simulated annealing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 1290–1294, New York, NY, USA, 2017. ACM.

[MFDW17] Dominik Moritz, Danyel Fisher, Bolin Ding, and Chi Wang. Trust, but verify: Optimistic visualizations of approximate queries for exploring big data. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 2904–2915, New York, NY, USA, 2017. Association for Computing Machinery.

[MHH13] Selin A. Malkoc, William Hedgcock, and Steve Hoeffler. Between a rock and a hard place: The failure of the attraction effect among unattractive alternatives. *Journal of Consumer Psychology*, 23(3):317 – 329, 2013.

[MI09] David C. McCallum and Pourang Irani. Arc-pad: absolute+relative cursor positioning for large displays with a mobile touchscreen. In *Proc. UIST '09*, pages 153–156. ACM, 2009.

[Mil00] David R. Millen. Rapid ethnography: Time deepening strategies for hci field research. In *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, DIS '00, pages 280–286, New York, NY, USA, 2000. Association for Computing Machinery.

[Mil03] Elizabeth Shown Mills. Genealogy in the 'information age': History's new frontier? *National Genealogical Society Quarterly*, 91:260–77, December 2003.

[MJ62] Bennet B Murdock Jr. The serial position effect of free recall. *Journal of experimental psychology*, 64(5):482, 1962.

[MLF+12] K.-L. Ma, I. Liao, J. Frazier, H. Hauser, and H. Kostis. Scientific storytelling using visualization. *IEEE Computer Graphics and Applications*, 32(1):12–19, Jan 2012.

[MMD+13] Seth A. Marvel, Travis Martin, Charles R. Doering, David Lusseau, and M. E. J. Newman. The small-world effect is a modern phenomenon. *arXiv e-prints*, page arXiv:1310.2636, Oct 2013.

[MPBW07] Laura G. Militello, Emily S. Patterson, Lynn Bowman, and Robert Wears. Information flow during crisis management: challenges to coordination in the emergency operations center. *Cognition, Technology & Work*, 9(1):25–31, Mar 2007.

[MSD+18] K. Marriott, F. Schreiber, T. Dwyer, K. Klein, N.H. Riche, T. Itoh, W. Stuerzlinger, and B.H. Thomas. *Immersive Analytics*. Lecture Notes in Computer Science. Springer International Publishing, 2018.

[MST+14] Michael R. Marner, Ross T. Smith, Bruce H. Thomas, Karsten Klein, Peter Eades, and Seok-Hee Hong. Gion: Interactively untangling large graphs on wall-sized displays. In Christian Duncan and Antonios Symvonis, editors, *Graph Drawing*, pages 113–124, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg.

[Mun15] Tamara Munzner. *Visualization Analysis & Design*. CRC Press, 2015.

[MW10] Winter Mason and Duncan J Watts. Financial incentives and the performance of crowds. *ACM SigKDD Explorations Newsletter*, 11(2):100–108, 2010.

[NASK18] Sabrina Nusrat, Muhammad Jawaherul Alam, Carlos Scheidegger, and Stephen Kobourov. Cartogram visualization for bivariate Geo-Statistical data. *IEEE Transactions on Visualization and Computer Graphics*, 24(10):2675–2688, October 2018.

[NC07] Daniele Nadalutti and Luca Chittaro. Visual analysis of users' performance data in fitness activities. *Computers & Graphics*, 31(3):429 – 439, 2007.

[NCG93] Richard Nadeau, Edouard Cloutier, and J-H Guay. New evidence about the existence of a bandwagon effect in the opinion formation process. *International Political Science Review*, 14(2):203–213, 1993.

[New02] Mark EJ Newman. Spread of epidemic disease on networks. *Physical review E*, 66(1):016128, 2002.

[NHC12] Miguel Nacenta, Uta Hinrichs, and Sheelagh Carpendale. FatFonts: Combining the Symbolic and Visual Aspects of Numbers. In *Proc. Advanced Visual Interfaces (AVI)*, pages 407–414. ACM, 2012.

[NKD06] Marc Nienhaus, Florian Kirsch, and Jürgen Döllner. Sketchy Illustrations for Presenting the Design of Interactive CSG. In *Proc. IV*, pages 772–777, Los Alamitos, 2006. IEEE Computer Society.

[NM17] Miguel A. Nacenta and Gonzalo Gabriel Méndez. Ivolver: A visual language for constructing visualizations from in-the-wild data. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, ISS '17, pages 438–441, New York, NY, USA, 2017. Association for Computing Machinery.

[NS92] Kai Nagel and Michael Schreckenberg. A cellular automaton model for freeway traffic. *Journal de physique I*, 2(12):2221–2229, 1992.

[NWP+11] Mathieu Nancel, Julie Wagner, Emmanuel Pietriga, Olivier Chapuis, and Wendy Mackay. Mid-air pan-and-zoom on wall-sized displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 177–186, New York, NY, USA, 2011. ACM.

[O+15] Open Science Collaboration et al. Estimating the reproducibility of psychological science. *Science*, 349(6251), 2015.

[OL03] Maria Cristina Ferreira de Oliveira and Haim Levkowitz. From visual data exploration to visual data mining: a survey. *IEEE Transactions on Visualization and Computer Graphics*, 9(3):378–394, 2003.

[OP95] Yigang Pan Sue O'Curry and Robert Pitts. The attraction effect and political choice in two elections. *Journal of Consumer Psychology*, 4(1):85–101, 1995.

[Org19] WorldHealth Organization. Ebola virus disease. `https://www.who.int/en/news-room/fact-sheets/detail/ebola-virus-disease`, 2019. Last accessed: 2019-09-04.

[OS97] Aude Oliva and Phillipe G. Schyns. Coarse Blobs or Fine Edges? Evidence that Information Diagnosticity Changes the Perception of Complex Visual Stimuli. *Cognitive Psychology*, 34(1):72–107, 1997.

[OTS06] Aude Oliva, Antonio Torralba, and Philippe G. Schyns. Hybrid Images. *ACM Transactions on Graphics*, 25(3):527–532, 2006.

[PAEE19] Evan M. Peck, Sofia E. Ayuso, and Omar El-Etr. Data is personal: Attitudes and perceptions of data visualization in rural pennsylvania. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, New York, NY, USA, 2019. Association for Computing Machinery.

[PBC15] Arnaud Prouzeau, Anastasia Bezerianos, and Olivier Chapuis. Road traffic monitoring on a wall display. In *Proceedings of the 27th Conference on l'Interaction Homme-Machine, IHM 2015, Toulouse, France, October 27-30, 2015*, pages 29:1–29:6, 2015.

[PBC16a] Arnaud Prouzeau, Anastasia Bezerianos, and Olivier Chapuis. Towards road traffic management with forecasting on wall displays. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces*, ISS '16, pages 119–128, New York, NY, USA, 2016. Association for Computing Machinery.

[PBC16b] Arnaud Prouzeau, Anastasia Bezerianos, and Olivier Chapuis. Visual immersion in the context of wall displays. In *Proceedings of the 2016 ACM Companion on Interactive Surfaces and Spaces*, ISS '16 Companion, pages 33–39, New York, NY, USA, 2016. Association for Computing Machinery.

[PBC17a] A. Prouzeau, A. Bezerianos, and O. Chapuis. Evaluating multi-user selection for exploring graph topology on wall-displays. *IEEE Transactions on Visualization and Computer Graphics*, 23(8):1936–1951, Aug 2017.

[PBC17b] Arnaud Prouzeau, Anastasia Bezerianos, and Olivier Chapuis. Trade-offs between a vertical shared display and two desktops in a collaborative path-finding task. In *Proceedings of the 43rd Graphics Interface Conference*, GI '17, pages 214–219, Waterloo, CAN, 2017. Canadian Human-Computer Communications Society.

[PBC18] Arnaud Prouzeau, Anastasia Bezerianos, and Olivier Chapuis. Awareness techniques to aid transitions between personal and shared workspaces in multi-display environments. In *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*, ISS '18, pages 291–304, New York, NY, USA, 2018. Association for Computing Machinery.

[PBJ93] John W Payne, James R Bettman, and Eric J Johnson. *The adaptive decision maker*. Cambridge University Press, 1993.

[PBML16] Stefania Pizza, Barry Brown, Donald McMillan, and Airi Lampinen. Smartwatch in vivo. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 5456–5469. ACM, 2016.

[PBP20] V. Peña-Araya, A. Bezerianos, and E. Pietriga. A comparison of geographical propagation visualizations. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)* - to appear. ACM, 2020.

[PC05] Peter Pirolli and Stuart Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. pages 2–4, 2005.

[PGG+09] Kristin Potter, Amy Gooch, Bruce Gooch, Peter Willemsen, Joe Kniss, Richard Riesenfeld, and Peter Shirley. Resolution Independent NPR-Style 3D Line Textures. *Computer Graphics Forum*, 28(1):52–62, March 2009.

[PKF+16] Anshul Vikram Pandey, Josua Krause, Cristian Felix, Jeremy Boy, and Enrico Bertini. Towards understanding human similarity perception in the analysis of large sets of scatter plots. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 3659–3669, New York, NY, USA, 2016. Association for Computing Machinery.

[PLM08] Riccardo Poli, William B. Langdon, and Nicholas Freitag McPhee. *A field guide to genetic programming*. Published via `http://lulu.com` and freely available at `http://www.gp-field-guide.org.uk`, 2008. (With contributions by J. R. Koza).

[Plo93] Scott Plous. *The psychology of judgment and decision making.* Mcgraw-Hill Book Company, 1993.

[Poh16] Rüdiger F Pohl. *Cognitive illusions: Intriguing Phenomena in thinking, judgement and memory*. Psychology Press, 2016.

[PPB20] V. Peña-Araya, E. Pietriga, and A. Bezerianos. A comparison of visualizations for identifying correlation over space and time. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):375–385, Jan 2020.

[Pre90] Fred H. Previc. Functional Specialization in the Lower and Upper Visual Fields in Humans: Its Ecological Origins and Neurophysiological Implications. *Behavioral and Brain Sciences*, 13(3):519–542, September 1990.

[Pro17] Arnaud Prouzeau. *Utilisation collaborative d'un mur d'écran en contexte critique. (Collaboration around wall displays in command and control contexts)*. PhD thesis, University of Paris-Saclay, France, 2017.

[PSTW+17] Stephan Pajer, Marc Streit, Thomas Torsney-Weir, Florian Spechtenhauser, Torsten Muller, and Harald Piringer. Weightlifter: Visual weight space exploration for multi-criteria decision making. *IEEE transactions on visualization and computer graphics*, 23(1):611, 2017.

[PVF13] Charles Perin, Frédéric Vernier, and Jean-Daniel Fekete. Interactive horizon graphs: Improving the compact visualization of multiple time series. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 3217–3226, New York, NY, USA, 2013. Association for Computing Machinery.

[PW06] Matthew D. Plumlee and Colin Ware. Zooming versus multiple window interfaces: Cognitive costs of visual comparisons. *ACM Trans. Comput.-Hum. Interact.*, 13(2):179–209, June 2006.

[RB10] Ronald Rensink and Gideon Baldridge. The perception of correlation in scatterplots. *Comput. Graph. Forum*, 29:1203–1210, 06 2010.

[RCJN17] Juan Rosso, Céline Coutrix, Matt Jones, and Laurence Nigay. Deformable tangible slider for eyes-free one-handed thumb interaction on mobile devices. In *Proceedings of the 29th Conference on l'Interaction Homme-Machine*, IHM '17, pages 21–31, New York, NY, USA, 2017. Association for Computing Machinery.

[RCP+16] Simon Robinson, Céline Coutrix, Jennifer Pearson, Juan Rosso, Matheus Fernandes Torquato, Laurence Nigay, and Matt Jones. Emergeables: Deformable displays for continuous eyes-free mobile interaction. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 3793–3805, New York, NY, USA, 2016. Association for Computing Machinery.

[Rei08] Hannes Reijner. The development of the horizon graph. available online at http://www.stonesc.com/Vis08_Workshop/DVD/Reijner_submission.pdf, 2008.

[Ren02] Ronald A. Rensink. Change detection. *Annual Review of Psychology*, 53(1):245–277, 2002.

[Ren17] Ronald Rensink. The nature of correlation perception in scatterplots. *Psychonomic bulletin & review*, 24:776–797, 06 2017.

[RHDC18] Nathalie Henry Riche, Christophe Hurter, Nicholas Diakopoulos, and Sheelagh Carpendale. *Data-Driven Storytelling*. A. K. Peters, Ltd., USA, 1st edition, 2018.

[RJPL15] Khairi Reda, Andrew E. Johnson, Michael E. Papka, and Jason Leigh. 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, pages 2759–2768, New York, NY, USA, 2015. Association for Computing Machinery.

[RLBR18] Donghao Ren, Bongshin Lee, Matthew Brehmer, and Nathalie Henry Riche. Reflecting on the evaluation of visualization authoring systems: Position paper. In *2018 IEEE Evaluation and Beyond-Methodological Approaches for Visualization (BELIV)*, pages 86–92. IEEE, 2018.

[RLL+05] Kathy Ryall, Neal Lesh, Tom Lanning, Darren Leigh, Hiroaki Miyashita, and Shigeru Makino. Querylines: Approximate query for visual browsing. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pages 1765–1768, New York, NY, USA, 2005. ACM.

[RLP10] Nathalie Henry Riche, Bongshin Lee, and Catherine Plaisant. Understanding interactive legends: A comparative evaluation with standard widgets. In *Proceedings of the 12th Eurographics / IEEE - VGTC Conference on Visualization*, EuroVis'10, pages 1193–1202, Chichester, GBR, 2010. The Eurographs Association & John Wiley & Sons, Ltd.

[ROF12] Patrick Riehmann, Jens Opolka, and Bernd Froehlich. The product explorer: decision making with ease. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pages 423–432. ACM, 2012.

[Ros06] Hans Rosling. The best stats you've ever seen. https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen, 2006. TED, Last accessed: 2019-03-27.

[RSG16] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "why should i trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 1135–1144, New York, NY, USA, 2016. ACM.

[RTR+16] Roy A. Ruddle, Rhys G. Thomas, Rebecca Randell, Philip Quirke, and Darren Treanor. The design and evaluation of interfaces for navigating gigapixel images in digital pathology. *ACM Trans. Comput.-Hum. Interact.*, 23(1), January 2016.

[RWM+15] Fateme Rajabiyazdi, Jagoda Walny, Carrie Mah, John Brosz, and Sheelagh Carpendale. Understanding researchers' use of a large, high-resolution display across disciplines. In *Proceedings of the 2015 International Conference on Interactive Tabletops and Surfaces*, ITS '15, pages 107–116, New York, NY, USA, 2015. Association for Computing Machinery.

[S+19] Ben Shneiderman et al. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. USA, 6th edition, 2019.

[SASS16] Hans-Jörg Schulz, Marco Angelini, Giuseppe Santucci, and H Schumann. An enhanced visualization process model for incremental visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22:1830–1842, 07 2016.

[SBMR12] Tobias Schwarz, Simon Butscher, Jens Mueller, and Harald Reiterer. Content-aware navigation for large displays in context of traffic control rooms. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, AVI '12, pages 249–252, New York, NY, USA, 2012. Association for Computing Machinery.

[SBT+16] Vidya Setlur, Sarah E. Battersby, Melanie Tory, Rich Gossweiler, and Angel X. Chang. Eviza: A natural language interface for visual analysis. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16, pages 365–377, New York, NY, USA, 2016. Association for Computing Machinery.

[SCB+19] A. Sarikaya, M. Correll, L. Bartram, M. Tory, and D. Fisher. What do we talk about when we talk about dashboards? *IEEE Transactions on Visualization and Computer Graphics*, 25(1):682–692, Jan 2019.

[SCI04] Stacey D. Scott, M. Sheelagh T. Carpendale, and Kori M. Inkpen. Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, CSCW '04, pages 294–303, New York, NY, USA, 2004. Association for Computing Machinery.

[SdWvW14] R Scheepens, H v d. Wetering, and J J v. Wijk. Non-overlapping aggregated multivariate glyphs for moving objects. In *IEEE Pacific Visualization Symposium*, pages 17–24, March 2014.

[SE19] Bahador Saket and Alex Endert. Investigating the manual view specification and visualization by demonstration paradigms for visualization construction. *Comput. Graph. Forum*, 38(3):663–674, 2019.

[SFLGM15] Iván Sánchez Fernández, Tobias Loddenkemper, Aristea S. Galanopoulou, and Solomon L. Moshé. Should epileptiform discharges be treated? *Epilepsia*, 56(10):1492–1504, 2015.

[SGL08] John Stasko, Carsten Görg, and Zhicheng Liu. Jigsaw: Supporting investigative analysis through interactive visualization. *Information Visualization*, 7(2):118–132, April 2008.

[SGL09] Y. B. Shrinivasan, D. Gotzy, and J. Lu. Connecting the dots in visual analysis. In *2009 IEEE Symposium on Visual Analytics Science and Technology*, pages 123–130, Oct 2009.

[SH10] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148, Nov 2010.

[SH14] Arvind Satyanarayan and Jeffrey Heer. Authoring narrative visualizations with ellipsis. In *Proc. of the 16th Eurographics Conference on Visualization*, EuroVis '14, pages 361–370, Aire-la-Ville, Switzerland, Switzerland, 2014. Eurographics Association.

[Shn83] Shneiderman. Direct manipulation: A step beyond programming languages. *Computer*, 16(8):57–69, Aug 1983.

[Shn96] B. Shneiderman. The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings 1996 IEEE Symposium on Visual Languages*, pages 336–343, Sep. 1996.

[SHPE20] B. Saket, S. Huron, C. Perin, and A. Endert. Investigating direct manipulation of graphical encodings as a method for user interaction. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):482–491, Jan 2020.

[Sim56] Herbert A Simon. Rational choice and the structure of the environment. *Psychological review*, 63(2):129, 1956.

[Sim78] Herbert A Simon. Rationality as process and as product of thought. *The American economic review*, 68(2):1–16, 1978.

[Sim89] Itamar Simonson. Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, 16(2):158–174, 1989.

[SK16] Drew Skau and Robert Kosara. Arcs, angles, or areas: Individual data encodings in pie and donut charts. *Comput. Graph. Forum*, 35(3):121–130, June 2016.

[SKL+16] Tarique Siddiqui, Albert Kim, John Lee, Karrie Karahalios, and Aditya Parameswaran. Effortless data exploration with zenvisage: An expressive and interactive visual analytics system. *Proc. VLDB Endow.*, 10(4):457–468, November 2016.

[SLR+20] A. Satyanarayan, B. Lee, D. Ren, J. Heer, J. Stasko, J. Thompson, M. Brehmer, and Z. Liu. Critical reflections on visualization authoring systems. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):461–471, Jan 2020.

[SMI99] Thomas Strothotte, Maic Masuch, and Tobias Isenberg. Visualizing Knowledge about Virtual Reconstructions of Ancient Architecture. In *Proc. CGI*, pages 36–43, Los Alamitos, 1999. IEEE.

[SMM12] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2431–2440, Dec 2012.

[SMY+05] Takafumi Saito, Hiroko Nakamura Miyamura, Mitsuyoshi Yamamoto, Hiroki Saito, Yuka Hoshiya, and Takumi Kaseda. Two-tone pseudo coloring: Compact visualization for one-dimensional data. In *Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization*, INFOVIS '05, pages 23–, Washington, DC, USA, 2005. IEEE Computer Society.

[SND05] Purvi Saraiya, Chris North, and Karen Duca. An insight-based methodology for evaluating bioinformatics visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 11(4):443–456, July 2005.

[SPCL+13] Y. Song, D. Pickup, P C. Li, Rosin, and P. Hall. Abstract art by shape classification. *IEEE Transactions on Visualization and Computer Graphics*, 19(8):1252–1263, Aug 2013.

[SRHH16] A. Satyanarayan, R. Russell, J. Hoffswell, and J. Heer. Reactive vega: A streaming dataflow architecture for declarative interactive visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):659–668, Jan 2016.

[SS19] H. Song and D. A. Szafir. Where's my data? evaluating visualizations with missing data. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):914–924, Jan 2019.

[SSRE18] B. Saket, A. Srinivasan, E. D. Ragan, and A. Endert. Evaluating interactive graphical encodings for data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 24(3):1316–1330, March 2018.

[SSS02] Lawrence J Sanna, Norbert Schwarz, and Shevaun L Stocker. When debiasing backfires: accessible content and accessibility experiences in debiasing hindsight. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3):497, 2002.

[SSSEA19] Thilo Spinner, Udo Schlegel, Hanna Schafer, and Mennatallah El-Assady. explAIner: A visual analytics framework for interactive and explainable machine learning. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–1, 2019.

[SSZ+16] D. Sacha, M. Sedlmair, L. Zhang, J. Lee, D. Weiskopf, S. North, and D. Keim. Human-centered machine learning through interactive visualization. In *ESANN 2016 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*. ESANN, 2016.

[Sto12] Maureen Stone. In Color Perception, Size Matters. *IEEE Computer Graphics and Applications*, 32(2):8–13, 2012.

[Sty01] Hilary C. Styron. CSX Tunnel Fire. Technical report, United States Fire Administration, 07 2001.

[SvW08] Yedendra Babu Shrinivasan and Jarke J. van Wijk. Supporting the analytical reasoning process in information visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1237–1246, New York, NY, USA, 2008. Association for Computing Machinery.

[Sza18] D. A. Szafir. Modeling color difference for visualization design. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):392–401, Jan 2018.

[Tak08] Hideyuki Takagi. New topics from recent interactive evolutionary computation researches. In *Knowledge-Based Intelligent Information and Engineering Systems*, page 14, 2008.

[TBBL13] Waldo Cancino Ticona, Nadia Boukhelifa, Anastasia Bezerianos, and Evelyne Lutton. Evolutionary visual exploration: experimental analysis of algorithm behaviour. In *Genetic and Evolutionary Computation Conference, GECCO '13, Amsterdam, The Netherlands, July 6-10, 2013, Companion Material Proceedings*, pages 1373–1380, 2013.

[TBJ15] Theophanis Tsandilas, Anastasia Bezerianos, and Thibaut Jacob. Sketchsliders: Sketching widgets for visual exploration on wall displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 3255–3264, New York, NY, USA, 2015. ACM.

[TC05] James J. Thomas and Kristin A. Cook. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. National Visualization and Analytics Ctr, 2005.

[TGK+17] C. Tominski, S. Gladisch, U. Kister, R. Dachselt, and H. Schumann. Interactive lenses for visualization: An extended survey. *Comput. Graph. Forum*, 36(6):173–200, September 2017.

[THSG04] Edward Tse, Jonathan Histon, Stacey D. Scott, and Saul Greenberg. Avoiding interference: How people use spatial separation and partitioning in sdg workspaces. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, CSCW '04, pages 252–261, New York, NY, USA, 2004. Association for Computing Machinery.

[TJ95] Frank Thomas and Ollie Johnston. *The illusion of life: Disney animation*. Abbeville Press Inc., 1995.

[TKBH17] Cagatay Turkay, Erdem Kaya, Selim Balcisoy, and Helwig Hauser. Designing progressive and interactive analytics processes for high-dimensional data analysis. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):131–140, January 2017.

[TLS+14] C. Turkay, A. Lex, M. Streit, H. Pfister, and H. Hauser. Characterizing cancer subtypes using dual analysis in caleydo stratomex. *IEEE Computer Graphics and Applications*, 34(2):38–47, Mar 2014.

[TM12] Yuzuru Tanahashi and Kwan-Liu Ma. Design considerations for optimizing storyline visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2679–2688, December 2012.

[TMB02] Barbara Tversky, Julie Bauer Morrison, and Mireille Bétrancourt. Animation: can it facilitate? *International Journal of Human-Computer Studies*, 57(4):247–262, 2002.

[TSA14] J. Talbot, V. Setlur, and A. Anand. Four experiments on the perception of bar charts. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2152–2160, Dec 2014.

[TTP+06] Anthony Tang, Melanie Tory, Barry Po, Petra Neumann, and Sheelagh Carpendale. Collaborative coupling over tabletop displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, pages 1181–1190, New York, NY, USA, 2006. Association for Computing Machinery.

[Tuf86] Edward R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, CT, USA, 1986.

[Tve72] A. Tversky. Elimination by aspects: A theory of choice. *Psychological Review*, 79:281–299, 1972.

[Uba13] Barbara Ubaldi. Open government data. (22), 2013.

[VB10] Daniel Vogel and Ravin Balakrishnan. Occlusion-aware interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 263–272, New York, NY, USA, 2010. ACM.

[VBS+06] Cécile Viboud, Ottar N. Bjørnstad, David L. Smith, Lone Simonsen, Mark A. Miller, and Bryan T. Grenfell. Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science*, 312(5772):447–451, 2006.

[vdEvW14] S. van den Elzen and J. van Wijk. Multivariate network exploration and presentation: From detail to overview via selections and aggregations. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2310–2319, Dec 2014.

[vHvW04] F. van Ham and J.J. van Wijk. Interactive visualization of small world graphs. In *IEEE Symposium on Information Visualization*, pages 199–206, Oct 2004.

[VLKS+11] Tatiana Von Landesberger, Arjan Kuijper, Tobias Schreck, Jörn Kohlhammer, Jarke Van Wijk, Jean-Daniel Fekete, and Dieter W. Fellner. Visual Analysis of Large Graphs: State-of-the-Art and Future Research Challenges. *Computer Graphics Forum*, 30(6):1719–1749, January 2011.

[VP04] Vivek Verma and Alex Pang. Comparative flow visualization. *IEEE Transactions on Visualization and Computer Graphics*, 10(6):609–624, November 2004.

[vV99] J. J. van Wijk and E. R. Van Selow. Cluster and calendar based visualization of time series data. In *Proceedings 1999 IEEE Symposium on Information Visualization (InfoVis'99)*, pages 4–9, Oct 1999.

[VWvH+07] Fernanda B. Viegas, Martin Wattenberg, Frank van Ham, Jesse Kriss, and Matt McKeon. Manyeyes: A site for visualization at internet scale. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1121–1128, November 2007.

[Wag06] Mark Wagner. *The Geometries of Visual Space*. Lawrence Erlbaum Associates, Mahwah, NJ, USA, 2006.

[War02] Matthew O. Ward. A Taxonomy for Glyph Placement Strategies for Multidimensional Data Visualization. *Information Visualization*, 1(6):194–210, 2002.

[War04] Colin Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2004.

[Wat01] Martin Wattenberg. Sketching a graph to query a time-series database. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '01, pages 381–382, New York, NY, USA, 2001. ACM.

[WBS+08] B. Watson, D. Brink, M. Stallman, D. Devajaran, and H. Rhyne, T.-M. amd Patel. Visualizing very large layered graphs with quilts. Technical Report TR-2008-10, NC State University, 2008.

[WBWK00] Michelle Q. Wang Baldonado, Allison Woodruff, and Allan Kuchinsky. Guidelines for using multiple views in information visualization. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '00, pages 110–119, New York, NY, USA, 2000. ACM.

[Wea04] Chris Weaver. Building highly-coordinated visualizations in improvise. In *Proceedings of the IEEE Symposium on Information Visualization*, INFOVIS '04, pages 159–166, USA, 2004. IEEE Computer Society.

[Wea16] Mark Wilkinson and et al. The fair guiding principles for scientific data management and stewardship. *Nature Scientific Data*, (160018), 2016.

[Wed91] Douglas H Wedell. Distinguishing among models of contextually induced preference reversals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(4):767, 1991.

[WF94] Stanley Wasserman and Katherine Faust. *Social Network Analysis*. Cambridge University Press, 1994.

[WHA07] Wesley Willett, Jeffrey Heer, and Maneesh Agrawala. Scented widgets: Improving navigation cues with embedded visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1129–1136, November 2007.

[WII+12a] J. Wood, P. Isenberg, T. Isenberg, J. Dykes, N. Boukhelifa, and A. Slingsby. Sketchy rendering for information visualization. *Visualization and Computer Graphics, IEEE Transactions on*, 18(12):2749–2758, Dec 2012.

[WII+12b] J. Wood, P. Isenberg, T. Isenberg, J. Dykes, N. Boukhelifa, and A. Slingsby. Sketchy rendering for information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2749–2758, Dec 2012.

[WL95] Colin Ware and Marlon Lewis. The dragmag image magnifier. In *Conference Companion on Human Factors in Computing Systems*, CHI '95, pages 407–408, New York, NY, USA, 1995. ACM.

[WMP+05] Pak Chung Wong, Patrick Mackey, Ken Perrine, James Eagan, Harlan Foote, and Jim Thomas. Dynamic Visualization of Graphs with Extended Labels. In *Proc. Symposium on Information Visualization (InfoVis)*, pages 73–80. IEEE, 2005.

[Wri75] Peter Wright. Consumer choice strategies: Simplifying vs. optimizing. *Journal of Marketing Research*, pages 60–67, 1975.

[WS92] Christopher Williamson and Ben Shneiderman. The dynamic homefinder: Evaluating dynamic queries in a real-estate information exploration system. In *Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '92, pages 338–346, New York, NY, USA, 1992. ACM.

[WSFB07] Daniel Wigdor, Chia Shen, Clifton Forlines, and Ravin Balakrishnan. Perception of elementary graphical elements in tabletop and multi-surface environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 473–482, New York, NY, USA, 2007. ACM.

[WSK+00] Bin Wei, Claudio Silva, Eleftherios Koutsofios, Shankar Krishnan, and Stephen North. Visualization research with large displays. *IEEE Comput. Graph. Appl.*, 20(4):50–54, July 2000.

[WSM13] James R. Wallace, Stacey D. Scott, and Carolyn G. MacGregor. Collaborative sensemaking on a digital tabletop and personal tablets: Prioritization, comparisons, and tableaux. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 3345–3354, New York, NY, USA, 2013. Association for Computing Machinery.

[WSS+09] James R. Wallace, Stacey D. Scott, Taryn Stutz, Tricia Enns, and Kori Inkpen. Investigating teamwork and taskwork in single- and multi-display groupware systems. *Personal Ubiquitous Comput.*, 13(8):569–581, November 2009.

[WW08] Leland Wilkinson and Graham Wills. Scagnostics Distributions. *Journal of Computational and Graphical Statistics*, 17(2):473–491, 2008.

[WWS20] M. Whitlock, K. Wu, and D. A. Szafir. Designing for mobile and immersive visual analytics in the field. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):503–513, Jan 2020.

[XHRC+18] Haijun Xia, Nathalie Henry Riche, Fanny Chevalier, Bruno De Araujo, and Daniel Wigdor. Dataink: Direct and creative data-oriented drawing. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, New York, NY, USA, 2018. Association for Computing Machinery.

[YFDH01] Ka-Ping Yee, Danyel Fisher, Rachna Dhamija, and Marti Hearst. Animated exploration of dynamic graphs with radial layout. In *Proceedings of the IEEE Symposium on Information Visualization*, pages 43–50, 2001.

[YHR+19] Fumeng Yang, Lane T Harrison, Ronald A Rensink, Steven L Franconeri, and Remco Chang. Correlation judgment and visualization features: A comparative study. *IEEE Transactions on Visualization and Computer Graphics*, 25(3):1474–1488, March 2019.

[YMSJ05] Ji Soo Yi, Rachel Melton, John Stasko, and Julie A Jacko. Dust & magnet: multivariate information visualization using a magnet metaphor. *Information Visualization*, 4(4):239–256, 2005.

[YN06] Beth Yost and Chris North. The perceptual scalability of visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):837–844, September 2006.

[ZBK15] Yunfeng Zhang, Rachel KE Bellamy, and Wendy A Kellogg. Designing information for remediating cognitive biases in decision-making. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2211–2220. ACM, 2015.

[ZGC+17] E. Zgraggen, A. Galakatos, A. Crotty, J. Fekete, and T. Kraska. How progressive visualizations affect exploratory analysis. *IEEE Transactions on Visualization and Computer Graphics*, 23(8):1977–1987, Aug 2017.

[ZGI+18] J. Zhao, M. Glueck, P. Isenberg, F. Chevalier, and A. Khan. Supporting handoff in asynchronous collaborative sensemaking using knowledge-transfer graphs. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):340–350, Jan 2018.

[ZIP16] Kostas Zoumpatianos, Stratos Idreos, and Themis Palpanas. Ads: The adaptive data series index. *The VLDB Journal*, 25(6):843–866, December 2016.

[ZK14] Björn Zimmer and Andreas Kerren. Sensemaking and provenance in distributed collaborative node-link visualizations. In *In IEEE VIS '14 Workshop on Provenance for Sensemaking*, pages 1–2, 11 2014.

[ZSBC20] Jiayi Eris Zhang, Nicole Sultanum, Anastasia Bezerianos, and Fanny Chevalier. Dataquilt: Extracting visual elements from images to craft pictorial visualizations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, *to appear*. ACM, 2020.

Curriculum Vitae

Current Situation & Research Positions

Sep 2011 - present	Associate Professor, Université Paris-Sud XI, UFR Sciences, Orsay member of the ILDA (Inria) team and HCC (LRI) team
2011-2018	Co-Responsable of Master 2 Research en Informatique, specialité Interaction Co-Responsable of European EIT ICT Master in Human Computer Interaction and Design (HCID)
2009-2011	Assistant Professor, École Centrale Paris SAP-BusinessObjects Chair
2007-2009	Researcher, National ICT Australia (NICTA) HxI Initiative between DSTO, CSIRO and NICTA (Braccetto project) Adjunct Professor, University of Sydney

Education

2001-2007	**PhD** Computer Science Department, University of Toronto, Canada Thesis: "Designs for single user, up-close interaction with Wall-sized displays" Supervisor: Ravin Balakrishnan Committee: Ken Hinckley (external reporter), Ron Baecker, Mark Chignell, Karan Singh, Khai Truong.
1999-2001	**MSc** Computer Science Department, University of Toronto, Canada Thesis: "Using Projection to Accelerate Ray-Tracing" Supervisor: Alejo Hausner, Reporter: James Stewart
1994-1999	**BS** Department of Informatics, University of Piraeus, Greece Thesis: "Defining and solving timetable scheduling problems using graphical constraint representations" Supervisor: Themis Panayiotopoulos

Fellowships & Distinctions

2018-now	Prime d'Encadrement Doctoral et de Recherche (PEDR) - top 20% of Associate Professors nationally
2017	IEEE VIS/InfoVis - Honorable mention (top 4 papers of conference)
2015	ACM CHI 2015 - Honorable mention (top 5% of papers)
2013	INTERACT 2013 - Brian Shackel Award (Best Paper)
2007	SIGGRAPH 2007 - Invited to Highlights from UIST session
2006	UIST 2006 - Best Student Paper Award
1999-2005	University of Toronto Open Fellowship
1999	Graduated top of class in Department of Informatics, University of Piraeus

Research Interests & Objectives

Research focused on the design and evaluation of novel interaction and visualization techniques, in the context of large displays and large visual datasets. Interested in such designs for single and multi-user interaction settings, collocated and distributed, and in the evaluation methods to validate them.

Research Supervision

PhD Supervision[3]

2018-now	Co-Supervisor of PhD student Raphael James (60%, supervisor E. Pietriga, O. Chapuis & T. Dwyer) Title: "Immersive Analytics in Wall-display Environments"
2018-now	Co-Supervisor of PhD student Tong Xue (80%, supervisor E. Pietriga) Title: "Knowledge Discovery in Data Journalism"
2016-2019	Supervisor of PhD student Anna Gogolou (60%, co-supervisor T. Palpanas) Title: "Iterative and Expressive Querying for Big Data Series Publications: IEEE VIS/InfoVis [J19], BigVis Workshop 2019 [O13] and 1 under submission
2014-2017	Co-Supervisor of PhD student Arnaud Prouzeau (60%, supervisor O. Chapuis) Title: "Collaboration around Wall-Displays in Command and Control Contexts Publications: TVCG [J14], ACM ISS [C14] and [C17], GraphicsInterface [C15], ISS workshop [O10], IHM [O8] Current Position: Post-Doc at University of Monash, Australia.
2014-now	Co-Supervisor of PhD student Evanthia Dimara (50%, supervisor P. Dragicevic) Title: "Information Visualization for Decision Making" Publications: TVCG/InfoVis [J15]-Honorable Mention, [J16] and [J17], TVCG [J20], ACM CHI [C16], IEEE VIS workshop [O6] Current Position: Research Scientist at University of Konstanz, Germany
2009-2012	Co-Supervisor of PhD student Micheline Elias (80%, supervisor M.-A. Aufaure) Title: "Enhancing Human Interaction with Business Intelligence Dashboards" Publications: INTERACT [C10]-Best Paper Award and [C8], ACM CHI [C9], and 2 patent submissions Current Position: Analytical Application Manager at Collective[i] New York.
2019	Internship supervisor of visiting PhD student Nicole Sultanum (3 months in U. Paris-Saclay) Title: "Data Visualization for Narrative Understanding in Journalism with Storifier" Publication: under submission
2007	Internship supervisor of visiting PhD student Nathalie Henry (3 months in NICTA) Title: "Exploring Large Social Networks with Matrix-Based Representations" Publication: IEEE VIS/InfoVis [J2] Current Position: Researcher at Microsoft Research Redmond USA.

Master Supervision

2014	Co-supervisor of MSc student Thanasis Taousakos Title: "Detecting and Simplifying User's routines on Mobile devices"
2013	Co-supervisor of MSc student Thibaut Jacob Title: "Sketching Interactions for Data Exploration" Publication: ACM CHI [C13]
2012	Co-supervisor of MSc student Stelios Frantzeskakis Title: "Navigating large information spaces on wall displays using mobile devices" Publication: ACM CHI [C12]
2012	Co-supervisor of MSc student Jeronimo Barbosa Title: "Combining interactive pen and tangibles for technical drawing"
2009	Supervisor of MSc student Micheline Elias Title: "Visual Exploration of Time-Varying Data Cubes" Publication: 1 patent submission

[3]In France professors without an HDR generally cannot supervise PhD thesis alone, in my list I give my percentage of supervision.

Scientific Community Involvement and Juries

Chairing	ACM CHI Conference: Subcommittee Chair[4](2019-2020) for Visualization
Program Committee	ACM CHI Conference (2015-2018) IEEE InfoVis Conference (2015-2017, 2018-now) ACM UIST Conference (2018-2019) GraphDrawing Conference (2014) IEEE PacificVis Conference (2010-2013)
Organization Committee	IEEE VIS Workshops co-chair (2019-now) BELIV Workshop co-organizer (2020) IEEE VIS Community co-chair (2017-2018) ACM CHI Interactivity co-chair (2009) IEEE VIS Workshop organizer (EValuation of Interactive VisuAl ML systems) IHM Demos Co-Chair, 2015
Other Committees:	IEEE VIS Best Short Paper committee - Chair (2019) IEEE VIS Best Poster committee (2018)
Other Community Involvement	IEEE VIS Ombuds representative (2019-now)
PhD Thesis Juries	Reporter for PhD of F. Rajabiyazdi, University of Calgary, Canada (2019) Jury member for PhD of A. Chalbi, Université Lille 1, France (2019) Reporter for PhD of S. Rufiange, École de technologie supérieure -ÉTS, Canada (2013)
Reviewing	NSERC, Canadian Discovery Grands (2018-2020) ACM Transactions on Computer-Human Interaction - ToCHI (since 2009) IEEE Transactions on Computer Graphics - TVCG (since 2009) IEEE Conferences: InfoVis & VAST, EuroGraphics (since 2007), PacificVis (since 2009) ACM Conferences: CHI, UIST, CSCW (since 2005), AVI, ISS and ITS (since 2007)
Local Community	President of the ACM SIGCHI Parisian chapter (2014-2016) Vice-President of the ACM SICCHI Parisian chapter (2012-2014) Treasurer of the ACM SICCHI Parisian chapter (2010-2012) Co-founder Parisian ACM SIGCHI Chapter (2010)
Student Volunteer	ACM UIST 2005

Funded Projects

iCoda 2018-now	funded by Inria (Inria Project Labs' initiative). Head of the HCI/Vis work package. Funding for 3 PhDs (e.g., PhD student T. Xue), 2 engineers, and 80K EU for expenses.
Digiscope 2011-2016	funded by ANR (French National Agency) Equipement d'Excellence (advanced infrastructure) Budget 6 million EU. Head for École Centrale Paris, continued involvement in Univ. Paris-Sud.
DaEv 2012	funded by Paris-Sud, Projet Attractivité (AAP). Primary and sole Investigator. 35K EU for my research (equipment and travel).
CUBIST 2010-2011	EU FP7 project. Wrote the HCI/VIS package and was member of the package. Budget 3 million EU (project ended in 2013).

[4]ACM CHI is organized in thematic subcommittees/areas, a Subcommittee Chair acts as a paper chair for the subcommittee. Decisions are taken in a live program committee meeting.

List of Publications

In the field of Human Computer Interaction, conferences are considered the primary method of publication. The ACM CHI and UIST conferences are particularly selective and are considered journal level work by the ACM/SIGCHI. See http://dx.doi.org/10.1145/1743546.1743569

In the field of Visualization, the major publication venue is the IEEE Transactions on Visualization and Computer Graphics journal. The proceedings of the top conferences in the domain (IEEE InfoVis/Vis/VAST) go through a journal reviewing process and are published as part of this journal.

International Peer-reviewed Journal Articles

[J.21] V. Peña-Araya, E. Pietriga and **A. Bezerianos** (2019). A comparison of visualizations for identifying correlation over space. In *IEEE InfoVis 2019 - the IEEE Transactions on Visualization and Computer Graphics*, 26(1), 2019, 10 pages.

[J.20] E. Dimara, S. Franconeri, C. Plaisant, **A. Bezerianos**, and P. Dragicevic (2018). A Task-based Taxonomy of Cognitive Biases for Information Visualization. In *TVCG - the IEEE Transactions on Visualization and Computer Graphics*, 20 pages.

[J.19] A. Gogolou, T. Tsandilas, T. Palpanas, **A. Bezerianos** (2018). Comparing Similarity Perception in Time Series Visualizations. In *IEEE InfoVis 2018 - the IEEE Transactions on Visualization and Computer Graphics*, 25(1), 2018, 10 pages. [25.7% acc. rate]

[J.18] T. Blascheck, L. Besan con, **A. Bezerianos**, B. Lee, P. Isenberg (2018). Glanceable Visualization: A Study of Perception Time for Data Comparison Tasks on Smartwatches. In *IEEE InfoVis 2018 - the IEEE Transactions on Visualization and Computer Graphics*, 25(1), 2018, 10 pages. [25.7% acc. rate]

[J.17] E. Dimara, G. Bailly, **A. Bezerianos**, S. Franconeri (2018). Mitigating the Attraction Effect with Visualizations. In *IEEE InfoVis 2018 - the IEEE Transactions on Visualization and Computer Graphics*, 25(1), 2018, 10 pages. [25.7% acc. rate]

[J.16] E. Dimara, **A. Bezerianos**, P. Dragicevic (2017). Conceptual and Methodological Issues in Evaluating Multidimensional Visualizations for Decision Support. In *IEEE InfoVis 2017 - the IEEE Transactions on Visualization and Computer Graphics*, 24(1), 2018, 10 pages. [23% acc. rate]

[J.15] E. Dimara, **A. Bezerianos**, P. Dragicevic (2016). The Attraction Effect in Information Visualization. In *IEEE InfoVis 2016 - the IEEE Transactions on Visualization and Computer Graphics*, 23(1), 10 pages. Best paper Honorable Mention (4 best papers of conference) [23% acc. rate]

[J.14] A. Prouzeau, **A. Bezerianos**, O. Chapuis (2016). Evaluating multi-user selection for exploring graph topology on wall-displays. In *TVCG - the IEEE Transactions on Visualization and Computer Graphics*, 14 pages.

[J.13] J. Fuchs, P. Isenberg, **A. Bezerianos**, F. Fischer, and D. Keim (2016). A Systematic Review of Experimental Studies on Data Glyphs. In *TVCG - the IEEE Transactions on Visualization and Computer Graphics*, 14 pages.

[J.12] N. Boukhelifa, A. Bezerianos, W. Cancino, E. Lutton (2015). Evolutionary Visual Exploration: Evaluation of an IEC Framework for Guided Visual Search. In Evolutionary Computation, Massachusetts Institute of Technology Press (MIT Press), 30 pages, 2015.

[J.11] J. Fuchs, P. Isenberg, **A. Bezerianos**, F. Fischer, and E. Bertini (2014). The Influence of Contour on Similarity Perception of Star Glyphs. In *IEEE InfoVis 2014 - the IEEE Transactions on Visualization and Computer Graphics*, 20(12), pp.2251-2260.[23% acc. rate]

[J.10] P. Isenberg, P. Dragicevic, W. Willett, **A. Bezerianos**, and J.-D. Fekete (2013). Hybrid- Image Visualization for Large Viewing Environments. In *IEEE InfoVis 2013 - the IEEE Transactions on Visualization and Computer Graphics*, 19(12): 2346-2355, (10 pages). [25% acc. rate]

[J.9] N. Boukhelifa, W. Cancino, **A. Bezerianos**, and E. Lutton (2013). Evolutionary Visual Exploration: Evaluation With Expert Users. In *Eurovis 2013 - Eurographics/IEEE- VGTC Symposium on Visualization*, 32(3): 31-40, (10 pages) [28% acc. rate]

[J.8] **A. Bezerianos** and P. Isenberg (2012). Perception of Visual Variables on Tiled Wall- Sized Displays for Information Visualization Applications. In *IEEE InfoVis 2012 - the IEEE Transactions on Visualization and Computer Graphics (Proceedings Scientific Visualization / Information Visualization 2012)*, 18(12): 2516-2525, (10 pages). [25% acc. rate]

[J.7] N. Boukhelifa, **A. Bezerianos**, T. Isenberg and J.-D. Fekete (2012). Evaluating Sketchiness as a Visual Variable for the Depiction of Qualitative Uncertainty. In *IEEE InfoVis 2012 - the IEEE Transactions on Visualization and Computer Graphics (Proceedings Scientific Visualization / Information Visualization 2012)*, 18(12): 2769-2778, (10 pages). [25% acc. rate]

[J.6] P. Isenberg, **A. Bezerianos**, P. Dragicevic and J.-D. Fekete (2011). A Study on Dual- Scale Data Charts. In *In IEEE InfoVis 2011- the IEEE Transactions on Visualization and Computer Graphics*, 17(12): 2469-2487, (10 pages) [26% acc. rate]

[J.5] **A. Bezerianos**, P. Dragicevic, J.-D. Fekete, J. Bae, B. Watson (2010). GeneaQuilts: A System for Exploring Large Genealogies. In *IEEE InfoVis 2010 - the IEEE Transactions on Visualization and Computer Graphics*, 16(6): 1073-1081, (10 pages) [26% acc. rate]

[J.4] **A. Bezerianos**, F. Chevalier, P. Dragicevic, N. Elmqvist and J.-D. Fekete (2010). GraphDice: A System for Exploring Multivariate Social Networks. In *Eurovis 2010 - Eurographics/IEEE-VGTC Symposium on Visualization*, 29(3): 863-872, (10 pages) [29% acc. rate]

[J.3] P. Isenberg, **A. Bezerianos**, N. Henry, S. Carpendale, and J.-D. Fekete (2009). CoCoNutTrix: Collaborative Retrofitting for Information Visualization. In *IEEE Computer Graphics & Applications: Special Issue on Collaborative Visualization*, 29(5):44-57, (14 pages).

[J.2] N. Henry, **A. Bezerianos**, J.-D. Fekete (2008). Improving the Readability of Clustered Social Networks by Node Duplication. In *IEEE InfoVis 2008 - the IEEE Transactions on Visualization and Computer Graphics*, 14(6): 1317-1324, (8 pages).

[J.1] **A. Bezerianos**, R. Balakrishnan (2005). Canvas Portals: View and Space Management on Large Displays. In *IEEE Computer Graphics & Applications, Special Issue on Applications of Large Displays*, 25(4): 34-43, (10 pages).

International Peer-reviewed Conference Articles

[C.21] E. Zhang, N. Sultanum, **A. Bezerianos**, and F. Chevalier (2020). A comparison of geographical propagation visualizations. In *ACM CHI 2020 - the ACM SIGCHI conference on Human Factors in Computing Systems*, 14 pages, to appear.

[C.20] V. Peña-Araya, **A. Bezerianos**, and E. Pietriga (2020). A comparison of geographical propagation visualizations. In *ACM CHI 2020 - the ACM SIGCHI conference on Human Factors in Computing Systems*, 14 pages, to appear.

[C.19] N. Boukhelifa, **A. Bezerianos**, I.C. Trelea, N. Mejean-Perrot, E. Lutton (2019). An Exploratory Study on Visual Exploration of Model Simulations by Multiple Types of Experts. In *ACM CHI 2019 - the ACM SIGCHI conference on Human Factors in Computing Systems*, 14 pages [23.8% acc. rate].

[C.18] J. Fuchs, P Isenberg, **A. Bezerianos**, M Miller, D Keim. EduClust -A Visualization Application for Teaching Clustering Algorithms. In *Eurographics - Education Papers*, May 2019, 8 pages.

[C.17] A. Prouzeau, **A. Bezerianos**, O. Chapuis (2018). Designs for Transitioning from Loose to Close Collaboration in Multi-Display Environments. In *ACM ISS 2018 - the ACM International Conference on Interactive Surfaces and Spaces*, 14 pages [26.7% acc. rate]

[C.16] E. Dimara, **A. Bezerianos**, P. Dragicevic (2017). Narratives in Crowdsourced Evaluation of Visualizations: A Double-Edged Sword? In *ACM CHI 2017 - the ACM SIGCHI conference on Human Factors in Computing Systems*, pages 5475-5484. [25% acc. rate]

[C.15] A. Prouzeau, A. Bezerianos, O. Chapuis (2017). Trade-offs Between a Vertical Shared Display and Two Desktops in a Collaborative Path-Finding Task. In GI-2017 - Proceedings of Graphics Interface, pages 214-219.

[C.14] A. Prouzeau, **A. Bezerianos**, O. Chapuis (2016). Towards road traffic management with forecasting on wall displays. In *ISS 2016 - Proceedings of the 2016 ACM on Interactive Surfaces and Spaces*, pages 119-128 (10 pages).

[C.13] T. Tsandilas, **A. Bezerianos**, T. Jacobs (2015). SketchSliders: Sketching Widgets for Visual Exploration on Wall Displays. To appear in *ACM CHI 2015 - the ACM SICGHI Conference on Human Factors in Computing Systems*, (10 pages), Honorable mention (top 5% of papers).

[C.12] O. Chapuis, **A. Bezerianos**, S. Frantzeskakis (2014). Smarties: An Input System for Wall Display Development. In *ACM CHI 2014 - the ACM SICGHI Conference on Human Factors in Computing Systems*, (10 pages). [23% acc. rate]

[C.11] T. Tsantilas, C. Appert, **A. Bezerianos**, and D. Bonnet (2014). Coordination of Tilt and Touch in One- and Two-Handed Use. In *ACM CHI 2014 - the ACM SICGHI Conference on Human Factors in Computing Systems*, (4 pages). [23% acc. rate].

[C.10] M. Elias, M.-A. Aufaure, and **A. Bezerianos** (2013). Storytelling in Visual Analytics tools for Business Intelligence. In *INTERACT 2013 - the 14th IFIP TC13 Conference on Human-Computer Interaction*, ParLNCS 8119, pages 280-297, (18 pages). Brian Shackel Award (best paper).

[C.9] M. Elias and **A. Bezerianos** (2012). Annotating BI visualization dashboards: needs & challenges. In *ACM CHI 2012 - the ACM SICGHI Conference on Human Factors in Computing Systems*, pages 1641-1650, (10 pages). [23% acc. rate]

[C.8] M. Elias and **A. Bezerianos** (2011) Exploration Views: understanding dashboard creation & customization for visualization novices. In *INTERACT 2011 - the 13th IFIP TC13 Conference on Human-Computer Interaction*, pages 274-291, (18 pages).

[C.7] P. Dragicevic, **A. Bezerianos**, W. Javed, N. Elmqvist, and J.-D. Fekete (2011). Temporal Distortion for Animated Transitions. In *ACM CHI 2011 - the ACM SICGHI Conference on Human Factors in Computing Systems*, pages 2009-2018, (10 pages). [23% acc. rate]

[C.6] C. Melo, B. Le Grand, M.-A. Aufaure, and **A. Bezerianos** (2011). Extracting and Visualizing Tree-like Structures from Concept Lattices. In *IV 2011 - the 15th International Conference on Information Visualization*, pages 261-266, (6 pages).

[C.5] F. Chevalier, P. Dragicevic, **A. Bezerianos**, and J.-D. Fekete (2010). Using Text Animated Transitions to Support Navigation in Document Histories. In *ACM CHI 2010 - the ACM SICGHI Conference on Human Factors in Computing Systems*, pages 683-692, (10 pages). [22% acc. rate].

[C.4] A. Collins, **A. Bezerianos**, G. McEwan, M. Rittenbruch, R. Wasinger, and J. Kay (2009). Understanding File Access Mechanisms for Embedded Ubicomp Collaboration Interfaces. In *Ubicomp 2009- the 11th international Conference on Ubiquitous Computing*, pages 135-144, (10 pages). [12% acc. rate].

[C.3] **A. Bezerianos** (2007). Using alternative views for layout, comparison and context switching tasks in wall displays. In *OzCHI 2007 - the Australasian Computer-Human Interaction Conference*, pages 303-310, (8 pages).

[C.2] **A. Bezerianos**, P. Dragicevic, and R. Balakrishnan (2006). Mnemonic Rendering: An Image-Based Approach for Exposing Hidden Changes in Dynamic Displays. In *ACM UIST 2006 - the 19th ACM Symposium on User Interface Software and Technology*, pages 159-168, (10 pages). Best Student Paper Award. [23% acc. rate].

[C.1] **A. Bezerianos** and R. Balakrishnan (2005). The Vacuum: Facilitating the manipulation of distant objects. In *ACM CHI 2005 - the ACM SICGHI Conference on Human Factors in Computing Systems*, pages 361-270, (10 pages) [25% acc. rate].

Book Chapters

[B.2] M. Billinghurst, M. Cordeil, **A. Bezerianos**, and T. Margolis. Collaborative Immersive Analytics. In Immersive Analytics. pp.221-257, ISBN:9783030013882, Kim Marriott, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger, Bruce H. Thomas (Eds.), Springer, 2018.

[B.1] N. Boukhelifa, **A. Bezerianos**, E. Lutton. Evaluation of Interactive Machine Learning Systems. In Human and Machine Learning Visible, Explainable, Trustworthy and Transparent. pp.341-360, ISBN:978-3-319-90403-0, Zhou, Jianlong, Chen, Fang (Eds.), Springer, 2018.

Theses

[3] **A. Bezerianos** Designs for single user, up-close interaction with Wall-sized displays. Ph.D. Thesis, Dep. of Computer Science, University of Toronto, 2007.

[2] **A. Bezerianos** Using Projection to Accelerate Ray Tracing. M.Sc. Thesis, Dep. of Computer Science, University of Toronto, 2001.

[1] **A. Bezerianos** Defining and solving timetable scheduling problems using graphical constraint representations. Hon. Thesis. Department of Informatics, University of Piraeus, 1999.

Other Peer-reviewed Articles

[O.13] A Gogolou, T Tsandilas, T Palpanas, A Bezerianos (2019). Progressive Similarity Search on Time Series Data. BigVis Workshop 2019 in EDBT/ ICDT.

[O.12] T. Blascheck, L. Besan con, A. Bezerianos, B. Lee, P. Isenberg (2018). Preparing for Perceptual Studies: Position and Orientation of Wrist-worn Smartwatches for Reading Tasks. Workshop on Data Visualization on Mobile Devices, ACM CHI, 2018, Montréal, Canada

[O.11] T. Blascheck, L. Besan con, A. Bezerianos, B. Lee, P. Isenberg (2018) Perception des visualisations sur smartwatch. Journées Visu 2018, May 2018, EDF-Saclay

[O.10] A. Prouzeau, **A. Bezerianos**, O. Chapuis (2016). Visual Immersion in the Context of Wall Displays. Interactive Surfaces and Spaces Surfaces Companion Proceedings.

- [O.9] P. Goffin, W. Willett, A. Bezerianos, P. Isenberg (2015). Exploring the Effect of Word-Scale Visualizations on Reading Behavior. In ACM CHI 2015 - Ext. Abstracts
- [O.8] A. Prouzeau, **A. Bezerianos**, and O. Chapuis (2015). Surveillance du trafic routier avec un mur d'écrans. In IHM TeC '15: Proceedings of the 27th international conference of the Association Francophone d'Interaction Homme-Machine, Travaux en Cours, 6 pages, ACM, 2015.
- [O.7] T. Onorati, P. Isenberg, **A. Bezerianos**, E. Pietriga, P. Diaz (2015). WallTweet: A Knowledge Ecosystem for Supporting Situation Awareness. ITS Workshop on Data Exploration for Interactive Surfaces (DEXIS). 2015
- [O.6] E. Dimara, P. Dragicevic, **A. Bezerianos** (2014). Accounting for Availability Biases in Information Visualization. In *IEEE VIS workshop on Dealing with Cognitive Biases in Visualisations (DECISIVe)*, (4 pages).
- [O.5] W. Cancino, N. Boukhelifa, **A. Bezerianos**, and E. Lutton (2013). Evolutionary Visual Exploration: Experimental Analysis of Algorithm Behavior. In *GECCO workshop on Genetic and Evolutionary Computation (VizGEC 2013)*, (8 pages).
- [O.4] **A. Bezerianos**, P. Isenberg, O. Chapuis, and W. Willett (2013). Perceptual Affordances of Wall-Sized Displays for Visualization Applications: Color. In *ACM CHI 2013- Extended Abstracts, Workshop on Interactive, Ultra-High-Resolution Displays (PowerWall)*, (6 pages).
- [O.3] C. Melo, B. Le Grand, **A. Bezerianos**, M.-A. Aufaure (2011). Parent Selection Criterion for Extracting Trees from Concept Lattices. In *CUBIST Workshop 2011*, (10 pages).
- [O.2] Y. Riche, N. Henry Riche, P. Isenberg, and **A. Bezerianos** (2010). Hard-To-Used Interfaces Considered Beneficial (Some of the Time). In *ACM CHI 2010 -Extended Abstracts (alt.chi)*, pages 2705-2714, (10 pages).
- [O.1] **A. Bezerianos** and G. McEwan (2008). Presence disparity in mixed presence collaboration. In *ACM CHI 2008 - Extended Abstracts*, pages 3285-3290, (6 pages).

Sample Puplications

*Sketch*Sliders: Sketching Widgets for Visual Exploration on Wall Displays

Theophanis Tsandilas[1,2]
fanis@lri.fr
[1]INRIA
F-91405 Orsay, France

Anastasia Bezerianos[2,1]
anab@lri.fr
[2]Univ Paris-Sud & CNRS (LRI)
F-91405 Orsay, France

Thibaut Jacob[1,2,3]
thibaut.jacob@telecom-paristech.fr
[3]Telecom ParisTech & CNRS (LTCI)
F-75013 Paris, France

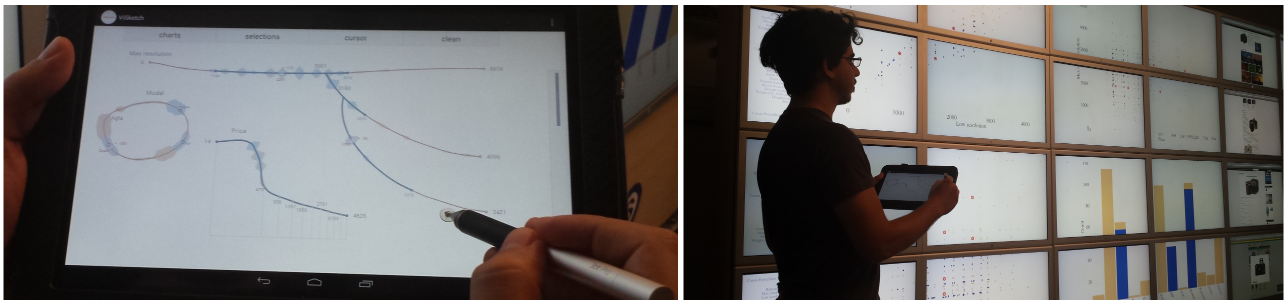


Figure 1. *Sketch*Sliders (left) allow users to directly sketch visualization controllers to explore multi-dimensional datasets. We explore a range of slider shapes, including branched and circular, as well as shapes that express transformations. *Sketch*Sliders control visualizations on a wall display (right).

ABSTRACT
We introduce a mobile sketching interface for exploring multi-dimensional datasets on wall displays. We demonstrate the idea of *Sketch*Sliders, range sliders that users can freely sketch on a mobile surface to customize their exploration. A small combination of sketches and gestures allows the creation of complex interactive sliders, such as circular sliders for periodic data, slider branches for detailed interaction, and fisheye transformation sliders. We augment sliders with a suite of tools, such as markers, slider cursors, and approximate views of data distributions. Our designs are inspired by a design study with three visualization experts and validated through a user study with six experts using our system. Our findings indicate that our sketching interface accommodates a wide range of exploration strategies, helping users customize as well as focus their visual explorations.

Author Keywords
Data visualization; sketching interfaces; wall displays

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

INTRODUCTION
High-resolution wall-sized displays [3, 6, 24] allow users to view a large amount of visual information, and thus have applications in a wide range of domains related to visual data analysis and exploration. Nevertheless, choosing appropriate techniques to explore data in such environments is not a simple matter. Viewers are often mobile, moving away to get an overview of complex visuals, and coming up-close to see details [3]. Exploring complex datasets also requires access to a large number of interactive controls in order to manipulate multiple dimensions and adjust their visual parameters.

We introduce a mobile sketching interface (Figure 1) that decouples control and visualization in a wall environment. Instead of having users interact with a large set of predefined exploration widgets, we let them customize their exploration by *sketching the controllers* that best suit their needs. To demonstrate the approach, we focus on range slider controllers, *Sketch*Sliders. We show how with a small gesture vocabulary users can: creatively use sketching to adjust exploration properties, such as precision, by drawing controllers of various sizes and shapes; focus on parts of the data by changing the control resolution in dense areas of data; explore variations of controllers by grafting alternative paths; and bookmark important results and points. Due to the nature of sketching, users can naturally customize the controller's appearance and its effect on the exploration. For example, they can draw larger sliders and branches for a finer control, circular sliders for periodic data, or shapes that describe transformation functions to focus on a smaller range of the data.

Our designs were inspired by design sessions with three visualization experts, whose feedback reinforced our choice of sketch-based range sliders as a versatile controller. Our sketching interface was then evaluated by six visualization experts. Our findings verify the flexibility of sketching controllers. They also demonstrate how the approach can accommodate different visual exploration strategies, and effectively support data exploration in interaction with wall displays.

Theophanis Tsandilas, Anastasia Bezerianos, and Thibaut Jacob. SketchSliders: Sketching Widgets for Visual Exploration on Wall Displays. In CHI'15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 3255-3264, ACM, April 2015.

http://dx.doi.org/10.1145/2702123.2702129.

Our main contributions are:

- We allow users to sketch directly the interactive controllers they require to conduct data exploration. We illustrate the potential of the approach with range sliders.
- We design mechanisms for sketching interactive sliders of arbitrary forms, including circular, branched and transformation sliders, that support complex queries over multiple data dimensions and multiple levels of control granularity.
- We introduce an environment that combines visualization plots on a wall display with mobile devices for sketching controllers. We study how experts use such a setup.

RELATED WORK

The search for appropriate interactions for wall displays has been largely investigated in HCI, focusing on questions mostly related to specific tasks, such as pointing [24] and maintaining awareness [6]. Recent work has looked at the use of mobile devices as an interaction medium that supports mobility. For example Smarties [10] customizes programmatically mobile interfaces to control a wall, while Jansen et al. [18] combine mobile devices with tangible controllers to explore visualizations. Our approach differs, as we allow users to customize their exploration interface using sketching.

Sketch-based input has been used among others for 3D modeling [17], note taking [14], and for domain specific applications such as MathPad2 [21] and Mus*ink* [28]. We focus here on two relevant research directions: using sketching to create interfaces, and sketching specifically for visual exploration.

Sketching to Create Interfaces

Wong [33] explains that sketching by hand to prototype interfaces allows users to ignore graphic details, focusing instead on the main goal and nature of the interaction. This is of special interest in data exploration, as often there is no clear goal (and thus an ideal interface), and hypotheses are formed progressively as users generate insight into their data [26].

SILK [20] was the first to allow users to sketch several interface controllers (buttons, sliders), and make them active by automatically detecting them and generating code. It relates to our work, as interface components are sketched, but are then "beautified" by the system, losing their informal and custom look. Given Wong's suggestions [33] and comments from our participants, this custom drawing and appearance is important in visual exploration.

Monet [23], allows users to sketch items and their states to create continuous widgets, keeping their informal look. They explore how to define new widgets and their states by example, while we examine how to customize existing widgets through sketching to aid visual exploration.

Sketching in Information Visualization

Sketching has been used in visualization as a visual rendering style (e.g. [34]), but less work has been done on using sketching as input for data exploration.

QuerySketch [30] and QueryLines [25] allow users to search and query data by sketching the desired result of their queries in the form of a graph. Relaxed selection techniques [16] go further by providing ways to implicitly define the level of similarity between sketches and real data. These approaches target mainly time-series data, focusing on searching for predefined patterns, rather than open exploration as we do.

In Transmogrification [7] selected regions from a 2D visualization can be transformed into a destination shape defined by sketching. This technique does not focus on interactive controllers. It rather acts directly on the visualization. Nevertheless, it relates to our work, as we also explore how an arbitrary slider shape can deform a visualization.

SketchStory [22] uses a small set of sketched gestures and touch interactions to construct a story for presentation and communication purposes. This tool does not support data exploration per se, it is rather a tool for helping play back the results of data exploration to others.

NapkinVis [9] uses sketching gestures to rapidly generate different types of charts, focusing mainly on data visualization authoring, not exploration. More relevant to our approach is SketchVis [8] that allows users both to load and to create visualizations using sketching, but it also provides simple ways of conducting data exploration. Users can switch between different data views, select specific data categories, or apply simple functions such as averages or maximum. SketchVis was later used in a Wizard of Oz study [29] to better understand the interplay between touch and pen interactions. Our work delves into using sketching for more complex query types during exploration, focusing in detail on one type of interactive controller, range sliders.

GOALS, CONCEPT AND DESIGN SESSIONS

Visual exploration and analysis tasks [2], such as filtering and range-selection, consist of selections and dynamic queries [1]. These are predominantly performed using slider widgets, each representing one data dimension. Our goal was to combine the simplicity of such widgets with the expressive power of sketch-based data exploration. We were particularly interested in identifying meaningful roles that shaped controllers can take, and explore how users could make use of sketching to augment controllers with new functions.

To explore this direction, we recruited three information visualization experts outside the research team, and ran an hour-long semi-structured design session with each of them. Their experience in visualization design and exploration ranged from 5 to 10 years. We first explained the main goal of the session, and then gave them three examples of sketch interaction, namely how to draw a straight slider and define new ranges to filter a data set, and how to sketch a data transformation function to acquire overview or details. These are fundamental tasks in visual exploration [27].

The experts then explored a multi-dimensional dataset. Their task was left intentionally open to encourage exploration, and was run in a *Wizard of Oz* setting. Participants had to think aloud, explaining what they were sketching on paper and what the expected behavior on the visualization was. One experimenter simulated this behavior on a real visualization

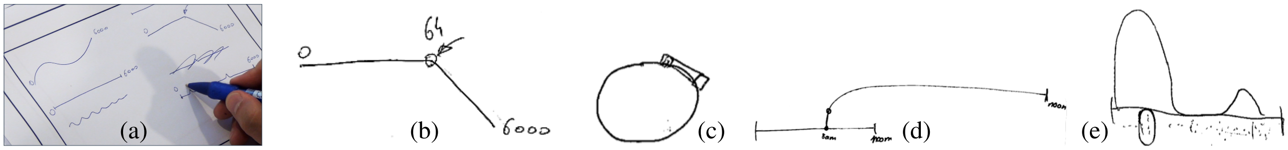


Figure 2. Designs from our three experts: (a) customized sliders, (b) bended slider with a mark, (c) circular slider, (d) grafting a slider, and (e) a slider combining different granularity ranges.

using controllers on her machine. A small brainstorming session followed, to discuss sketch interactions that could benefit data exploration in the context of our experts' own work.

Participants gravitated naturally to variations of sliders as their main interaction tool to narrow down their exploration, further validating our choice to focus on this controller. They all felt that sketching sliders for filtering data provided several benefits for visual exploration:

Customization (B1). Each stroke is unique. This aspect of sketching can be used to visually differentiate one slider from another (Figure 2a). Moreover, as the shape of a stroke can be arbitrary, one can encode information inside it, for example bending the slider in a given point of interest (Figure 2b).

Granularity (B2). The length of sketched slider affects the level of its control on the data. Long sliders allow fine-grained control and more precise filtering of data, while smaller ones are only appropriate for coarse-grained control. Two experts commented on how they often need both fine and coarse-grained control, e.g., for timelines that contain periods with both heavy and low activity. Each expert sketched a different solution to this problem. One created a coarse-grained slider and then *grafted* on top of it a second fine-grained slider (Figure 2d). The other expert drew a bulge at the location he required finer control (Figure 2e).

Parametrization (B3). Sketched sliders can support multiple ranges, giving users the possibility to filter the visualization in a discontinuous manner. This is occasionally a constraint in existing predefined range sliders. Another way of parametrizing a slider is to write by hand possible slider extremums or link specific data values to specific locations of the slider, which controls completely the mapping between the slider and the data. All our experts suggested a variation of this.

Special Shapes (B4). Our experts took advantage of the fact that they can sketch sliders in any shape they wanted. One expert sketched a circular slider to be used in the exploration of periodic dimensions such as hours of the day or angle values (Figure 2c). In this way, the slider represents a complete period with no start or end.

Reusability (B5). All participants requested the ability to deactivate controllers while still have them accessible for later use, either "faded-out" on the canvas or in a separate side panel. These solutions ensure that users keep copies of their sketched components, allowing them to explore alternative aspects of their data without losing their past work.

Annotation (B6). All our experts appreciated how a sketching environment naturally supports annotating and bookmarking of important information that is crucial for long term visual analysis tasks. Such an environment can combine notes on the analysis process taken by users, as well as traces of the interaction exploration, e.g., sketched controllers and values, that led to specific insights and findings.

One motivation behind our work is the need for an interface that is mobile, to accommodate users working in front of high-resolution wall displays. Nevertheless, all experts commented that a tool for sketching controllers would be valuable even in desktop environments, where visualization systems are often overloaded with numerous controllers.

SYSTEM OVERVIEW

We designed a sketch-based interface for data exploration that implements many ideas from our design sessions. The interface runs on mobile devices and communicates with a visualization dashboard on a high-resolution wall display (see Figure 1). The dashboard displays plots such as scatterplots and bar charts that present different views of a dataset. Plots are completely synchronized (coordinated views). If data points are selected or filtered, these effects are applied to all plots.

Users can sketch sliders on the mobile interface to explore different data dimensions, and create queries to filter data. Queries are communicated through the network to a server, which propagates the appropriate data filters to a wall cluster for rendering. Figure 3 shows the effect of filters on the two dimensions of a scatterplot. Blue points are data points inside all active filter ranges, while pale orange ones are outside.

Our setup supports personal exploration (tablet) over a shared viewing space (wall). This decoupling serves two main goals. First, users can move freely around the wall display and interact with sliders at the desired viewing distance and level of data overview. Second, large datasets can be comfortably visualized on a visualization cluster without overloading the limited capacities of personal devices. To a lesser extent, the setup accommodates wall displays that lack touch support.

Interacting with Gestures

Our sketch-based interface runs on Android devices, tablets and smartphones. Its design combines free writing, sketching of interactive controllers, and interaction with gestures. Since our interaction model only requires the use of a passive stylus or a finger, we combine stroke delimiters [13], crossing-based selection [4], and simple gesture recognition [32] to correctly infer the type and function of pen strokes.

Delimiters. We use dwells as delimiters to identify special command-strokes. As seen in Figure 4, a dwell invokes a contextual menu with a list of possible actions, e.g., create a

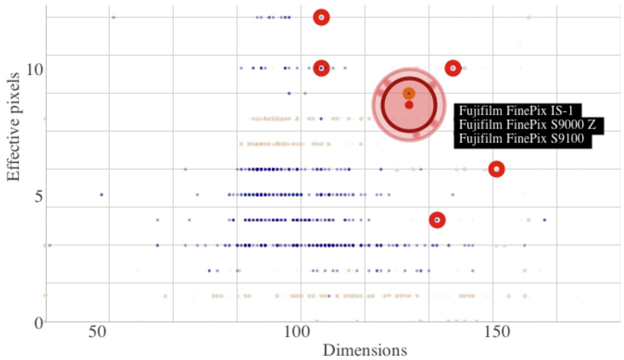


Figure 3. A scatterplot displays filtered points that represent camera models. Blue points are inside the active filter ranges, pale orange points are outside. A red area cursor can point to, and select, blue data points. Circled points correspond to previous selections. The dots in the ring around the cursor communicate the relative position of selected points.

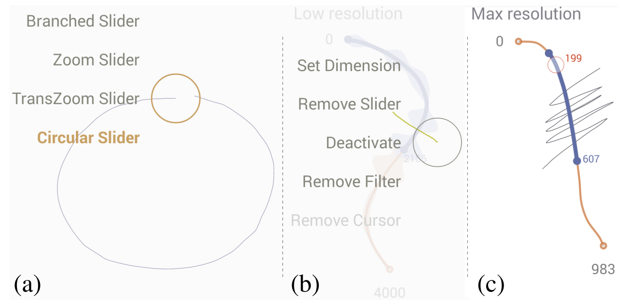


Figure 4. Combining crossing, dwelling, and gesture recognition. (a) Dwelling after drawing a curve activates a contextual menu that lets the user create a new slider. In this example, the system has recognized a circular form and therefore suggests a "Circular Slider" in addition to other types. (b) Dwelling after crossing a *Sketch*Slider shows a contextual menu with a set of possible actions. (c) Scribbling out a filter.

new slider. We found that short dwells of 250 to 350 ms do not disrupt the flow of regular writing. Hinckley et al. [13] report that dwells result in fewer errors but are slower than pigtails. But since pigtails interfere with symbols of regular handwriting, we only use them in conjunction with crossing.

Crossing-Based Selection. Users can cross existing interactive components and then dwell to choose a function from a context menu, e.g., deactivate or activate a slider, remove a filter or a marker, and change a filtering dimension.

Gesture Recognition. We facilitate interaction by recognizing three special forms of gestures: zig-zag scribbles, circular strokes, and pigtails. Zig-zag gestures serve as erasers of individual or groups of strokes as well as erasers of interactive slider components, e.g., whole sliders and filters. Circular strokes are generally considered as candidates for periodic sliders unless their trace crosses a slider, in which case they create slider cursors. Finally, pigtails create filters. We give more details about these features in the next section.

Pointing on the Wall Display and Data Selection

Users can activate additional functionality through bezel and contextual menus to reconfigure the plots on the wall display. They can also interact directly with their filtered content by turning the mobile device into a touch pad. As shown in Figure 3, we support pointing and selection through a circular area cursor with *excentric labeling* [12], for previewing and selecting data points directly on the plots. Users can pinch with two fingers to resize the area cursor and reduce or increase the active area of selection. We use the GlideCursor [5] and the acceleration function of Nancel et al. [24] to control the cursor position on the wall display.

Implementation

Our application has been built on Java for Android 4.3. The visualization software on the wall runs on Processing 1.5 - 2.0 and uses the Most Pixels Ever library[1] for spanning the charts on multiple machines and screens on the wall display, and a modified giCenter Utilites library for the charts[2]. Our wall consists of 32 30-inch LCDs arranged in a 8×4 matrix (size 5.5 x 1.8m and effective resolution 20480 x 6400 pixels) and is driven by a cluster of 16 computers. A Java server takes care of the communication between the Android and the wall application through the Open Sound Control protocol[3].

SKETCHSLIDERS

As regular sliders, *Sketch*Sliders control specific dimensions of a dataset, but can take arbitrary shapes. We have explored a number of designs that take advantage of their free form, including periodic sliders, grafted sliders, and transformation sliders whose path describes a fisheye function. *Sketch*Sliders can serve as classic controllers, but also as alternative mechanisms for direct data exploration. As we support embedded data distributions, slider cursors, and markers, users can focus on the tablet to complete part of their tasks without having to frequently move their attention to the wall display.

Basic SketchSliders and Filters

After drawing the path of a *Sketch*Slider, a popup dialog asks the user to assign a data dimension. We support both ratio and ordinal dimension variables (numerical, or textual sorted in alphabetical order). For ratio variables, we differentiate between decimals and integers. We also differentiate between periodic, e.g., months, and non-periodic variables (Figure 5).

When first created, a slider extends between the two extrema values of the dimension's active range. Users can interactively change the active range (*B3*) by long-pressing the label of the starting or ending extremum value, and then dragging it leftwards or rightwards. This decreases or increases the associated value, but also updates the range of the data dimension on the plots of the wall display. This way, users can control the charts by zooming in within a subrange of a dimension.

Figure 5 shows examples of *Sketch*Sliders and explains how users interact with their widgets. Sliders can host one or more interactive filters (*B3*), that define the union of ranges or individual values. Filter ranges are shown in blue. We support *delta filters* and *range filters*, created with pigtailed strokes that cross either once or twice the path of a slider. A delta

[1] http://github.com/shiffman/Most-Pixels-Ever-Processing
[2] http://www.gicentre.net/software/#/utils/
[3] http://opensoundcontrol.org

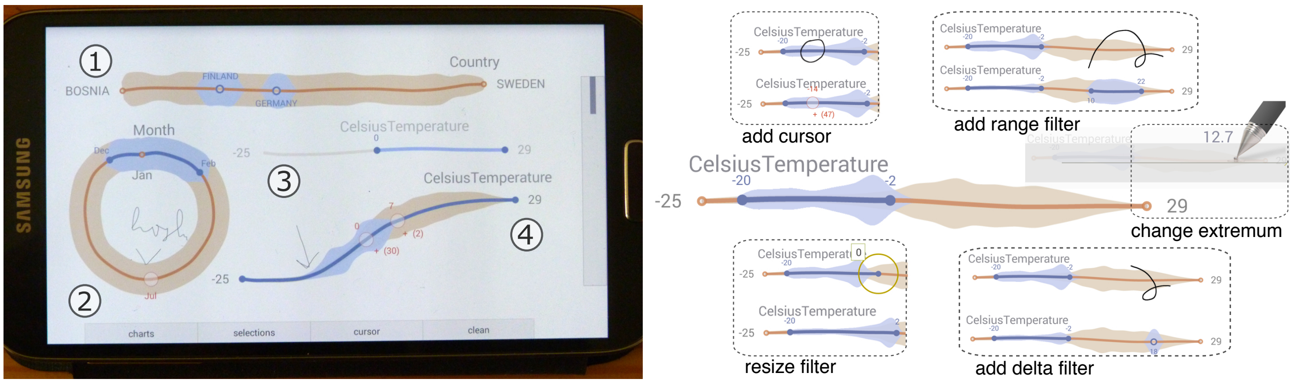


Figure 5. Left: Four *Sketch*Sliders on the screen of a 5-inch Samsung Galaxy S4: (1) a slider over an ordinal variable (Country), (2) a circular slider over a periodic variable (Month), (3) a slider that is currently inactive, and (4) a slider over a ratio variable (Temperature). The query of the three active sliders selects winter mean temperatures for Finland and Germany. Slider 4 shows how the distribution of active temperatures (blue) leans towards lower values. Right: Gestures and interactions to create and manipulate basic slider widgets: add a cursor (crossing circle), add a range or a delta filter (crossing pigtail), change a slider extremum, and resize a filter. Observe how active (blue) density distributions change in response to these actions.

filter has a single control point and represents either a unique value (ordinal variables), or a small delta range around a value (ratio variables), whose precision depends on the slider size and range of values. Range filters have two control points, and determine an active range between two values. Users can manipulate the control points to change the start and end values, or pick and move the entire range along the slider.

For complex queries involving multiple data dimensions, users can create multiple *Sketch*Sliders. Their result is the intersection of the results of the individual sliders' queries. During data exploration, users may draw several sliders and switch between them by deactivating ones or activating others (*B5*). Our implementation allows for a single active slider for each dimension. Sliders become automatically inactive when a new slider for the same dimension is drawn.

Guided by the results of our design study (*B4*), we support circular sliders dedicated to periodic variables (Figure 5: Left). Circular sliders have no ends, and thus do not constrain the translation of their filters. For example, a single range filter can specify the winter months, from December to February.

Embedded Density Distributions

We augment *Sketch*Sliders with density distributions that appear as shadows along the path of each slider, and communicate data overviews. We get inspiration from *scented widgets* [31], but opt for a different visualization approach that derives from *violin plots* [15]. Violin plots show a density trace that extends symmetrically along the length of a box plot. The density trace provides rich information about the underlying distribution, and when compared to histogram-based visualizations [31] it results in less visual clutter. It can also better generalize to curved paths of arbitrary shapes.

As with histograms, this approach requires the selection of an interval (bin) width h. We try to optimize h by considering the type of the slider's variable (ordinal, integer, decimal) and its range. For a given range, longer *Sketch*Sliders contain a larger number of intervals, and thus show more detailed density distributions. Density distributions are normalized between zero and the highest available density. We visualize two distributions: (1) a static distribution of the entire dataset within the slider's range (in light orange), and (2) a dynamic one of the active data subset, defined by the filters of the currently active sliders (in blue). A dynamic distribution is updated every time the user adds, removes or changes a filter on an active *Sketch*Slider. This allows users to directly observe the effects of their filters on multiple dimensions.

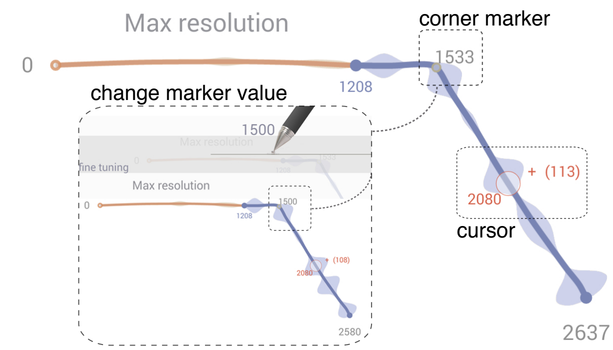


Figure 6. Corners can serve as markers of values of interest. As with extrema, the user can adjust the value of a corner by long-pressing and dragging its label along the horizontal axis. A slider cursor reveals both the value of its center and the number of active data points it covers.

Density distributions are approximate but can support data exploration directly on the mobile surface. For instance, users can draw multiple sliders to get quick information about how data points are distributed along different dimensions, and identify areas of interest within the sliders' ranges, common tasks in data analysis [2]. Thus, users can focus on their actual filters without having to continuously shift their attention to the wall display. Or, the distributions can help users identify functional dependencies between dimensions as they manipulate filters of different sliders. For example, the Temperature slider in Figure 5 (Left) shows that registered winter temperatures for Finland and Germany lean towards the low range of values and include a high proportion of negative values.

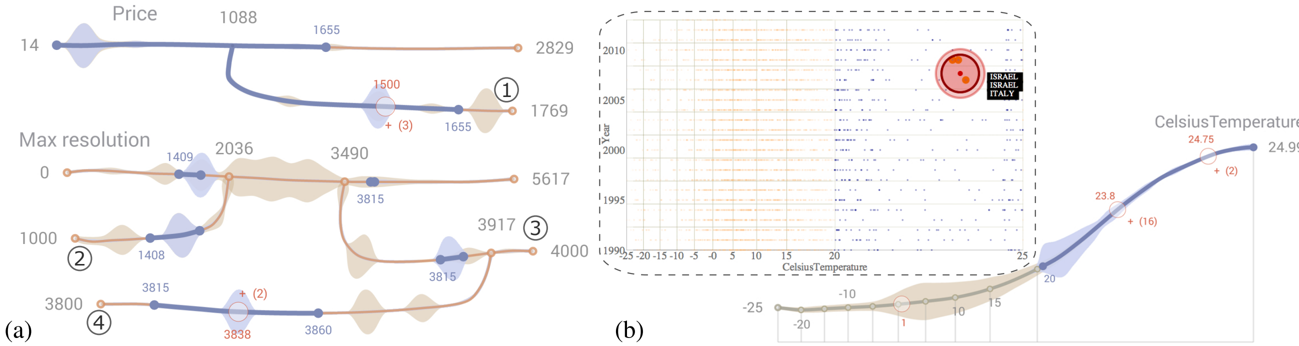


Figure 7. (a) Branched *Sketch*Sliders. Branches (1, 2, 3, 4) serve as proxies, providing a more detailed view of a range. A branch can start from (1, 3) or end at (2, 4) the main slider or another branch. (b) A Transformation *Sketch*Slider applies a fisheye deformation to a scatterplot. The slider contains grid lines that correspond to the transformed grid of the scatterplot. In this example, the transformation allows for zooming into the range of temperatures between 20 and 25°C. For this range, the user can make more precise selections on the scatterplot and the slider itself.

Slider Cursors and Markers

A *Sketch*Slider can host one or more *slider cursors* that serve as one-dimensional navigation widgets. Slider cursors can move along the length of a slider and display the data value that corresponds to its centre, aiding value retrieval tasks [2]. Similarly to the 2D area cursor on the wall charts, they also serve as data-point selectors. As seen in Figure 6, a slider cursor covers a delta area over the slider's range, and counts the active data points queried by the filters. Users can tap on the $+$ symbol to display the list of the covered data points, and optionally, add a subset of these points to their selections. The granularity of the cursor depends on the length of the *Sketch*Slider and its range of values.

Since *Sketch*Sliders can take any shape, distinctive features can act as *markers* of values of interest, as suggested in our design study (*B1*). We enhance this property by running a corner detection algorithm that automatically identifies corners on the slider's path and displays their value. As with extrema values, users can long press on a marker label to further adjust its value. Figure 6 shows that changing the value of a marker shifts the extreme values (one or both) of the slider and adjusts its scale, ensuring that the slider remains linear.

Grafting: Branched Sliders

We apply the concept of grafting from our study (*B2*), in a new slider type, *branched* sliders. Branched sliders support infinite nesting of branches to allow multiple levels of granularity within a slider's range. Grafting a branch is similar to creating a new *Sketch*Slider, but the stroke for the new branch has to cross an existing slider or branch. The joint point can be either the starting or the ending extreme of the branch, depending on the direction of the stroke that creates it. Thus, branches can form arbitrary tree and polytree structures. We also allow for closed branches between any two points of the root slider or a branch. Note that the extrema values of a branch cannot exceed the extrema values of its parent.

As seen in Figure 7a, branches act as slider proxies, and filters have copies in all the branches of a slider. The user can manipulate any of these copies according to the level of intended precision. By grafting branches and adjusting their extrema, users can zoom in partial ranges of the slider. The density distribution of a branch reflects its individual range rather than the range of the root slider, offering a different view of the data. As users zoom in a subrange, distributions get more fine grained, while filters and cursors become more precise. This can reveal distribution anomalies or clusters, that are important components of data analysis [2]. For example, branches 1, 3, and 4 in Figure 7a reveal small clusters of data points not visible on the root slider. Notice how the filter on branch 4 has increased in length and granularity compared to its original copy on the root slider. Branched sliders let users dig incrementally into ranges of a dimension, without removing the trace of previous explorations.

Transformation Sliders

Sketching is commonly used for drawing curves to communicate trends or mathematical functions. Inspired by our study (*B2*), the path of *transformation* sliders similarly describes an 1D transformation function. To define transformations over slider paths, we use a curvilinear l-y coordinate system, where l is the arc length of the partial curve at point $p(x, y)$. This approach overcomes the problem of curves that do not describe valid functions in Cartesian coordinates.

We examine *focus+context* transformation functions that affect the dimensions of plots on the wall display. More specifically, peaks of a slider curve represent areas of focus while valleys represent areas of context. To model this behavior, we define a fisheye function $F(y, g) = (1 - g)f(1) + gf(y)$, where $y \in [0, 1]$ is the vertical curve position normalized between the higher and lower extreme of the curve, while $g \in [0, 1]$ is a gain function. In our implementation, the gain is $g = h/h_{max}$, where h is the height of the curve and h_{max} is a maximum height value. Finally, f is a monotonically increasing function. We produce a more aggressive fisheye deformation by taking an exponential function $f(y) = a^y$.

Figure 7b shows a *Sketch*Slider that applies a fisheye transformation to the $x-$axis of a scatterplot. We can see that the transformation applies to the scatterplot but also to the slider itself. This means that values are not uniformly distributed along the path of the slider, i.e., they are sparser around peaks and denser around valleys. Similarly, the granularity of slider

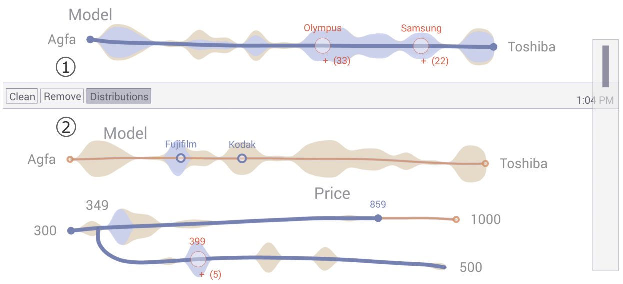


Figure 8. A document can consist of multiple sections, where each section (1, 2) forms a unique query. A movable bar (separator) denotes the bottom of a section and includes functionality for managing the section.

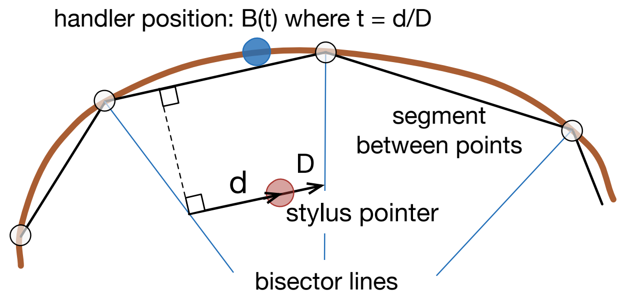


Figure 9. Deriving the position of a handle (blue circle) on the cubic Bézier curve of a slider from the position of the pen (red circle).

widgets varies along the slider. As a result, delta filters and cursors become more precise closer to higher peaks.

To apply a fisheye transformation over the slider's path, we divide it into small arcs and estimate the transformed value $v(l) \in [0, 1]$ at position l as the sum of its partial delta values:

$$v(l) = \sum_{k=1}^{i-1} \Delta v(l_k, l_{k+1}) + \Delta v(l_i, l) \qquad (1)$$

We then approximate the fisheye transformation value for a position s within each arc i as a linear function $F(s) = F_i + \lambda_i s$, where F_i is its value at the start point of the arc and $\lambda_i = (F_{i+1} - F_i)/(l_{i+1} - l_i)$. Based on this approximation, we can calculate the delta values in Equation 1 as follows:

$$\Delta v(l_k, l) = \beta \int_{l_k}^{l} F(s)ds = \beta(F_k \delta l + \frac{\lambda_k \delta l^2}{2}) \qquad (2)$$

where $\delta l = l - l_k$. The parameter β is a normalization factor derived from Equations 1 and 2 by setting $v(L) = 1$, where L is the total slider length.

Exploration Sections

A user can divide an exploration session into sections, where each section can contain any number of sliders that form a query. Sections let users start a new exploration without losing the history of their previous queries (*B5*). New sections are created by drawing a long horizontal line with a pigtail. This creates a *separator* (Figure 8) that serves as the lower boundary of a section. The widget can be dragged up or down to adjust the size of two adjacent sections.

CURVE MODELING AND DRAGGING

We model the path of *Sketch*Sliders and range filters as cubic Bézier curves. We uniformly decrease the number of path points, which results in smoother curves and better performance. We also use cubic Bézier curves to model the trace of distributions, by connecting density points between bins at the normals of the *Sketch*Sliders. If the lines of two neighboring density points intersect, we consider a single density point at their intersection, avoiding the creation of path loops.

Figure 9 explains how we derive the position of a handler, e.g., a slider cursor, on the Bézier curve from the position of the finger or stylus pointer when dragging the handler. We find the line that crosses the pointer and is parallel to the closest segment between two consecutive path points. We then take the vector $\vec{D}$ defined by its intersection with the previous and the next bisector of the between-segment angles. We also take the position $\vec{d}$ of the pointer on this vector. From this, we derive the t parameter of the corresponding Bézier curve $B(t)$ as the normalized pointer position $t = d/D$. To deal with slider curves that contain closed loops, we can set a threshold distance that avoids accidental jumps between non-neighboring segments.

In contrast to our approach, DimpVis [19] simply takes the minimum-distance point to project the position of the finger to a path. Dragicevic et al. [11] have proposed a more generic solution by modeling curvilinear dragging as an optimization problem that minimizes a 3D instead of a 2D distance. Our solution is less generic, yet fast and simpler, while it still results in very smooth behavior for meaningful slider shapes.

USER STUDY

To validate our prototype we conducted a user study with six visualization researchers (2 female), who used regularly visual analytics tools, recruited through chain-sampling. Their experience in visualization research and analysis ranged from 2 to 15 years (median 10), and their age from 24 to 41 (median 34.5). All reported being highly familiar with the visualization plots used in our study and the notion of range sliders. Three had taken part in our early design study.

Goals: The goal of the study was threefold: (*G1*) verify the findings from the design study regarding the value of sketching controllers for visual exploration; (*G2*) observe how experts can appropriate *Sketch*Sliders; and (*G3*) examine the benefits and shortcomings of divided attention. We wanted to observe how the sketching interface supports visual exploration in combination, but also in separation, from the wall.

Setup: Participants were seated in front of the wall display, which displayed three scatterplots and one bar chart. Participants interacted with the plots using a Nexus 10-inch tablet running our prototype (Figure 10). We also provided a capacitive pen. We chose a seated setup, as opposed to free walking, to ensure user comfort for the duration of the study, and to avoid any confounds due to uncomfortable hold positions of the tablet without proper hand support.

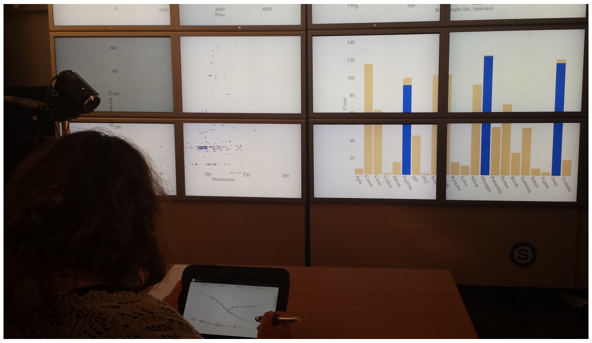


Figure 10. Study setup with wall display and *Sketch*Sliders on a tablet.

T1	Tell us your observations about the distribution of Prices (e.g. extremes, clusters, similarities or gaps inside clusters at low detail).
T2	Which camera Models are in the Price range 3000-6000? Approximate as much as you can the most frequent Max resolution values for them.

Table 1. The two open tasks. T1 was conducted by using both the wall and the tablet. For T2, participants were asked to not look at the wall.

Procedure and Tasks: Each session lasted approximately one hour and was videotaped. Participants were given a detailed training session of 20-30 minutes, using a dataset on the history of temperatures in different countries. They then performed alone, using a think-aloud protocol, two open exploration tasks on another dataset. The two open tasks (Table 1) were related to a multidimensional camera dataset with a total of 1038 digital cameras and 13 dimensions. For the second task, participants were asked to not look at the wall. We expected that this would encourage them to explore the capabilities of *Sketch*Sliders in more depth, and assess their strengths and limitations both with and without the wall. The session was followed by a semi-structured interview.

Results

We report here on observations gathered during the study and participants' responses during the interview.

G1: Value of Sketching in Visual Exploration

All participants were very enthusiastic about sketching their own controllers as needed. They each commented on different aspects, which we briefly summarize here:

Customization. Participants reinforced findings from our design study, stating that sketched sliders can be *"customized to what I need"*, *"easy to tell apart"*, and *"feel personal"*. As one participant mentioned *"there is something very compelling about sketching your own tools"*.

Personal space. All participants could envision sharing the wall plot result but would be reluctant to share the sliders themselves. As one participant mentioned, *"it is similar to sharing my notes, I would be reluctant, it is too personal"*.

Flexibility. All participants appreciated the flexibility of sketching different shapes and sizes to: (i) better fit their view of the data, *"this is vertical [indicating a dimension on a plot] so I drew a vertical slider"*; (ii) focus on parts of the data as needed, *"I can focus either with branches or transformation, you don't have that in other interfaces"*; and (iii) generally allow users to feel that the interface can accommodate their needs. As one participant stated *"existing interfaces are not flexible enough, they only show you predefined controllers"*.

Annotation. All participants mentioned they liked the ability to be able to sketch controllers and also take personal notes related to their analysis and findings (see Figure 11a,b). However, three participants stated that they would have preferred a different mode for analysis and a different one for annotation, as in some cases they wanted to use gestural marks, such as circling a slider, as a free-form note to indicate emphasis.

G2: Use of Sketch*Sliders functionality by Experts*

Participants were able to effectively use the interface without aid, adopting different strategies to perform their tasks and using different combinations of *Sketch*Sliders' functionalities.

Branches. Five participants used extension branches to increase the granularity of their exploration. As P1 mentioned *"I draw a branch when I couldn't interact precisely [with the slider], and adjusted its end to be more precise"*. Branches were also used to compare different parts of the data. P2 explained that *"I made a second branch to see if I have the same detailed pattern as in the other [branch]"*. P4 combined branches with other strategies to explore different aspects of the data. She drew multiple sliders for the same dimension and adjusted their ranges. She then compared their effect on the wall plots by switching between them. As the activation and deactivation of sliders happens through a contextual menu, she asked for a faster way to do so.

Transformation sliders. Four participants used transformation sliders to focus on part of their data. As P6 mentioned, this gave *"a better view of the densely packed data points in this range"*. Figure 11a shows how P6 used a transformation slider to focus on the lowest range of camera prices, while he filtered release dates with a regular slider. This participant did not use any branches, as he could complete the tasks with transformation sliders by adjusting their extremes. Others used transformation sliders early on to get a quick view of clusters in their data, or to focus on some interesting areas of certain dimensions, while using branched sliders for others.

Slider cursors. Participants made a frequent use of the slider cursor and found it very useful. Some created a large number of these cursors (Figure 11) to locate items within an area of the slider and then display them as selected on the wall. Two participants commented that they would have liked to change the cursor's size, as they could do with the wall cursor.

Density distributions. All participants mentioned that slider distributions were helpful to *"identify possible high level patters"*, *"get a quick overview of data"*, or to *"quickly see how much data are filtered"*. P1 stated that the distribution was enough on its own for a high level and approximate data exploration. On the other hand, P4 and P5 suggested showing individual points rather than distributions at fine granularity, as they *"expected clusters not points, and was not sure if I should focus more"*. P4 observed that the rendering can be misleading, expecting that data are normally distributed within the range covered by bins. A possible solution to this problem is to use beeswarm-like visualizations[4] when the number of data points around a bin becomes very low.

[4]http://www.cbs.dtu.dk/~eklund/beeswarm/

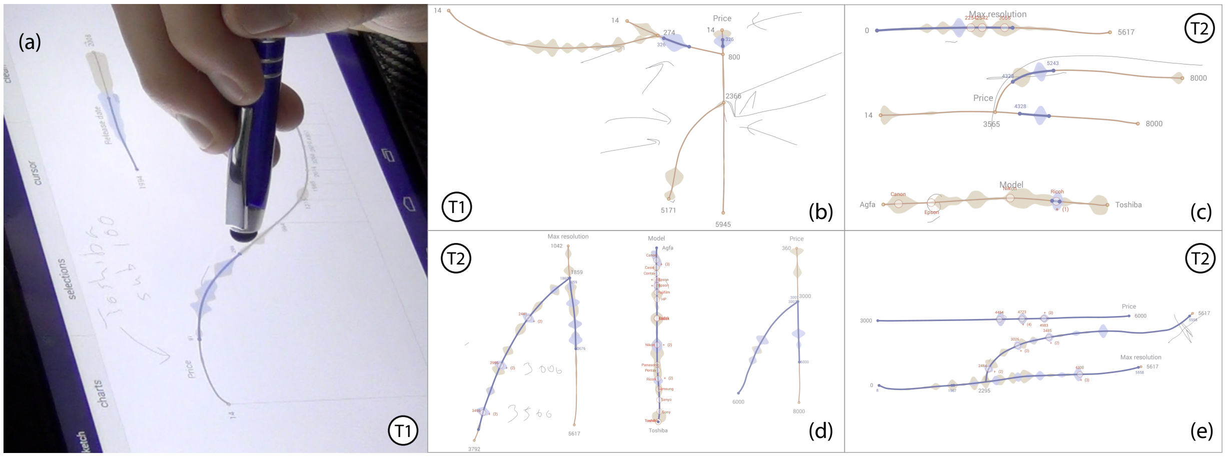


Figure 11. Complex queries sketched by different participants of our study. P6 applies a filter to lower prices by using a transformation slider (a). Use of a mix of simple and branched *Sketch*Sliders by P2 (b, d), P3 (c), and P4 (e). We indicate the task (T1, T2) for which the query was created.

G3: Exploration using Sketch*Sliders and Wall Display*

All participants mentioned that the combination of sketching controllers and decoupled wall plots worked really well for visual exploration and were *"very well integrated"*. According to P5, *"the setup really works well for me, they are complementary, as you cannot show all the information on the tablet, and I don't really want to directly interact with the wall all the time"*. Observing how participants worked on task T1, we were able to identify the following strategies.

Participants found that sketching one controller at a time helped their analysis process. P1 used almost exclusively the tablet, as plots on the wall *"are at first overwhelming, too much data"*. She explained that sketching sliders as needed let her *"focus on one dimension, and see the result on other dimensions [using the slider distributions], which reduced my cognitive load"*. She explained that the interface *"helped me filter not just my data, but also my controllers"*.

A similar sentiment was expressed by P3, who explained that *"existing interfaces are too cluttered and it is hard to decide where to focus on, here I can focus on one thing, and drawing a slider is a way for myself to decide what will be most likely of interest next"*. It is interesting to note that P3 and P6 completed the task by focusing mainly on the wall, manipulating their sliders eyes-free. According to P6, this helped him *"avoid splitting attention"* between the two displays.

P2, P4, and P5 divided their analysis between tablet and wall. They used slider distributions for approximate answers, and the wall for detailed ones. P4 mentioned she always *"checked to see the result [of a slider] on the plots"*, while P2 explained that *"when I wanted detail I would go here [plots]"*.

Overall, participants were very enthusiastic, and were able to retain and use the interface successfully. Their choices on what controllers to create, and how to explore data using the tablet and wall varied greatly. Clearly experts have their own analysis approaches, and one interface does not fit all. Given the varied ways our setup was used, we feel *Sketch*Sliders are flexible enough to support different working styles.

CONCLUSION AND PERSPECTIVES

We presented *Sketch*Sliders, range sliders that users can freely sketch directly on a mobile, in order to parametrize and customize their exploration of data on a wall display. With a small combination of sketches and gestures users can create complex interactive components, such as slider branches and data transformation sliders, to investigate detailed aspects and subsets of their datasets. Apart from their natural custom shape, our sketched sliders are enhanced by interaction aids such as slider cursors, markers and distribution visualizations.

Results from a user study with visualization experts indicate that *Sketch*Sliders are flexible, support different exploration strategies, while the fact that they are sketched as needed can focus and aid the visual analysis. Our system currently does not support collaboration. Nevertheless, participants indicated that the sketched sliders are very personal, and although they would be willing to share their results on the wall visualization, they are less open to sharing the controllers themselves. This opens interesting questions regarding sharing mechanisms and privacy that need further study.

We were motivated by a scenario where users view large datasets on wall displays, and require mobile interaction support. Nevertheless, our experts commented that sketching customized controllers can be useful even in desktop settings, where visualizations are usually laden with numerous inflexible controls. We plan to investigate further how users can sketch and use sliders directly on their visualizations.

One benefit of *Sketch*Sliders is the combination of regular note taking that is crucial in data exploration, and the trace of the interactive controllers and their values that were used to reach insights. This coupling of analysis process and recording could be further enhanced. We could envision users taking snapshots of their data to add to their sketched environment, and detailed history mechanisms for reverting to previous steps in their exploration. These "traces" of the interaction exploration could then be used as a logging mechanism for analysts, or shared with others without sharing the data themselves that could be too large or sensitive.

REFERENCES

1. Ahlberg, C., Williamson, C., and Shneiderman, B. Dynamic queries for information exploration: an implementation and evaluation. *CHI '92*, ACM (1992), 619–626.
2. Amar, R., and Stasko, J. A knowledge task-based framework for design and evaluation of information visualizations. *Symposium on Information Visualization*, INFOVIS '04, IEEE (2004), 143–150.
3. Andrews, C., Endert, A., Yost, B., and North, C. Information Visualization on Large, High-resolution Displays: Issues, Challenges, and Opportunities. *Information Visualization 10*, 4 (2011), 341–355.
4. Apitz, G., and Guimbretière, F. Crossy: A crossing-based drawing application. *UIST '04*, ACM (2004), 3–12.
5. Beaudouin-Lafon, M., Huot, S., Olafsdottir, H., and Dragicevic, P. Glidecursor: Pointing with an inertial cursor. *AVI '14*, ACM (2014), 49–56.
6. Bezerianos, A., Dragicevic, P., and Balakrishnan, R. Mnemonic rendering: an image-based approach for exposing hidden changes in dynamic displays. *UIST '06*, ACM (2006), 159–168.
7. Brosz, J., Nacenta, M. A., Pusch, R., Carpendale, S., and Hurter, C. Transmogrification: Causal manipulation of visualizations. *UIST '13*, ACM (2013), 97–106.
8. Browne, J., Lee, B., Carpendale, S., Riche, N., and Sherwood, T. Data analysis on interactive whiteboards through sketch-based interaction. *ITS '11*, ACM (2011), 154–157.
9. Chao, W. O. NapkinVis: Rapid Pen-Centric Authoring of Improvisational Visualizations. *IEEE InfoVis Poster* (2009), 1–9.
10. Chapuis, O., Bezerianos, A., and Frantzeskakis, S. Smarties: An input system for wall display development. *CHI '14*, ACM (2014), 2763–2772.
11. Dragicevic, P., Ramos, G., Bibliowitcz, J., Nowrouzezahrai, D., Balakrishnan, R., and Singh, K. Video browsing by direct manipulation. *CHI '08*, ACM (2008), 237–246.
12. Fekete, J.-D., and Plaisant, C. Excentric labeling: Dynamic neighborhood labeling for data visualization. *CHI '99*, ACM (1999), 512–519.
13. Hinckley, K., Baudisch, P., Ramos, G., and Guimbretiere, F. Design and analysis of delimiters for selection-action pen gesture phrases in scriboli. *CHI '05*, ACM (2005), 451–460.
14. Hinckley, K., Zhao, S., Sarin, R., Baudisch, P., Cutrell, E., Shilman, M., and Tan, D. Inkseine: In situ search for active note taking. *CHI '07*, ACM (2007), 251–260.
15. Hintze, J. L., and Nelson, R. D. Violin plots: A box plot-density trace synergism. *The American Statistician 52*, 2 (May 1998), 181–184.
16. Holz, C., and Feiner, S. Relaxed selection techniques for querying time-series graphs. *UIST '09*, ACM (2009), 213–222.
17. Igarashi, T., Matsuoka, S., and Tanaka, H. Teddy: A sketching interface for 3d freeform design. *SIGGRAPH 2007 Courses*, SIGGRAPH '07, ACM (2007).
18. Jansen, Y., Dragicevic, P., and Fekete, J.-D. Tangible remote controllers for wall-size displays. *CHI '12*, ACM (2012), 2865–2874.
19. Kondo, B., and Collins, C. Dimpvis: Exploring time-varying information visualizations by direct manipulation. *Trans. Vis. Comput. Graph. PrePrints*, IEEE (2014), 2003–2012.
20. Landay, J. A. Silk: Sketching interfaces like krazy. *CHI '96*, ACM (1996), 398–399.
21. LaViola, Jr., J. J., and Zeleznik, R. C. Mathpad2: A system for the creation and exploration of mathematical sketches. *SIGGRAPH '04*, ACM (2004), 432–440.
22. Lee, B., Kazi, R. H., and Smith, G. Sketchstory: Telling more engaging stories with data through freeform sketching. *Trans. Vis. Comput. Graph. 19*, 12 IEEE (2013), 2416–2425.
23. Li, Y., and Landay, J. A. Informal prototyping of continuous graphical interactions by demonstration. *UIST '05*, ACM (2005), 221–230.
24. Nancel, M., Chapuis, O., Pietriga, E., Yang, X.-D., Irani, P., and Beaudouin-Lafon, M. High-precision pointing on large wall displays using small handheld devices. *CHI '13*, ACM (2013), 831–840.
25. Ryall, K., Lesh, N., Lanning, T., Leigh, D., Miyashita, H., and Makino, S. Querylines: Approximate query for visual browsing. *CHI EA '05*, ACM (2005), 1765–1768.
26. Saraiya, P., North, C., and Duca, K. An insight-based methodology for evaluating bioinformatics visualizations. *Trans. Vis. Comp. Graph. 11*, 4 IEEE (2005), 443–456.
27. Shneiderman, B. The eyes have it: A task by data type taxonomy for information visualizations. *Symposium on Visual Languages*, IEEE (1996), 336–.
28. Tsandilas, T., Letondal, C., and Mackay, W. E. Musink: Composing music through augmented drawing. *CHI '09*, ACM (2009), 819–828.
29. Walny, J., Lee, B., Johns, P., Riche, N. H., and Carpendale, S. Understanding pen and touch interaction for data exploration on interactive whiteboards. *Trans. Vis. Comput. Graph. 18*, 12 IEEE (2012), 2779–2788.
30. Wattenberg, M. Sketching a graph to query a time-series database. *CHI EA '01*, ACM (2001), 381–382.
31. Willett, W., Heer, J., and Agrawala, M. Scented widgets: Improving navigation cues with embedded visualizations. *Trans. Vis. Comp. Graph. 13*, 6 IEEE (2007), 1129–1136.
32. Wobbrock, J. O., Wilson, A. D., and Li, Y. Gestures without libraries, toolkits or training: A $1 recognizer for user interface prototypes. *UIST '07*, ACM (2007), 159–168.
33. Wong, Y. Y. Rough and ready prototypes: Lessons from graphic design. *CHI '92*, ACM (1992), 83–84.
34. Wood, J., Isenberg, P., Isenberg, T., Dykes, J., Boukhelifa, N., and Slingsby, A. Sketchy rendering for information visualization. *Trans. Vis. Comput. Graph. 18*, 12 IEEE (2012), 2749–2758.

Perception of Visual Variables on Tiled Wall-Sized Displays for Information Visualization Applications

Anastasia Bezerianos and Petra Isenberg

Fig. 1. Two viewers analyzing data visualizations from different viewpoints in front of a large high-resolution wall display (a). A participant conducting a trial during our first experiment (b).

Abstract—We present the results of two user studies on the perception of visual variables on tiled high-resolution wall-sized displays. We contribute an understanding of, and indicators predicting how, large variations in viewing distances and viewing angles affect the accurate perception of angles, areas, and lengths. Our work, thus, helps visualization researchers with design considerations on how to create effective visualizations for these spaces. The first study showed that perception accuracy was impacted most when viewers were close to the wall but differently for each variable (*Angle*, *Area*, *Length*). Our second study examined the effect of perception when participants could move freely compared to when they had a static viewpoint. We found that a far but static viewpoint was as accurate but less time consuming than one that included free motion. Based on our findings, we recommend encouraging viewers to stand further back from the display when conducting perception estimation tasks. If tasks need to be conducted close to the wall display, important information should be placed directly in front of the viewer or above, and viewers should be provided with an estimation of the distortion effects predicted by our work—or encouraged to physically navigate the wall in specific ways to reduce judgement error.

Index Terms—Information Visualization, Perception, Wall Displays

1 Introduction

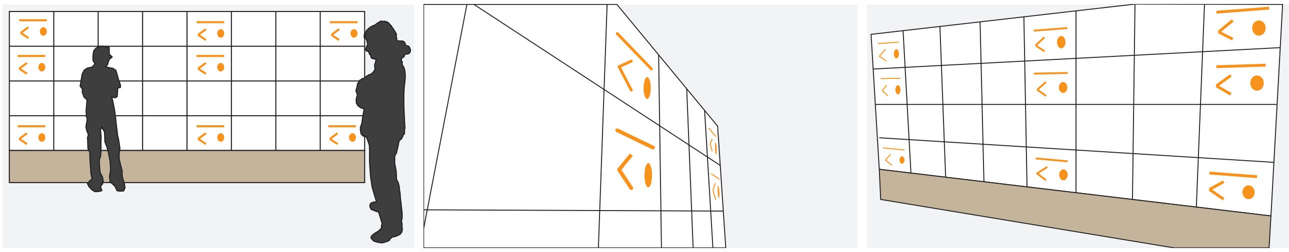
Mega- and Giga-pixel wall-sized displays (henceforth referred to as wall-sized displays) offer the opportunity to engulf viewers in very large high-resolution information spaces. They form intriguing new environments for data analysis and information visualization due to several inherent benefits: physical rather than virtual navigation affords a natural pan-and-zoom in the information space, an enlarged physical space in front of the display enables collaborative viewing and data analysis, and millions of pixels support viewing tremendous amounts of data in one shared environment [6, 16]. To fully leverage wall-sized displays for *data analysis*, however, we need to design wall-sized *visualizations* and workspaces based on a sound understanding of how human's perceptual and cognitive capabilities are affected by this new work environment. At the most basic level, visualization workspaces for wall displays have to incorporate what we already know about the design of information visualizations for desktop-sized displays. Beyond this knowledge, wall-specific design recommendations have to be developed. One important criterion for the development of information visualization techniques for wall-sized displays is their immense physical size. It is not uncommon to see wall displays of over 5m (16') × 2m (6.5') in width and height [7, 16]. Even complete rooms covered on all sides by high-resolution displays are being constructed for visualization research and applications [35].

With physically large display-walls, physical navigation becomes an important means of accessing an information visualization [6, 16, 41]. Viewers choose close or far viewpoints to zoom in and out, and pan physically by moving left and right to see different parts of the display. This type of movement may involve a physical relocation as well as a change of head orientation, as depicted in Fig. 1. Thus, viewers fluidly and frequently switch viewing distances and angles which may lead to systematic discrepancies between the actual appearance of displayed information in *physical* space (as can be measured by rulers) and its psychophysical appearance in a person's *visual space*.

Understanding discrepancies and where and when they occur is important for information visualization design, as fundamental data analysis tasks involve the correct assessment and comparison of elementary visual variables such as areas, angles, positions, slopes, or lengths [12]. To read a bubble chart, for example, one has to compare the sizes of circles to one another and to a legend, as well as relate positions in a

- *Anastasia Bezerianos is with Univ Paris-Sud & CNRS (LRI) and INRIA E-mail: anastasia.bezerianos@lri.fr.*
- *Petra Isenberg is with INRIA, E-mail: petra.isenberg@inria.fr.*

Manuscript received 31 March 2012; accepted 1 August 2012; posted online 14 October 2012; mailed on 5 October 2012.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.



(a) The viewpoints of two observers. (b) Wall as seen by the left observer. (c) Wall as seen by the right observer.

Fig. 2. Two observers looking at the same angles, lengths, and circles displayed across a large wall display.

2*D* coordinate space. Fig. 2 gives an example of how the appearance of three visual variables is affected when seen from different viewpoints and viewing angles. The question arises whether comparisons such as these are affected by the oblique viewing angles which occur when viewing data from different positions in front of a wall-sized display.

To-date many high-resolution wall-sized displays, including ours, are assembled from multiple LCD monitors [7, 16, 35]. These setups include clear visible bezels which form part of our study context. The research we report on in this paper, thus, takes a first step towards assessing the implications of changes in viewpoint on the assessment of data representations on *tiled* wall-sized displays with visible bezels.

Our research is motivated by three main questions:

- *Are all areas of a wall equally effective for close scrutiny and comparison of data items?*
- *What is the effect of viewing distance and angle on the perception of visual variables in large viewing spaces?*
- *What are the benefits of walking in comparison tasks?*

We began addressing these questions by studying how perception of elementary visual variables (*Angle*, *Area*, *Length*) was affected by varying viewing distances and angles. We contribute two studies: the first assessed static viewing conditions and identified different parameters that can help predict the perceived magnitude of the tested visual variables. The second contributes an understanding of the influence of allowing participants to move in front of the display. Our final contribution is a set of design implications about placement of data items on wall displays and the characteristics of effective physical navigation.

2 Related Work

We can draw from a variety of past research for the design of our experiments. A large chunk of the literature comes from the field of psychophysics. We report on the related background in this field separately in the following section as we lead into the study design. In this section, we concentrate on the related literature on large displays and perception of graphical elements in HCI and information visualization.

2.1 Viewpoints and Interaction with Large Displays

The problems of viewing and interacting with information on physically large displays has been investigated in HCI, focusing on several different questions: how to acquire targets across large distances [33], how to view far areas up-close [8], how to maintain awareness [10, 25], how large displays influence performance in spatial orientation tasks [36], or how a larger field of view influences user performance [15]. In contrast to these questions we want to learn how varying viewing distances and angles affect the *accurate perception* of a virtual object's properties such as its area, length, or angles. We know of no large-display literature that asks this question but the problem has already been recognized [3]. Several researchers have instead considered the influence of varying viewpoints on other large-display tasks:

Jota et al. [24] studied the impact of viewing angles on pointing performance on a 3m × 1m wall. Several studies in the tabletop literature assessed the relationship of view position and 2*D* object rotation on coordination, comprehension, and collaboration [27, 28]. Viewpoints have also been studied for viewing 3*D* objects on tabletops [21]. In multi-display environments, Nacenta et al. [30] showed that dynamically correcting perspective based on a viewer's viewpoint improved performance for tasks such as targeting, steering, aligning, pattern-matching, and reading. These studies relate to ours in that they corroborate the importance of view positions and angles to task performance.

2.2 Information Visualization and Large Displays

Several researchers have considered the influence of a viewer's position in front of a large display on information visualization tasks. For tabletops, Wigdor et al. [39] studied how varying screen orientation from a horizontal to up-right position influenced the accurate perception of elementary graphical elements. They found perception to be least accurate in the horizontal position. This study resembles ours in that elementary elements were tested using study techniques from psychophysics [20]. We relate some of their findings more closely to ours in our Discussion Section. Alallah et al. [2] tested how the perception of simple charts was impacted by varying viewing angles around a horizontal screen. They found that reading charts right-side up was fastest and least error-prone, and proposed a new chart design to alleviate orientation problems.

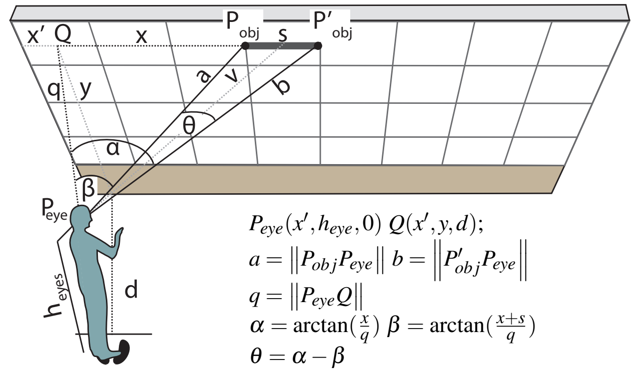
For wall-sized displays several studies explore how changes in a viewer's position affect how visualizations are read. Endert et al. [16] discuss how a viewer's distance from a large display influences the visual aggregation of displayed information. They found encodings based on a color ramp to visually aggregate particularly well across viewing distances for a visual search task. Yost and North [41] tested several data visualizations for their ability to effectively display large amounts of data on large displays. They found their visualizations to scale well for the tasks of finding detailed and overview information and note that spatial encoding of information was particularly important on large displays. In a follow-up experiment Yost et al. [40] studied how scaling visualizations beyond visual acuity affected user performance. For almost all tested tasks they found performance improvements and argue for design guidelines that take visual aggregation and physical navigation into account. Ball and North [5] compared the benefits of added peripheral vision vs. physical navigation for large displays, and found that physical navigation influenced task performance while added peripheral vision did not. The authors further stress the importance of physical navigation for visualization tasks. The stream of research on physical navigation relates to ours as a strong motivation for studying the influence of changing viewpoints and angles on accurate perception of data representations.

3 Background in Psychophysics

Psychophysics is a sub-discipline of psychology that is concerned with measuring the relationships between perceived and actual properties of a visual object [20, 37]. Much research in psychophysics is concerned with the study of spatial perception and the comparison of physical and visual space. Unfortunately no one model exists which clearly describes visual space and would allow us to predict how elementary graphical elements will be perceived in a variety of viewing conditions [37]. While it has been proposed to model visual space using hyperbolic, euclidean, or other geometries, no single geometry has been shown to work under all viewing conditions. Instead, researchers have attempted to mathematically describe the differences between physical and perceived magnitude of objects as collected from user studies. One popular function describing this difference is Stevens' [34] power law: $J = \lambda D^{\alpha}$, with J = judged magnitude, D = actual magnitude, α = exponent, λ = scaling constant. It has been tested under varying conditions, and several values for α have been proposed for judging elementary graphical elements (*visual variables*) such as length, area, or position. Wagner [37] gives a recent meta-analysis of 104 articles reporting 530 values for α collected under different conditions. No combination of conditions matched those of viewing elements on wall-sized displays. The reported exponents can, thus, help us hypothesize but not predict how reading elementary graphical variables may be affected in our work environment. As no previous study matches our viewing conditions, we conducted our own experiments under conditions close to how one would work in front of a wall-sized display. Our conditions involved: binocular vision, eye movement, changing head positions and viewing distances, and a back-lit viewing surface.

Psychophysics has developed several methods to help assess a viewer's visual perception of an object and to, thus, compare its magnitude (e. g., size) in the *physical space* to its subjectively experienced magnitude in a person's perceived *visual space*. Methods include numeric estimation, magnitude production, and sensitivity measures [37]. There is a debate as to which method is the best to measure the perceived magnitude of a given object. The methods of numeric estimation have been used in many experiments in the past (e. g., [14, 20, 23, 37, 39]). In our experiment we chose to use a magnitude production methodology. Here observers are asked to match two types of perceptions. Participants are shown a "standard" modulus object and are asked to change the intensity of a second object (the stimulus) until it is perceived to be equivalent to the modulus. We chose a magnitude production methodology for our experiment as the comparison judgements it requires are extremely frequent in information visualization [19]. We give additional justification in Section 4.

It is known that no exponent for Steven's law holds under all viewing conditions [38]. Given the large number of varying factors, none that matched our study setup in its entirety, we have to use average exponents for forming study hypotheses. Wagner [37] reports the following average exponents for studies on perception tasks: 1.02 for position and length, 0.84 for area, and 0.76 for angle. These state that generally people's judgement for position and length is consistent with actual positions and lengths, while angles and areas are underestimated compared to their real sizes. It has further been investigated how the visual angle—the angle a viewed object creates on the retina—and viewing distance influences the perception of visual variables [18, 29]. In order to derive hypotheses from articles suggesting an influence, we calculated viewer-object distances and visual angles for distinct regions on our wall size display as can be seen in Fig. 3.



$P_{eye}(x', h_{eye}, 0)\ Q(x', y, d);$
$a = \left\|P_{obj}P_{eye}\right\|\ b = \left\|P'_{obj}P_{eye}\right\|$
$q = \left\|P_{eye}Q\right\|$
$\alpha = \arctan(\frac{x}{q})\ \beta = \arctan(\frac{x+s}{q})$
$\theta = \alpha - \beta$

(a) Calculations for visual angle θ.

47° 78	29° 104	13° 158	6° 221	4° 286	3° 353	2° 420	1° 488
60° 60	30° 91	11° 150	5° 215	3° 282	2° 349	1° 417	1° 485
51° 73	30° 100	12° 156	6° 219	4° 285	2° 352	2° 419	1° 487
36° 106	26° 126	14° 174	8° 232	5° 295	3° 360	2° 426	2° 493

(b) θ, v for $d = 60cm$ and $h_{eye} = 160cm$ on our wall.

12° 324	12° 331	10° 352	9° 384	7° 425	6° 473	5° 525	4° 580
12° 320	12° 327	10° 348	9° 381	7° 422	6° 470	5° 522	4° 578
12° 323	12° 330	10° 351	9° 383	7° 424	6° 472	5° 524	4° 580
12° 332	11° 339	10° 359	9° 391	7° 431	6° 478	5° 530	4° 585

(c) θ, v for $d = 320cm$ and $h_{eye} = 160cm$ on our wall.

1-8	8-16	16-23	23-30	30-38	38-45	45-53	53-60

(d) Heatmap legend for visual angle in Fig. 3(b) and 3(c).

Fig. 3. Calculations of visual angles θ and viewer-object distance v (in cm) on our wall display. We tested screens with red borders in our study.

4 Study Motivation

Given previous work we expect that locations with smallest visual angles (resulting from object size, position, and viewing distance) will result in larger visual distortion of the perceived visual variables. To understand the effect of different display locations and viewing distances in detail we conducted two magnitude production experiments.

In *Experiment 1*, our goal was to determine how different object positions and sizes affect perception, by asking participants to interactively decrease the magnitude of an object's visual variable to match the magnitude of another object's visual variable at another area in the display. This is motivated by the following scenario: People position themselves in front of information of interest to facilitate their tasks [6, 16]. When assessing information of interest, the data elements often have to be placed within their larger context, to determine how they compare to others (e. g., compared to a legend). Although viewers could walk to get a closer look at data and walk back, this type of interaction comes at a cost of efficiency, especially when data needs to be quickly compared. Furthermore, collaborative settings may require viewers to quickly achieve common ground by comparing what someone else is viewing. For these data analysis scenarios, it is unclear how the perception of informations is affected by different static viewer placements around the wall. In *Experiment 1* we, thus, compare distortion across three visual variables (*Angle*, *Area*, *Length*) and try to determine if it is predictable. We attempt to characterize this effect and determine when quick comparisons from a stationary viewpoint, by turning one's head, are acceptable, and when the potential distortion errors are such that they require physical navigation or additional interface widgets to bring remote information closer. The goal of *Experiment 2* was to investigate free movement as an alternative to static viewer placement. In contrast to Experiment 1, participants were allowed to move freely in front of the wall display. We were interested in the movement choices and strategies participants followed when allowed to walk, as well as time vs. accuracy trade-offs.

5 Hypotheses

From an assessment of the psychophysics and information visualization literature we derived a number of hypotheses for our experiments:

- *H1:* Accuracy results for visual variables follow those of previous work with lowest absolute error for *Length*, followed by *Area*, and *Angle* (upright) [39].
- *H2:* The nature of judgement errors will differ between different visual variables. Based on our visual angle calculations (Fig. 3) distant objects look smaller and the only depth cues available to viewers are bezels. We thus expect areas to be underestimated on average. Angles oriented towards the biggest axis of distortion (Fig. 2) will be overestimated: their line segments look smaller and they will seem more obtuse. As in previous studies [37] lengths will correspond approximately to their actual sizes.
- *H3:* Accuracy decreases with growing distance between viewer and remote object. H3 contrasts H6 in Wigdor et al.'s study [39] that found no such effect, as we test much larger left-right distances.
- *H4:* Performance (accuracy and task time) decreases for close viewpoints as differences in visual angles are more extreme following H2 and the visual angle calculations in Fig. 3 that show smaller visual angles for remote objects.
- *H5:* The accuracy and nature of judgments of different visual variables is impacted differently for increasing object distances and viewing distances from the wall, but in a predictable way.
- *H6:* Accuracy increases when free movement is allowed, at a cost of temporal efficiency.

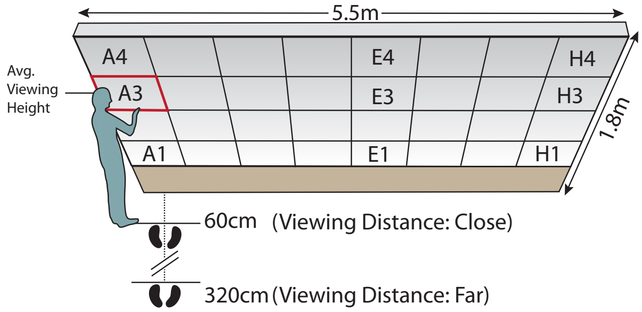


Fig. 4. The physical experiment setup showing the dimensions of our wall and modulus locations using chess notation (stimulus always in *A3*).

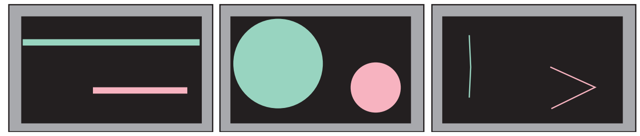


Fig. 5. Example screens showing the large interactive stimulus (green) the viewers adjusted to match the remote modulus (red) for *Length*, *Area* and *Angle*. The stimulus was always close to the participant's location.

6 Experiment 1: static comparisons

Participants were placed at two fixed positions left-most in front of the wall. We chose left-most positions instead of centered ones as we expect results to be symmetric left and right and because we could test the most extreme distances. At each position participants were requested to engage in magnitude production tasks and interactively adjust the magnitude of an object's visual variable close to their location, to match the magnitude of the same variable on a remote modulus object. As a control condition, the two objects were occasionally drawn on the same screen. The works of Cleveland and McGill [12] and Wigdor et al. [39] differ slightly to ours as they used a magnitude *estimation* methodology. We followed this approach in an original pilot of 16 participants, but found that they tended to round their results to the closest 10%. This produced very noisy data and as a consequence results that were not accurate enough when attempting to predict perspective distortion. Thus, we decided on a magnitude *production* experiment that bypasses the mental conversion of a size to a number.

6.1 Apparatus

We used a $5.5m(18') \times 1.8m(5.9')$, tiled wall-sized display consisting of 32 LCD screens of 2560×1600 resolution each. Screens are arranged as seen in Fig. 4 resulting in an effective resolution of 20480 x 6400 pixels, and are driven by a 18 workstation cluster. Software was written using the ZVTM toolkit with cluster extension [31]. Lights inside the experiment room were dimmed to reduce glare effects.

6.2 Factors

Our study included three main factors: *visual variable*, *viewing distance* from the wall, and modulus *location* and *size*.

6.2.1 Visual Variable

We used a subset of Cleveland's [12] elementary graphical perception tasks, namely assessing *Length*, *Angle* and *Area* as they are among the most highly ranked by Cleveland [12] and because we hypothesized them to be impacted by perspective changes. We did not test position, slope, and color for the following reasons. In our pilot study we tested position and found it to be largely unaffected by distortion. Furthermore, testing position is highly impacted by the presence of bezels, as positions can be easily compared within one single screen from a bezel onward. We thus decided not to include it in our final study to reduce time constraints on participants. Slope was not considered as previous work suggests a close relationship to angle judgments [14]. Finally, similar to Wigdor et al. [39], color was not investigated, as color consistency across the wall is hard to achieve under differing viewing angles, creating a likely confound. This is especially true in our setup, as color perception is heavily influenced by the viewing angles of particular LCD models [22, 26], and some viewing angles can even invert color perception.

The interactive object and the remote modulus were drawn with a distinct color of $\sim 81\%$ intensity (#7FFFD4, #FFB6C1). The arms of the angles were of different length for modulus and stimulus in order to avoid participants making vertical length judgements on the angle's open side. Participants were informed about this choice. To minimize possible influences due to the presence of bezels [4, 11], objects were drawn fully within a wall tile on a black background. For *Length*, objects were oriented horizontally. *Angle* judgements are known to be affected by angle orientation [37] so we chose to keep a consistent *Angle* orientation that follows the axis of biggest distortion (the angle bisector was horizontal). In results reported by Wigdor et al. [39] this "upright" angle orientation lead to larger errors than an orientation rotated by 90º. Fig. 5 shows examples of how the interactive object and modulus were drawn if sharing the same screen.

6.2.2 Viewing Distance

Participants performed tasks at two distances from the wall: *DistanceClose* $= 60$cm($\sim 24''$) and *DistanceFar*$=$ 320cm($\sim 126''$). *DistanceClose* is within the recommended range for desktop monitor viewing [17]. We chose it because it affords viewing objects in great detail at regular monitor distance, as well as direct-touch interaction. Given a conservative number of 60º for the human visual field outward from the nose for each eye, *DistanceFar* was chosen so that viewers had the entire wall in view when looking straight at it. Fig. 4 gives an overview of the two viewing distances.

6.2.3 Modulus locations and sizes

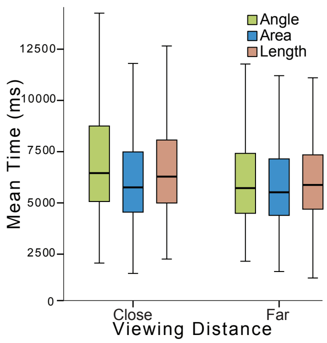
We used 9 modulus *locations*, described in chess notation (Fig. 4). From the left we used columns A, E, and H and rows 1, 3, and 4 from the bottom. Given the height of our wall and the average height of our participants, location $A3$ was always parallel to the viewer's frontal plane and had the shortest viewing distance in both *DistanceClose* and *DistanceFar* (Fig. 3). We refer to $A3$ as the *frontal screen*.

For each visual variable, participants were presented with 6 modulus *sizes* (intensities/magnitudes) to produce. These were 10%, 20%, 30%, 40%, 60% and 70% of the initial size of the interactive stimulus for each visual variable. These initial stimulus sizes were always 180^o for *Angle*, 2560 pixels for *Length* (a single screen width), and 1280 pixels for the diameter of *Area* (half the screen width) respectively. We ensured that these initial sizes allowed the modulus to be visible in the smallest increments, while still be able to fit on the same screen as the interactive stimulus for the $A3$ frontal screen *location* condition. During each trial, the interactive stimulus had to be interactively reduced in size until it perceptually matched the remote modulus.

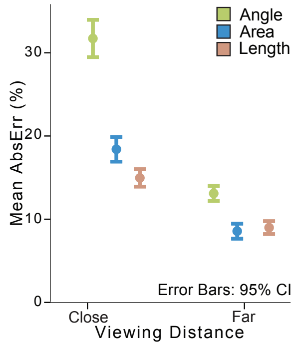
6.3 Participants and Procedure

Fifteen participants (5 female) took part in the study, recruited from our research institute. They were not paid for their involvement. Participants ranged from 24–33 years in age (mean & median age 29), 7 were students and 8 non-students with technical occupations. All participants reported normal or corrected-to-normal vision. Twelve participants reported experience with wall-sized displays in work tasks or games; the remaining 3 participants reported no previous experience.

Visual variable presentation order was randomized using a latin square. Presentation of modulus locations and sizes was also randomized, and their exact position within their screen location was varied between trials. Participants adjusted the size of the interactive object using the UP and DOWN arrow keys of a wireless keyboard on a stand in front of them. When the desired size was achieved they hit ENTER to terminate the trial. Before each trial started, the screens containing the stimulus and modulus were highlighted to ensure participants did

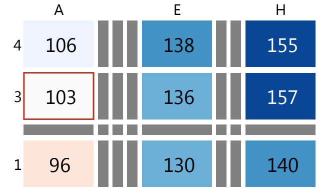


(a) Time.

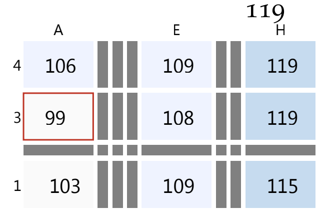


(b) Absolute Error.

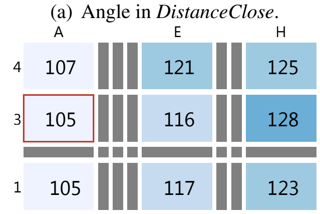
Fig. 6. Time and Absolute error across all visual variables for Close and Far viewing distances.



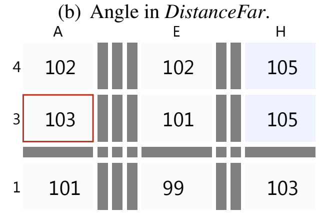
(a) Angle in *DistanceClose*.



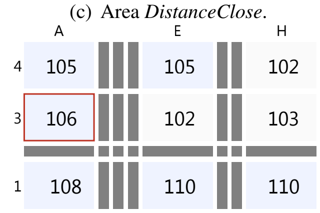
(b) Angle in *DistanceFar*.



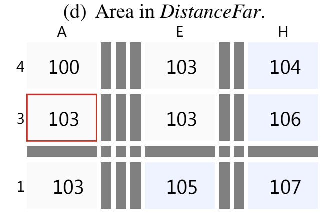
(c) Area *DistanceClose*.



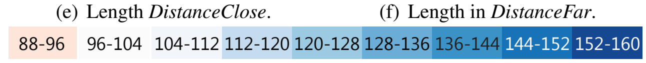
(d) Area in *DistanceFar*.



(e) Length *DistanceClose*.



(f) Length in *DistanceFar*.



(g) Legend for the heatmaps.

Fig. 7. Results for magnitude estimations (100+*EstErr*) for each *visual variable* and *viewing distance* in Experiment 1. Values per screen indicate the percentage difference in average judgments for this modulus location. Values $> 100\%$ are overestimations and values $< 100\%$ underestimations. The frontal screen is highlighted with a red border.

not spend time on visual searches. Timing started once the stimulus and modulus appeared on the screen and stopped when ENTER was hit. After the study, participants filled out a questionnaire eliciting demographic information and subjective data on their performance and preference. Overall Experiment 1 consisted of:

3	tasks (*Angle*, *Area*, *Length*)	×
9	modulus locations	×
2	viewing distances (*DistanceClose*, *DistanceFar*)	×
6	modulus magnitude *sizes*	=
324	trials per participant	×
15	participants	=
4860	**trials in total**	

7 Experiment 1 Results

Metrics used in our analysis were *Time*, *AbsErr* and *EstErr*. We define *AbsErr* similarly to magnitude estimation studies [13, 39]. *AbsErr* is the absolute percentage of estimation error over the real magnitude of the modulus object. Thus if participants report stimulus magnitude m_u for a modulus of true magnitude m_t, then $AbsErr = \mid \frac{m_u - m_t}{m_t} * 100 \mid$. This metric expresses the overall error in judgement (irrespective of over- or under-estimation tendencies). It is a skewed distribution, and as suggested by Cleveland [13], we conducted our analysis on its log variation $log_2(\frac{1}{8} + AbsErr)$. Means reported are before normalization.

EstErr represents the direction of estimation error, i. e. the tendency to over- or under-estimate the magnitude of the modulus and by how much. It is defined as $EstErr = \frac{m_u - m_t}{m_t} * 100$, with $EstErr > 0$ when magnitude is overestimated, and $EstErr < 0$ when underestimated.

Trials were marked as outliers when metrics were beyond two standard deviations from the mean for a given *visual variable*, *viewing distance*, *size* and *location*. 186 trials (3% of all trials) were identified as outliers and removed from further analysis. Similar to Cleveland and McGill [13] the remaining trials were aggregated per *participant* and factors for all *size*s. All metrics followed the normal distribution. All analyses were performed using an ANOVA, and post-hoc pair-wise mean comparison p-values are adjusted using the Bonferroni criterion.

7.1 Results Across Visual Variables

We first analyzed effects across visual variables *Area*, *Angle* and *Length* and compared their performance.

Time (Fig. 6.a)

There was no significant effect of *visual variable* or *location* on time. Mean *Time* was longer for *Angle* (7.12 sec), followed by *Length* (6.6 sec) and *Area* (6.39 sec). There was a significant effect of *viewing distance* ($F_{1,14} = 17.3$, $p < .001$). Tasks performed at *DistanceClose* took significantly longer than those at *DistanceFar* (all $p < .05$).

AbsErr (Fig. 6.b) and *EstErr* (Fig. 7)

There was a significant effect of *visual variable* on *AbsErr* ($F_{2,28} = 95.2$, $p < .0001$). Pair-wise comparisons showed that errors in judgement were significantly larger for *Angle* than all others ($p < .001$) with no other differences. Mean values for *Angle* (22%) were larger, followed by *Area* (13%) and *Length* (11%).

Our ordering of visual variables according to accuracy is different than that reported by Cleveland [13] (where angles have smaller errors than areas), but similar to Wigdor [39] for upright angles. We examined this order separately when both objects were placed in frontal screen *A*3, to investigate if the effect was only present in remote location, but found this ordering to be present even on the frontal screen.

EstErr gives us the tendencies (direction) of estimation error. There was a significant effect for *visual variable* ($F_{2,28} = 25$, $p < .0001$). *EstErr* was different for all *visual variable*s (all $p < .05$), with the modulus being consistently overestimated, but by different amounts. Mean overestimation was significantly larger for *Angle* (19%), followed by *Area* (9%) and *Length* (4%).

The somewhat stronger differences of *EstErr* than *AbsErr* indicate that although the different *visual variable*s were affected somewhat differently in terms of absolute magnitude, it is the tendencies to over- and under-estimate that are different, with clear tendencies to overestimate in *Angle* but less consistent tendencies for *Length* and *Area*.

Viewing Distance: There was a significant effect of *viewing distance* on *AbsErr* ($F_{1,14} = 199.5$, $p < .0001$), with less *AbsErr* in the *DistanceFar* condition ($p < .001$). There was no significant *visual variable* × *viewing distance* interaction, indicating accuracy did not vary differently for the different *visual variable*s at different distances.

There was a significant effect of *viewing distance* on *EstErr* ($F_{1,14} = 73.5$, $p < .0001$). Participants overestimated overall, with larger overestimations in *DistanceClose* than *DistanceFar* ($p < .001$). A significant *viewing distance* × *visual variable* interaction ($F_{2,28} = 34.5$, $p < .0001$) indicates that the direction of error was affected differently by *viewing distance* for each of the *visual variable*s. Pair-wise comparisons showed all *visual variable*s to be different for *DistanceClose* (all $p < .05$) following the global trends described before. In *DistanceFar* there was no difference between *Length* and *Area*, indicating *Angle* was overestimated significantly more even in the *DistanceFar*.

LOCATION: There was a significant effect of screen *location* on *AbsErr* ($F_{8,112} = 54.8$, $p < .0001$). *AbsErr* increased with column distance (all reported effects $p < .01$): $A1, A3, A4$ had significant less *AbsErr* than all others, with no difference between screens in that column. Similarly *AbsErr* in the medium column $E1, E3, E4$ was significantly higher than the screens in A, and lower than remote screens in H. Finally the remote screens $H1, H3, H4$ had the largest *AbsErr*. There is, thus, a clear screen grouping across columns in terms of *AbsErr*.

The effects of direction of estimation are similar, with significant effect of screen *location* on *EstErr* ($F_{8,112} = 26.3$, $p < .0001$). Overall participants overestimated, and overestimation increased with column distance $A < E < H$ (Fig. 7). Overestimation on the column $A1, A3, A4$ was significantly less than all others, screens in the middle $E1, E3, E4$ had significantly larger *EstErr* than the first column, and significantly smaller *EstErr* than the two upper screens in the last column (all $p < .05$). We also observed a tendency ($p < .1$) for row 1 (lower screens) to have a lower average *EstErr* than the other screens in the same column.

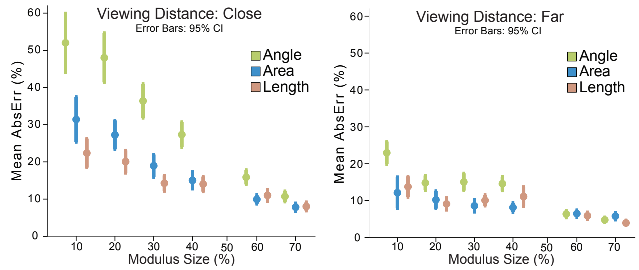
*Visual variables*s were affected differently by *location*. There was a significant *location* × *visual variable* interaction on *AbsErr* ($F_{16,224} = 1.9$, $p < .01$). Pair-wise comparisons (all $p < .05$) showed no difference between techniques in column A. Nevertheless the overall error of *Angle* increased compared to the others in the middle E and far column H. In E *Angle* has significantly larger *AbsErr* than *Length*, and in H larger than *Area* as well. No significant difference between *Area* and *Length* was found, nor significant effects depending on screen height.

The direction of error had clearer effects. There was a significant *location* × *visual variable* interaction on *EstErr* ($F_{16,224} = 30.2$, $p < .0001$). Pair-wise comparisons (all $p < .05$) showed that *Angle* was overestimated more compared to other *visual variable*s in most *location*s, but that this is not the case in screen $A1$ (lower screen close to the participant). In this location the estimation of *Angle* was significantly less than in all other screens for **all** *visual variable*s (the inverse trend from all other *location*s). Moreover, *Length* which tends to have a small overestimation, had one of the largest overestimations in screen $A1$. In the middle column E, we found no difference between *visual variable* at the lower screen $E1$, although for the 2 higher screens *Angle* was significantly overestimated. In the far column H, all techniques were different at $H3, H4$. But again for the lower screen $H1$ effects were less pronounced, with only *Angle* being different from the others. The effects stem mainly from the *DistanceClose* condition (all $p < .05$), but similar trends appear in *DistanceFar* ($p < .1$).

In summary, the effects of the screen height seen in *EstErr* were not as strong in *AbsErr*, indicating that it was the tendencies to over- and under-estimate that changed with screen height, not the absolute error.

SIZE: As in previous work [13, 39] we aggregated the results of different modulus *size*s for the main analysis above. In a separate analysis, we also tested for effects of *size* (a separate factor of 6 possible values). We found a significant effect of *size* on *AbsErr* ($F_{5,70} = 12.1$, $p < .0001$). Overall *AbsErr* decreased with the increase of object size, although only the two larger targets had significantly less *AbsErr* than other target *size*s ($p < .05$).

A significant *size* × *visual variable* interaction ($F_{10,140} = 39.2$, $p < .0001$) was also present. When looking at the effect of *size* on the different *visual variable*s we found that for the four smaller *size*s *Angle* had significantly larger *AbsErr* than all others, with no difference between visual variables for the 2 large *size*s. Fig. 8 shows that (especially in *DistanceClose*) the *AbsErr* drops for larger sizes, with a difference in the amount of decrease between *visual variable*s until there is little difference on larger object sizes.



(a) Size for DistanceClose. (b) Size for DistanceFar.

Fig. 8. Absolute error across all visual variables for the 6 modulus (target) sizes tested, for Close and Far viewing distances.

7.1.1 Summary

Our analysis showed no significant difference between *visual variables* for task *Time*, but a difference for the two error metrics.

The absolute error follows the ordering reported by Wigdor et al. [39], with *Angle* being the most error-prone and *Length* the least. *AbsErr* tends to increase when viewers are close to the screen, and when the distance to the remote object increases, with *Angle* being most affected. This last finding merits further discussion. An effect of stimulus-modulus distance was also reported by Cleveland [13], but Wigdor et al. [39] suggested that it may have been due to a possible confound in the original study. We discuss these findings in Section 9. We also found that the absolute estimation error decreased with the increase of object size. The rate of decrease was more steep for *Angle* (and somewhat less for *Area*), until errors were similar across visual variables for the largest object sizes.

The nature of over- or under- estimation was different per *visual variable*: *Angle* was consistently overestimated, except on screen $A1$, whereas *Length* and *Area* were less consistent in their tendencies (especially *Length*). The generally observed nature of overestimation was less pronounced in lower screens. Nevertheless, as we move upwards on the wall the overestimation becomes more pronounced for *Angle* followed by *Area*. Looking at the estimation averages for *Angle* and to a lesser degree *Area* (Fig. 7) the amount of overestimation is lowest in the lowest screens of the same column, whereas *Length* tended to be overestimated by a larger degree at lower screens, thus balancing *EstErr* across *visual variable*s in these locations. This indicates that lower screens are perceived differently. Looking at horizontal screen location, *Angle* was affected the most, and *Length* the least, with estimates going up faster with horizontal distance.

7.2 Predicting Visual Variables

In the previous section we compared the *visual variable*s. We now examine each *visual variable* in an attempt to predict their observed behavior for our study setup. More specifically we examine the effect of the different factors related to *perceived magnitude PerMag* (that is the participant's answer m_u) given the true magnitude m_t. In our previous findings, effects were similar across rows or columns of the wall, thus we express screen location as a combination of *column* A, E, H (horizontal displacement), and *row* $1, 3, 4$ (vertical displacement).

7.2.1 Results for Visual Variable: Angle

We found a significant effect of *viewing distance* ($F_{1,14} = 89.7$, $p < .0001$), *column* ($F_{2,28} = 96.8$, $p < .0001$) and *row* ($F_{2,28} = 65.5$, $p < .0001$) on *PerMag*, as well as a significant *viewing distance* × *column* ($F_{2,28} = 81.5$, $p < .0001$) and *viewing distance* × *row* ($F_{2,28} = 19.8$, $p < .0001$) interaction. Pair-wise comparisons (all $p < .05$) showed that overestimation of *Angle* was significantly different between the three different *column*s, increasing with column distance. This effect was present both in *DistanceClose* and *DistanceFar*, although less pronounced in *DistanceFar* (all $p < .05$). For *row*s, the lower screens (row 1) always had significantly less overestimation, with no other differences. This effect was only present in *DistanceClose*, with no difference due to screen *row* in *DistanceFar*. There was also an effect of *size* on *PerMag* ($F_{5,70} = 1895.3$, $p < .0001$), with all *size*s being perceived differently (all $p < .0001$).

ANGLE PREDICTION Based on the above, we expect that the perceived *Angle* increases with the increase of the factors *size*, *column* and *row*, and decreases with the factor *viewing distance*. Indeed, we found a positive Pearson correlation between the dependent variable *PerMag*, the true object *size* ($r = 0.943$, $n = 1620$, $p < .0001$) and screen *column* ($r = 0.161$, $n = 1620$, $p < .0001$), and a negative correlation with *viewing distance* ($r = -0.12$, $n = 1620$, $p < .0001$). We found no significant correlation with *row*, but there was a clear trend ($p = .07$). No correlations were found between the predictor variables, indicating they are mutually independent. Thus we feel these factors are enough to predict the perceived angles. To verify this hypothesis for our setup, we ran a multiple linear regression analysis using the above factors. We obtained a very good fit for predicting the reported angles ($R^2 = .93$, Adjusted $R^2 = .93$). In our regression analysis we expressed *viewing distance*, *column* and *row* in cm, and the predicted and actual angles in angle degrees. Table 1 summarizes the coefficients that predict *Angle* in our setup.

7.2.2 Results for Visual Variable: Area

A significant effect of *viewing distance* ($F_{1,14} = 60.3$, $p < .0001$), *column* ($F_{2,28} = 9.2$, $p < .01$) and *row* ($F_{2,28} = 5.4$, $p < .01$) on *PerMag* was present, and a significant *viewing distance* × *column* ($F_{2,28} = 28.4$, $p < .0001$) interaction. Overestimation of *Area* was significantly different between the three different *column*s, increasing with column distance. The effect was due to the *DistanceClose* condition (all $p < .05$). For screen *row*, the lower screens (row 1) had significantly less overestimation than the higher ones, with no other differences (all $p < .05$). There was also an effect of *size* on *PerMag* ($F_{5,70} = 3847.4$, $p < .0001$), with all *size*s being perceived differently (all $p < .0001$).

AREA PREDICTION We expected that perceived *Area* will increase with increasing factors *size*, *column* and *row* and decrease when increasing *viewing distance*. Indeed, we found a positive Pearson correlation between the dependent variable *PerMag*, the actual *size* ($r = 0.969$, $n = 1620$, $p < .0001$) and *column* ($r = 0.05$, $n = 1620$, $p < .05$), and a negative correlation with *viewing distance* ($r = -0.096$, $n = 1620$, $p < .0001$). We found no significant correlation with *row* and no correlations between the predictor variables. Thus, these factors (excluding *row*) are enough to predict the perception of areas. To verify this hypothesis, we ran a multiple linear regression analysis using the above factors. We obtained a very good fit ($R^2 = .925$, Adjusted $R^2 = .925$), although *column* had a very small influence. In our analysis we expressed *viewing distance* and *column* in *cm*, and the predicted and actual areas in cm^2. Table 1 summarizes the coefficients that predict *Area* in our setup.

7.2.3 Results for Visual Variable: Length

We found a significant effect of *viewing distance* ($F_{1,14} = 8.3$, $p < .05$) and of *row* ($F_{2,28} = 6.7$, $p < .01$) on *PerMag*. Results show that participants overestimated to a larger extend in *DistanceClose*. Moreover, objects in the lower screens (Row 1) were significantly overestimated compared to the other two rows (all $p < .05$). There was also a significant effect of *size* on *PerMag* ($F_{5,70} = 3953.6$, $p < .0001$) with all *size*s being significantly different.

LENGTH PREDICTION Given these results, we expect that the perceived *Length* increases with increasing *size*, and decreases with the increase of factors *row* and *viewing distance*. Nevertheless, a correlation analysis (over all factors), only shows a significant positive Pearson correlation between the dependent variable *PerMag* and the actual *size* ($r = 0.971$, $n = 1620$, $p < .0001$). Thus factor *size* should be enough to predict perceived lengths. To verify this hypothesis for our setup, we ran a linear regression analysis using *size* only as a factor. We obtained a very good fit ($R^2 = .939$, Adjusted $R^2 = .939$) for predicting the reported lengths. For our analysis we expressed the predicted and actual *size* in cm. We summarize the coefficients of the linear relationship that predicts lengths in our setup in Table 1.

	Perceived Size Magnitude Coefficients		
	Angle (*degrees*)	Area (cm^2)	Length (*cm*)
Constant	4.286* (0.931)	0.022* (0.003)	-3.124* (0.167)
True Magnitude m_t	0.977* (0.007)	1.027* (0.007)	0.944* (0.006)
Viewing Distance (*cm*)	-0.35* (0.002)	-0.11* (0.001)	
Screen X (*cm*)	0.32* (0.001)	$3.768 \cdot 10^-5*$ (0.000)	
Screen Y (*cm*)	0.3* (0.004)		
R-square	0.932	0.925	0.939
Adjuster R-square	0.932	0.925	0.939
Number of observations	1620	1620	1620

Standard errors are reported in parentheses.
* indicates significance at the 99% level.

Table 1. Regression analysis coefficients C. For our setup the perceived size can be predicted using the following equation $PerMag = Constant + C_{m_t} * m_t + C_{Distance} * Distance + C_{screenX} * ScreenX + C_{screenY} * ScreenY$.

7.2.4 Discussion on Prediction

We note that in our setup a linear relationship between *size* and the other factors is enough to provide a very accurate model of the perceived magnitudes. Even though perception of magnitude of visual variables follows a power low relationship with their true magnitude [37], an initial curve fitting (per visual variables, and *viewing distance*) showed an almost linear relationship (α very close to 1). We believe this is due to the fairly small amount of *size*s tested (6) compared to other perception studies. We expect that with an increase of sizes tested we will be able to observe such a power law behavior and further improve our model.

Although not reported, we tested visual angle and viewer-object distance (Fig. 3) as predictors of perceived magnitude. An inverse correlation was present (smaller visual angles lead to larger overestimation, larger viewer-object distances to smaller overestimation), but their influence is different at the two user distances. For example, Column *E* and *H* have similar visual angles at *DistanceClose* and *DistanceFar*, and Column *H* and *E* similar viewer-object distances (Fig. 3) but magnitude estimations were quite different (Fig. 7). Thus we feel the reported models are better predictors.

7.3 Questionnaire

We were further interested in the influence of the bezels. As we could not measure their influence directly, we asked participants for their strategies in solving the tasks and if they involved bezels. Thirteen participants reported to have used bezels, most of them for the *Length* task, but some noticed that bezels were only useful as landmarks for the larger sizes. It would be interesting to study the influence of bezels further in a dedicated experiment with the use of an eye-tracker.

8 EXPERIMENT 2: STATIC VS. MOVING

In Experiment 1 our goal was to understand and predict the effect of visual distortion while viewers stand at close and far locations in front of the wall display. We motivated this choice by scenarios in which viewers stand in specific locations conducting detailed tasks, and want to occasionally make quick visual comparisons with objects at distant locations (such as a comparison to a legend placed elsewhere). Nevertheless, we acknowledge that if the main task of the viewer is the comparison itself, they may decide to move in front of the wall to gain a better perspective of the information to compare. We, thus, conducted a follow-up study where participants were able to move freely, tracked using a Vicon motion capturing system (`www.vicon.com`).

Nine participants of the original study (3 female) took part in the second study a week later. Given that the stronger effects observed in our first study were in the farthest *column*, we only tested these *location*s (and the frontal screen *A*3) - 4 *location*s overall. Participants started each trial close to the screen (as in *DistanceClose*) and were then able to move freely to perform their task. We analyzed these results with the results for *DistanceClose* and *DistanceFar* of our first study for the specific modulus locations.

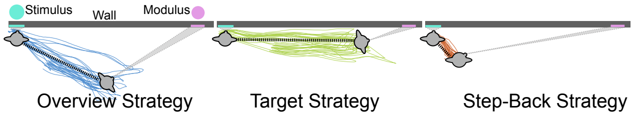
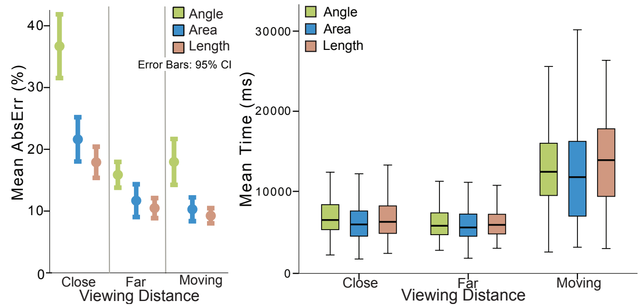


Fig. 9. Three participants' actual motion paths showing the three different types of moving strategies. We also illustrate possible modulus and stimulus locations and participants' viewing angles.



(a) AbsErr (b) Time

Fig. 10. Absolute Errors and Times for visual variables in Experiment 2.

Overall Experiment 2 had:

3	tasks (*Angle*, *Area*, *Length*)	×
4	modulus *locations* (including *A*3)	×
3	viewing distances (*DistanceClose*, *DistanceFar*, *ViewerMove*)	×
6	modulus magnitude *sizes*	=
216	trials per participant	×
9	participants	=
1944	**trials in total**	

8.1 Results

8.1.1 Moving Strategy

Three moving strategies emerged during our experiment. Four participants adopted an *overview* strategy, walking to the center of the display at a far distance ($\sim$ 3m), to observe both stimulus and modulus under a comparable visual angle in each direction. Three participants adopted a move to *target* strategy, walking until they arrived almost in front of the remote modulus. Finally, two participants performed a *step-back* strategy, moving slightly backwards from their original position ($\sim$ 1m) to look at the remote modulus. Sample strategy profiles can be seen in Fig. 9. Participants tended to be consistent in their strategies throughout the experiment. We observed changes only in the *target* strategy, where towards the end of the experiment participants tended to stop partway ($\sim$ 1m) before completely reaching the target. All participants performed tasks by first making an approximate judgement and then used walking to verify or adjust their initial judgement. All participants moved only once per trial.

AbsErr means were larger with the *step-back* strategy (20.4%), followed by the *target* strategy (11.1%), and with *overview* being the most accurate (9.5%). A Kruscal Wallis non-parametric test showed a significant effect of strategy on *AbsErr* ($Chi-square(2) = 13.1$, $p < .01$). Pair-wise comparisons showed that *step-back* was significantly more error phone than the others (all $p < .001$).

We also looked for learning effects between trials for each strategy, to see if participants' accuracy increased over time. Although we found no significant learning effect, when asked, five out of nine participants mentioned that after the end of the walking experiment they felt they could more accurately make estimations (even without walking). This leads us to believe that viewers can learn to self correct for visual distortion, a topic we plan to explore further in the future.

8.1.2 Static vs. Moving (Fig. 10)

ABSERR: There was a significant effect of *viewing distance* ($F_{2,16} = 18.2$, $p < .0001$) and *visual variable* ($F_{2,16} = 9.2$, $p < .0001$) on *AbsErr*, as well as a *visual variable* × *viewing distance* interaction ($F_{4,32} = 3.3$, $p < .05$). Pair-wise comparisons (all $p < .05$) showed that *AbsErr* was significantly higher in *DistanceClose* (25%), with no difference between *DistanceFar* (12.6%) and *ViewerMove*(12.4%). Again, *AbsErr* was significantly higher for *Angle* (23%), followed by *Area* (14.5%), and *Length* (12.5%). This difference between *visual variable*s was due to *DistanceClose* mainly, with no difference between *visual variable*s in *DistanceFar* and *ViewerMove* (all $p < .05$).

TIME: There was a significant effect of *viewing distance* ($F_{2,16} = 10.3$, $p < .0001$) on *Time*. Pair-wise comparisons (all $p < .05$) showed mean times to be significantly different between all *viewing distance*s for the tested *location*s. *DistanceFar* was faster (6 sec), followed by *DistanceClose* (6.7 sec), and was almost double for *ViewerMove* (13.1 sec). Thus, the accuracy benefits for *ViewerMove* come with a time cost, while *DistanceFar* is both faster and has similar accuracy.

9 Discussion

Our studies showed several interesting results in regards to our initial hypotheses. In **H1** we had hypothesized that results would follow previous work [39] and rank visual variables with increasing absolute error for length, area and angle (upright). Since we chose an angle orientation that is proven to be very error prone [39], our findings also follow this order, with angle being the most error prone visual variable (in all screens, including the frontal screen *A*3).

Based on previous work of Wagner who had conjectured that visual space was compressed in the in-depth dimension leading to angle overestimation [37], we had hypothesized (**H2**) that angles would be overestimated. This was indeed the case. We had also hypothesized that areas would be underestimated. This was contrary to our findings, with areas being overestimated. A possible explanation comes from related work. Aks and Enns [1] found that the addition of a grid to a scene of objects placed in 3D lead viewers to make object size corrections, hinting at a possibility that bezels may be used similarly. One of our participants accordingly stated "I compensated for my perspective." It is possible that our participants in an attempt to self-correct for perspective distortion did in fact self-correct too much. This effect was not as pronounced in length estimations and perhaps participants used bezels more successfully to estimate lengths than areas. Given participants' comments it is likely that results on length estimation may differ for a similar study on a seamless wall without bezels. Nevertheless results on angles and areas will most likely hold.

In **H3** we hypothesized that the effects of **H2** would increase with distance between stimulus and modulus. This was the case when participants were close to the wall, while the effect was less visible when they were far. The effect was present for both *Angle* and *Area*, and to a lesser extend for *Length*, confirming **H3**. In the work of Cleveland et al. [13] such an effect was observed as well, but not by Wigdor et al. [39], where it was shown that left-right distances did not lead to increasing error. We showed that the effect exists, and it was most likely not observed in [39] because they tested much smaller left-right distances and, thus, differences in visual angles, than our Experiment 1.

We partially confirmed **H4**. Participants were slower when standing close to the wall, but not significantly so. For accuracy, we confirmed an increase in absolute error when standing close to the wall for all *visual variable*s. In the questionnaire, all participants also reported that the tasks were easier to accomplish when standing far away.

In **H5** we hypothesized that visual variables would be independently impacted by changes in viewing angle and distance but were less sure about the nature of the impact. Indeed we found very different behavior across *visual variable*s. In general *Angle* was most impacted with highest inaccuracy. Although *Angle* judgments had a consistent overestimation tendency, estimations were smaller in the lower screens compared to other screens in the same column. *Area* had similar effects of increasing overestimation with left-right distance and a tendency for smaller estimations at lower screens, although the effects were mainly pronounced in *DistanceClose*. The judgements for *Length* were also affected by screen height, but in the inverse way. *Length* estimations that were on average closer to the real object values were overestimated in lower screens (an effect most visible in *DistanceClose*). Previous work in the physiology literature found a difference for visual activities in the upper- and lower visual fields [32], pointing to an interesting venue for further investigation for wall-sized displays.

Finally, in Experiment 2 we tested **H6** related to walking, focusing on extreme distance and distortion conditions (last column on the wall). As expected, accuracy for estimations from a static position close to the wall was worse than when participants were allowed to move. However, we found no difference in perception accuracy between moving and standing on a static position far from the wall. Moreover, the task completion time was more than twice as long in the moving condition and participants complained about fatigue.

10 Implications for Design

The results of our experiments apply to explicit comparison tasks that involve a quantitative comparison component [19] such as finding similarities, differences and trends, spotting outliers, or acquiring a quick overview. One of the goals of our experiments was to derive design considerations for visualizations for wall-sized displays that require these tasks. Our main questions in regards to workspace design were:

Are all areas of a wall equally effective for close scrutiny and comparison of data items?
Our analysis showed that indeed it is not recommended to compare data in certain screen locations as the error introduced reached as far as 157%/128%/110% overestimation for the three visual variables. We also found that lower screens tend to be somewhat unpredictable in their perception trends. We suggest that task-relevant data representations should not be placed at the lower positions on the wall. This is of importance to visualization designers, given that traditionally visual legends are placed at the bottom of visualizations and these legends often require visual comparisons (e. g., in a bubble chart the quantities represented by bubble sizes). Lower screens should be dedicated to widgets or contextual data that does not need to be reliably compared, such as visualization titles or numeric information about the data.

Should we redesign visualizations for walls for better comparison?
When viewers were close to the wall, we found that judgement accuracy for *Angle*, and to a lesser degree *Area*, started to drop for targets placed as far way as half the wall width ($\sim 3m$). Length was least affected by screen position and distances. When magnitude comparison tasks are expected to be performed regularly close the wall (e. g., comparing pie or sector charts) we recommend not to design visualizations such that they require comparison across large distances (more than $3m$), especially for *Angle*. Given the fairly predictable behavior of *visual variable*s we were able to identify factors affecting them and to provide approximation models for their perceived sizes that fit our observed data very accurately. These models can be used by visualization designers to predict visual variable distortion and decide on acceptable distortion effects in their visualizations.

In our experiments we did not test every possible visual variable (for time reasons). The use of color intensity was previously recommended as being particularly stable across viewing distances for a visual *search* task [16]. Its effect for comparison tasks, however, will have to be further investigated. Yet, given the high influence of LCD screens' viewing angles on color perception [22, 26] results may be difficult to generalize for other large wall setups.

How can we support data comparison at close viewing distances without visualization redesign?
We generally found comparisons across long distances when standing close to the wall most error-prone. If physical movement in front of the wall is not possible (e. g., while interacting using direct-touch or multiple viewers are occupying the area in front of the wall) specific widgets could be designed to bring far information up-close [9], enabling accurate comparisons with remote locations. Moreover, designers should provide additional aids to help viewers make judgements (e. g. use of tick marks, or value labels inside the visualization), that can act as guides much as the bezels did in our study. Alternatively, designers can use our prediction models to infer perceptual differences, and add additional meta-data on their visualizations about these calculations. At the very least viewers should be warned about distortion effects if designers deem comparison tasks important in their visualizations. For example a simple small text field could be added with a warning that "remote angles may appear twice as large" (similar to car mirror warnings for remote objects). Due to visual acuity these text fields could be made small to be only visible when needed, i. e. when viewers are close to the wall.

Should we encourage walking for comparison tasks?
Using interactive widgets in comparison tasks comes with an interaction cost. An alternative is physical navigation, which is flexible and natural to viewers, but in turn comes with a time cost, as our findings indicate. More surprising was that the mean accuracy was not better when participants could walk compared to a static viewpoint far from the wall. Participants' walking strategies may offer an explanation: as walking is tiring, some participants walked minimally and were thus still affected by visual distortion, resulting in higher error rates. Thus, recommendations for walking need to be more specific. Viewers need to either move far from the wall ($\sim$ 3m back), place themselves in the middle of the two objects to compare, or approach both objects to compare. Our models can be used to give viewers an approximate understanding of the distortion magnitude across the wall to help them decide when to make quick judgements turning their head, when to use interaction mediators to bring remote content closer to their focus of attention, or physically navigate. However, our discussion of walking guidelines is specific to quantitative comparison tasks. Physical navigation has been shown to be beneficial to other tasks such as zooming-in and -out to visually aggregate information [40]). The tradeoffs with these benefits need to be further investigated.

11 Conclusions

We conducted two studies to understand distortion effects for information visualizations placed on large high-resolution wall-sized displays. In the first, we tested two static locations in front of the display and found that viewing distance from the wall, as well as horizontal and vertical placement, affected errors. Participants performed tasks better when the information was in full view, despite the fact that they stood further away from the display and the objects to compare were visually smaller. We tested three visual variables and found that length was relatively unaffected by changes in viewing distance and placement on the wall—but area and angle judgments were significantly affected. Moreover, performance on the lower locations of the display was found not to be consistent with other locations. Finally we proposed prediction indicators of how large variations in viewing distances and object placement affect the accurate perception of these visual variables.

In the second study, we examined the trade-offs involved when allowing viewers to walk. We found that—although moving was as accurate as static comparisons from afar—it took twice as long and viewers occasionally chose non-optimal moving strategies. Based on these findings we derived design considerations which recommend to encourage viewers to stand further back from the display when conducting quantitative comparison tasks. As such, we support previous recommendations for different data analysis tasks for wall-sized displays [3] that promoted physical navigation.

If tasks need to be conducted close to the wall display, however, viewers should either be encouraged to physically navigate the wall in *specific ways* to reduce judgement error, or important information should be placed directly in front of the viewer or above, and viewers should be provided with an estimation of the distortion effects predicted by our work.

12 Acknowledgements

We'd like to thank Jean-Daniel Fekete for his comments on an earlier draft of the paper, our reviewers and our study participants.

REFERENCES

[1] D. J. Aks and J. T. Enns. Visual Search for Size is Influenced by a Background Texture Gradient. *Journal of Experimental Psychology: Human Perception and Performance*, 22(6):1467–1481, Dec. 1996. doi> 10.1037/0096-1523.22.6.1467

[2] F. Alallah, D. Jin, and P. Irani. Oa-Gaphs: Orientation Agnostic Graphs for Improving the Legibility of Charts on Horizontal Displays. In *Proc. ITS*, pp. 211–220. ACM, New York, 2010. doi> 10.1145/1936652.1936692

[3] C. Andrews, A. Endert, B. Yost, and C. North. Information Visualization on Large, High-resolution Displays: Issues, Challenges, and Opportunities. *Information Visualization*, 10(4):341–355, Oct. 2011. doi> 10.1177/1473871611415997

[4] R. Ball and C. North. Analysis of User Behavior on High-Resolution Tiled Displays. In *Proc. INTERACT*, volume 3585 of *LNCS*, pp. 350–363. Springer Verlag, Berlin / Heidelberg, 2005. doi> 10.1007/11555261_30

[5] R. Ball and C. North. The Effects of Peripheral Vision and Physical Navigation on Large Scale Visualization. In *Proc. Graphics Interface (GI)*, pp. 9–16. CIPS, Toronto, 2008.

[6] R. Ball, C. North, and D. A. Bowman. Move to Improve: Promoting Physical Navigation to Increase User Performance with Large Displays. In *Proc. CHI*, pp. 191–200. ACM, New York, 2007. doi> 10.1145/1240624.1240656

[7] M. Beaudouin-Lafon. Lessons Learned from the WILD Room, a Multisurface Interactive Environment. In *Proc. IHM*, pp. 18:1–18:8. ACM, New York, 2011. doi> 10.1145/2044354.2044376

[8] A. Bezerianos. Using Alternative Views for Layout, Comparison and Context Switching Tasks in Wall Displays. In *Proc. OZCHI*, pp. 303–310. ACM, New York, 2007. doi> 10.1145/1324892.1324956

[9] A. Bezerianos and R. Balakrishnan. The Vacuum: Facilitating the Manipulation of Distant Objects. In *Proc. CHI*, pp. 361–370. ACM, New York, 2005. doi> 10.1145/1054972.1055023

[10] A. Bezerianos, P. Dragicevic, and R. Balakrishnan. Mnemonic Rendering: An Image-based Approach for Exposing Hidden Changes in Dynamic Displays. In *Proc. UIST*, pp. 159–168. ACM, New York, 2006. doi> 10.1145/1166253.1166279

[11] X. Bi, S.-H. Bae, and R. Balakrishnan. Effects of Interior Bezels of Tiled-Monitor Large Displays on Visual Search, Tunnel Steering, and Target Selection. In *Proc. CHI*, pp. 65–74. ACM, New York, 2010. doi> 10.1145/1753326.1753337

[12] W. S. Cleveland. *The Elements of Graphing Data*. Wadsworth Advanced Book Program, 1985.

[13] W. S. Cleveland and R. McGill. Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association*, 79(387):531–554, Sept. 1984. doi> 10.2307/2288400

[14] W. S. Cleveland and R. McGill. Graphical Perception and Graphical Methods for Analyzing Scientific Data. *Science*, 229(4716):828–833, Aug. 1985. doi> 10.1126/science.229.4716.828

[15] M. Czerwinski, D. S. Tan, and G. G. Robertson. Women Take a Wider View. In *Proc. CHI*, pp. 195–202. ACM, New York, 2002. doi> 10.1145/503376.503412

[16] A. Endert, C. Andrews, Y. H. Lee, and C. North. Visual Encodings that Support Physical Navigation on Large Displays. In *Proc. GI*, pp. 103–110. CHCCS, Waterloo, Canada, 2011.

[17] R. Fisher. The Force of Contraction of the Human Ciliary Muscle During Accommodation. *Journal of Physiology*, 270(1):51–74, Aug. 1977.

[18] J. M. Foley. Depth, Size and Distance in Stereoscopic Vision. *Attention, Perception, & Psychophysics*, 3(4):265–274, July 1968. doi> 10.3758/BF03212742

[19] M. Gleicher, D. Albers, R. Walker, I. Jusufi, C. D. Hansen, and J. C. Roberts. Visual Comparison for Information Visualization. *Information Visualization*, 10(4):289–309, Oct. 2011. doi> 10.1177/1473871611416549

[20] E. B. Goldstein. *Sensation and Perception*. Brooks/Cole Publishing, Pacific Grove, USA, 5th edition, 1999.

[21] M. Hancock, M. Nacenta, C. Gutwin, and S. Carpendale. The Effects of Changing Projection Geometry on the Interpretation of 3D Orientation on Tabletops. In *Proc. ITS*, pp. 157–164. ACM, New York, 2009. doi> 10.1145/1731903.1731934

[22] C. Harrison and S. E. Hudson. A New Angle on Cheap LCDs: Making Positive Use of Optical Distortion. In *Proc. UIST*, pp. 537–540. ACM, New York, 2011. doi> 10.1145/2047196.2047266

[23] P. Isenberg, A. Bezerianos, P. Dragicevic, and J.-D. Fekete. A Study on Dual-Scale Data Charts. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2469–2478, Nov./Dec. 2011. doi> 10.1109/TVCG.2011.160

[24] R. Jota, M. A. Nacenta, J. A. Jorge, S. Carpendale, and S. Greenberg. A Comparison of Ray Pointing Techniques for Very Large Displays. In *Proc. Graphics Interface (GI)*, pp. 269–276. CIPS, Toronto, 2010.

[25] A. Khan, J. Matejka, G. Fitzmaurice, and G. Kurtenbach. Spotlight: Directing Users' Attention on Large Displays. In *Proc. CHI*, pp. 791–798. ACM, New York, 2005. doi> 10.1145/1054972.1055082

[26] S. Kim, X. Cao, H. Zhang, and D. Tan. Enabling Concurrent Dual Views on Common LCD Screens. In *Proc. CHI*, pp. 2175–2184. ACM, New York, 2012. doi> 10.1145/2207676.2208369

[27] R. Kruger, S. Carpendale, S. D. Scott, and S. Greenberg. Roles of Orientation in Tabletop Collaboration: Comprehension, Coordination and Communication. *Journal of Computer Supported Collaborative Work*, 13(5–6):501–537, Dec. 2004. doi> 10.1007/s10606-004-5062-8

[28] R. Kruger, S. Carpendale, S. D. Scott, and A. Tang. Fluid Integration of Rotation and Translation. In *Proc. CHI*, pp. 601–610. ACM, New York, 2005. doi> 10.1145/1054972.1055055

[29] C. A. Levin and R. N. Haber. Visual Angle as a Determinant of Perceived Interobject Distance. *Attention, Perception, & Psychophysics*, 54(2):250–259, Mar. 1993. doi> 10.3758/BF03211761

[30] M. A. Nacenta, S. Sakurai, T. Yamaguchi, Y. Miki, Y. Itoh, Y. Kitamura, S. Subramanian, and C. Gutwin. E-conic: A Perspective-Aware Interface for Multi-Display Environments. In *Proc. UIST*, pp. 279–288. ACM, New York, 2007. doi> 10.1145/1294211.1294260

[31] E. Pietriga, S. Huot, M. Nancel, and R. Primet. Rapid Development of User Interfaces on Cluster-Driven Wall Displays with jBricks. In *Proc. Engineering Interactive Computing Systems*, pp. 185–190. ACM, New York, 2011. doi> 10.1145/1996461.1996518

[32] F. H. Previc. Functional Specialization in the Lower and Upper Visual Fields in Humans: Its Ecological Origins and Neurophysiological Implications. *Behavioral and Brain Sciences*, 13(3):519–542, Sept. 1990. doi> 10.1017/S0140525X00080018

[33] G. Shoemaker, A. Tang, and K. S. Booth. Shadow Reaching: A New Perspective on Interaction for Large Displays. In *Proc. UIST*, pp. 53–56. ACM, New York, 2007. doi> 10.1145/1294211.1294221

[34] S. S. Stevens. *Psychophysics*. Transaction Publishers, New Brunswick, USA, 2nd edition, 1975.

[35] Stoney Brook University. Stony Brook University Receives NSF Grant to Design Revolutionary Reality Deck. Press Release, July 13 2010.

[36] D. S. Tan, D. Gergle, P. Scupelli, and R. Pausch. With Similar Visual Angles, Larger Displays Improve Spatial Performance. In *Proc. CHI*, pp. 217–224. ACM, New York, 2003. doi> 10.1145/642611.642650

[37] M. Wagner. *The Geometries of Visual Space*. Lawrence Erlbaum Associates, Mahwah, NJ, USA, 2006.

[38] W. M. Wiest and B. Bell. Steven's Exponent for Psychophysical Scaling of Perceived, Remembered, and Inferred Distance. *Psychological Bulletin*, 98(3):457–470, Nov. 1985. doi> 10.1037/0033-2909.98.3.457

[39] D. Wigdor, C. Shen, C. Forlines, and R. Balakrishnan. Perception of Elementary Graphical Elements in Tabletop and Multi-surface Environments. In *Proc. CHI*, pp. 473–482. ACM, New York, 2007. doi> 10.1145/1240624.1240701

[40] B. Yost, Y. Haciahmetoglu, and C. North. Beyond Visual Acuity: The Perceptual Scalability of Information Visualizations for Large Displays. In *Proc. CHI*, pp. 101–110. ACM, New York, 2007. doi> 10.1145/1240624.1240639

[41] B. Yost and C. North. The Perceptual Scalability of Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):837–844, Sept./Oct. 2006. doi> 10.1109/TVCG.2006.184

Evaluating Multi-User Selection for Exploring Graph Topology on Wall-Displays

Arnaud Prouzeau, Anastasia Bezerianos, Olivier Chapuis

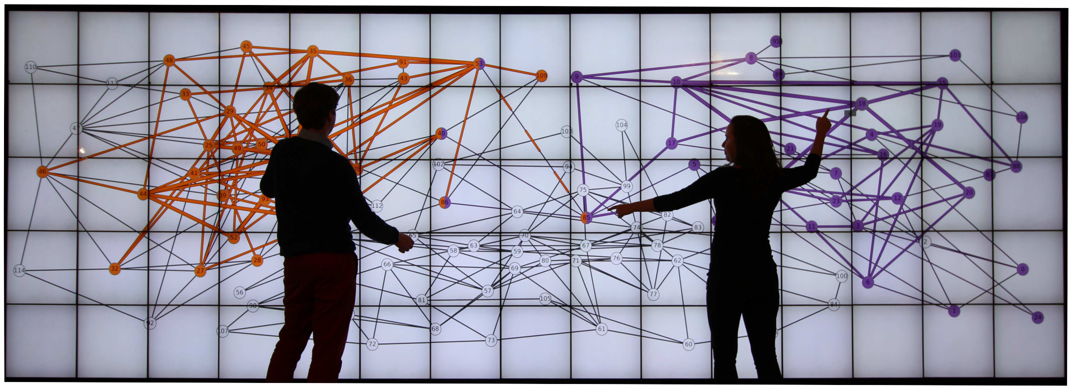


Figure 1. A pair using the propagation technique to explore a graph. They discuss two communities, in orange and purple, selected using the propagation technique. The communities are linked by a specific node shown by the right user. The remaining 3 orange-purple nodes show how by propagating the purple community, it flows into the orange one through this node.

Abstract—Wall-displays allow multiple users to simultaneously view and analyze large amounts of information, such as the increasingly complex graphs present in domains like biology or social network analysis. We focus on how pairs explore graphs on a touch enabled wall-display using two techniques, both adapted for collaboration: a basic localized selection, and a propagation selection technique that uses the idea of diffusion/transmission from an origin node. We assess in a controlled experiment the impact of selection technique on a shortest path identification task. Pairs consistently divided space even if the task is not spatially divisible, and for the basic selection technique that has a localized visual effect, it led to parallel work that negatively impacted accuracy. The large visual footprint of the propagation technique led to close coordination, improving speed and accuracy for complex graphs only. We then observed the use of propagation on additional graph topology tasks, confirming pair strategies on spatial division and coordination.

Index Terms—Wall-Displays; Multi-user interaction; Graph visualization; Selection techniques; Co-located Collaboration

1 Introduction

Graph structures, consisting of vertices and edges, exist in various application areas: in social networks they are used to represent people and their relationships, in molecular biology proteins and their interactions, in transport networks they can represent airflight routes, etc. Graph data structures are frequently represented as node-link diagrams, but like many visual representations of large datasets today, they can be too wide to view comfortably on regular screen monitors [64].

High-resolution wall-sized displays [8, 54] are promising data analysis environments, as their size and high pixel density allow simultaneous viewing, comparison, and exploration of large amounts of data. Their size can also comfortably accommodate multiple viewers, supporting collaborative analysis [60]. Despite their promise as collaboration platforms, they have received little attention for graph exploration. We take a step in this direction.

- *Arnaud Prouzeau is with Univ Paris-Sud & CNRS (LRI), Inria, and Université Paris-Saclay. E-mail: arnaud.prouzeau@lri.fr*
- *Anastasia Bezerianos is with Univ Paris-Sud & CNRS (LRI), Inria, and Université Paris-Saclay. E-mail: anab@lri.fr*
- *Olivier Chapuis is with Univ Paris-Sud & CNRS (LRI), with Inria, and Université Paris-Saclay. E-mail: chapuis@lri.fr*

We present a systematic study of how pairs use a wall-display to solve topology based tasks, that are common components of more complex graph analysis tasks [37]. We study how the choice of interaction technique supports or hinders pairs collaborating on these tasks. We focus on techniques for selection, a fundamental visualization task, as it is a pre-requisite to many interactions such as filtering, comparisons, details on demand, etc.

We adapt two general purpose graph selection techniques for use by multiple users on a touch-enabled wall-display. Our baseline is an extension of basic node/edge selection for multiple users. It is easy to master, and has a limited, and thus fairly localized, visual footprint on the wall display, that does not interfere with colleagues' work. The propagated selection extends for multiple users the idea of transmitting a selection to neighboring nodes/edges [21, 42]. It highlights the connectivity structure of the graph (Figure 1), but may have a large visual footprint that disturbs colleagues.

We first assess the impact of selection technique on pairs conducting a specific topology analysis task, namely identifying a shortest path. As there is no work on pairs working on such tasks on wall-displays, we tease out effects due to the technique or due to collaboration, by also studying single user selections. We

This is the authors' accepted version of the article (sent 4 March 2016; revised 17 June 2016; accepted 14 July 2016). The definitive version will appear in IEEE Transactions on Visualization and Computer Graphics X(X), XXXX-XXXX, 201X. DOI: http://dx.doi.org/10.1109/TVCG.2016.2592906

then examine how propagation, the most promising technique, is used by pairs on other graph analysis tasks [37]. Our studies are conducted on a touch enabled wall-display, instead of interacting using mice and keyboards, as mobility allows viewers to perform implicit zooming [7] and correct for visual distortions [10].

We contribute: (i) The adaptation of two graph selection techniques for collaboration on wall-displays. (ii) The controlled study of how pairs use these techniques on a graph topology task (shortest path identification) on wall-displays. (iii) A discussion and observational study on how one technique, propagation, supports different topology tasks. And (iii) a set of design implications: as pairs divide the work spatially, even when tasks are not spatially divisible, the use of a localized selection technique may be detrimental to performance in complex graphs; while a technique with global reach leads to tighter collaboration and coordination, that is more effective and accurate for such graphs.

2 Related Work

A wide range of topics surrounding large displays have been studied in HCI and Visualization. We focus on the most relevant, namely visual exploration and collaboration on wall-displays, in particular exploration of graphs, and the idea of transmission.

2.1 Walls in Visual Exploration

Wall-sized displays have been studied in the context of information visualization and analysis, as they can naturally display a large amount of visual information. Previous work comparing large displays to traditional desktops [40, 58] or to smaller displays [52] has shown performance improvements when moving to larger displays. Considering visual analysis in particular, Yost and North [68] tested several data visualizations for their scalability when moving from small to large displays. They found their visualizations to scale well for the tasks of finding detailed and overview information, and note that spatial encoding of information was particularly important on large displays. Jakobsen and Hornbæk [31] examined the interplay between display size, information space size and scaling, and found that all these factors need to be taken into account, and that increased display size did not improve navigation performance in tasks where targets are visible at all scales. Reda et al. [52] found that larger displays encourage longer visual analysis sessions, and result in deeper and more complex insights. Finally, Rajabiyazdi et al. [51] observed that they can lead to previously missed insights in multiple disciplines.

Beyond their benefits, researchers have studied specific issues related to visual perception of wall-displays due to their scale. Endert et al. [17] discuss how a viewer's distance from the wall influences the visual aggregation of displayed information. Bezerianos et al. [10] showed large discrepancies in the perception of basic visual encodings depending on viewing distances and angles, that nevertheless decrease if appropriate physical navigation is used. Ball et al. [6] compared the benefits of added peripheral vision vs. physical navigation, and found that physical navigation influenced task performance while peripheral vision did not. Isenberg et al. [28] blended two visualizations so that each is perceived at a different viewing distance from the wall. Collectively this work stresses the importance of physical navigation for visualization and visual perception tasks, even if it is not necessarily better than virtual navigation in classification tasks [33].

Despite the importance of physical navigation, a large body of this past work either assumes the use of mouse and keyboard, or simply does not study interaction. Nevertheless, recent work supports both interaction and physical navigation using handheld devices or direct touch. Handhelds are used as touch-pads to conduct classification tasks [40], or as a support for physical controllers [34] or for explicitly sketching interactive slider controllers [61] to conduct multi-dimensional data exploration. In a sense-making task, Jakobsen and Hornbæk [32] allow users to move freely and use direct touch to interact with the wall. We similarly use touch to support pairs working on graph topology tasks.

2.2 Walls in Collaborative Analysis

When it comes to co-located collaborative work and visual analysis (see [22, 29] for reviews), work has focused mostly on tabletops, and "small" vertical displays (SDG and whiteboards). For example, researchers have explored how collaborators shift from tight to loose work coupling [59], how users divide space (territoriality) [62], how they analyze text documents [30], compare tree visualizations [26], etc.

There are few works on co-located collaborative work on large (ultra-high resolution) walls. Notable exceptions are recent work by Jakobsen and Hornbæk [32] that studies the behavior of a pair of users in a sense-making task, and Liu C. et al. [39] that studies the effect of different collaboration styles and interaction in a classification task. Both these works stress the importance of users' coordination (possibly at distance) in these environments. Our work is along the same lines, but focuses on graph analysis in particular, and the effect of different interaction techniques on it, a topic that so far has not been studied.

2.3 Graph Exploration on Large Surfaces

Collaborative analysis is one of the next challenges of the analysis of graphs [64]. Existing systems support mainly remote collaboration (e.g. [69]). Less work has targeted co-located analysis, like that by Isenberg et al. [27] that retrofitted an existing graph visualization application for use by multiple analysts with mice and keyboards. We are not aware of any work that studies analysis of graphs by multiple users moving freely in front of wall-displays.

Although work on graph exploration using wall-displays is limited, researchers have identified their potential early on. For example, Abello et al. [1] used a wall display to visualize communication network data. Later, Mueller et al. [47] designed an algorithm to interactively layout graphs optimized for tiled displays and distributed environments, while Marner et al. [41] let users interactively adapt the layout on the wall using a mouse and keyboard. Finally, Lehmann et al. [38] leverage physical navigation as an implicit interaction, using the viewer's distance from the wall to adjusted the level of detail of a graph, and Kister et al. [36] use it to move a lens with contextual information. This past work on wall displays does not study the use of explicit interactions (e.g., selections) during collaboration, as we do.

Finally, although not explicitly testing collaboration, researchers have introduced multi-touch techniques for manipulating graphs on interactive tabletops. For example, Henry Riche et al. [23] use multi-touch interactions to fan out links leaving a node, to bundle them, or use link magnets to attract certain types of links. Schmidt et al. [55] alter link trajectories, pin, or make them vibrate by plucking them. This work introduces multi-touch techniques on tabletops for different purposes. While we also use touch, we focus specifically on selection and study how pairs use it to perform graph topology tasks on wall-displays.

2.4 Graph Exploration using Transmission

Visual analysis of graphs is a long standing field, with numerous research questions (see [24, 64] for reviews). We focus on techniques related to our propagation selection (section 3.2), that use the idea of propagating/transmitting information to neighboring nodes or links that is central to graph analysis (e.g., [53]).

As graph structures can be very large, exploration is often localized on interesting nodes and their neighbors. For example, van Ham and Perer [63] designed a Degree-of-Interest function for graph exploration that first proposes interesting nodes, and lets the user indicate interesting nodes to expand to. Archambault et al. [4] use specifically the notion of distance to progressively reveal and render nodes proximal to a node of interest from within a larger graph hierarchy. Moscovich et al. [46] propose interaction techniques for panning within a graph, or bringing neighbors closer, based on the graph's connectivity. Finally, egocentric techniques (e.g., [67]) re-layout graphs by focusing around one node and laying out the rest based on their distance from it; or focus on two nodes [13] and highlight their common neighbors. This work can lead to a user-driven re-layout of the graph, that may disrupt the work of other viewers in a multi-user setting.

Other techniques related to propagation preserve the layout. Heer et al. [21, 20], allow users to highlight the contour of the 1st or 2nd degree neighbors, or the connected component of a node, by hovering over it or by using repeated mouse clicks. McGuffin and Jurisica [42] propose techniques to locally select and manipulate nodes, including a menu option that selects a node's neighbors of increasing distance progressively. Ware and Bobrow [65] evaluate different means of highlighting connections to neighbors of arbitrary degree specified by a text field, and found that motion representations are not better than static highlighting. We extend this notion of propagated selection to multiple origin nodes, providing appropriate input and visual design, to support such selections by multiple users.

3 Interaction Techniques

Our goal is to investigate how interaction techniques affect multiple users working on graphs. We focus on selection, as it is a required first step for many other visualization tasks, such as filtering, comparison, details on demand, etc. Two techniques were considered, a simple selection (*Basic*), and one based on the graph's connectivity structure (*Propagation*). These techniques were chosen due to their properties: they can benefit graph exploration differently but also face different challenges when adapted for multiple users on wall-displays. We describe next how we adapted the techniques for touch interaction on wall-displays, and for collaborative use. Each description finishes with a summary of the technique's properties, motivation for their use, and possible challenges when used in a multi-user context on wall displays.

3.1 Basic Selection

Basic is inspired by colored selections available in graph visualization software extended for multiple users. We chose it to investigate the limits of basic selections in collaborative settings.

3.1.1 Interaction and Visual Design

A node (or link) is selected by tapping on it once, and deselected if tapped again. Inspired by previous work [4, 21], we also highlight the links (or nodes) attached to it so as to demonstrate its connections, but do not re-layout the graph to avoid disrupting collaborators. Given that we do not have keyboard modifiers, and wanted to keep the touch input vocabulary simple, we decided to allow users to modify this selection in the following way: if the user taps on a node adjacent to an existing selection (direct neighbor), then this node is added to the selection and it, and its links, are highlighted with the selection's color. If the node is adjacent to more than one existing selections it takes the color of the last edited selection. Tapping on a selected node removes it from the selection. This way users can edit their selections with simple taps, keeping the input vocabulary very simple. We chose to not use lasso-type selections that require dragging to select multiple items, as they are not well suited to large interactive surfaces, such as walls, where prolonged dragging is inaccurate, fatiguing [25], and often disrupted by bezels in tiled walls.

Our wall, similar to many touch enabled surfaces, does not differentiate between users. Nevertheless, it is important for colleagues to differentiate their work. Thus, if users tap on nodes that are not adjacent to existing selections, we assume a new selection is being made (potentially by a different user) and assign it a new color, chosen randomly from a set that is easily distinguishable.

3.1.2 Summary

Basic extends the simple selection available in graph visualization software to selection of multiple nodes/edges by multiple users. It is familiar, easy to understand, and our design ensures it relies on simple taps. It has a small visual footprint as it selects a single node and its edges at a time, and thus will likely not disrupt collaborators when used in a multi-user context on wall displays. Nevertheless, it may require extensive physical movement if users need to select multiple nodes that are far away on the wall.

3.2 Propagation Selection

As an alternative, we investigate *Propagation* selection, based on the idea of progressive transmission of a selection to neighboring nodes. Propagation allows local interaction on a node that can highlight its influence across a larger area on the graph (and wall), without requiring extensive physical movement that can be tiring. Nevertheless, it may have a large visual footprint if neighboring nodes are far away, potentially disrupting collaboration.

Variations of the propagation selection from past work (e.g. [21, 42]) allow a single user to either highlight neighboring nodes up-to a specific degree only [21], usually 2, or use a menu or text option to select a node and its neighbors of a certain degree [42]. We explain how we adapted the technique allowing multiple users to easily expand the selection to the *n*-th degree using simple touch interactions. We finally describe its properties and how it can be used to perform topology-based tasks [37] when analyzing graphs.

3.2.1 Interaction

Propagation allows users to select a node, which we will refer to as the *origin*, and then propagate the selection first to its neighbors, then to their neighbors, and so on. Propagation of a selection is done through a series of taps (clicks) on a node. The first tap selects the node itself (Figure 2-a), and the following taps propagate the selection progressively to the neighboring edges and nodes: the second tap adds to the selection outgoing links and first-degree neighbors of the origin (Figure 2-b), and so on for all following taps[1] (Figure 2-c). If users continue tapping,

[1] Propagation starts either from a node or a link. To simplify the discussion we talk about node propagation, but we use a similar selection pattern for links: link selected first, adjacent nodes and their links next, and so on.

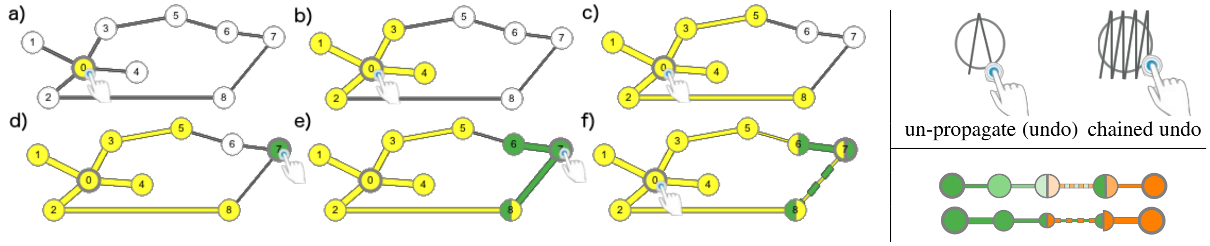


Figure 2. On the **left** multiple propagations: (a) a first tap on node 0 selects it; (b) a second tap propagates the selection to immediate neighbors; (c) and a third tap to 2nd degree neighbors (notice the difference in link width according to distance); (d) a tap on node 7 selects it with a new color; (e) a second tap selects its neighbors, one of which (node 8) is shared with the first propagation and has both colors; (f) a fourth tap on node 0 propagates the first selection a third time, resulting in nodes 6,7,8, and link 8-7 being shared between propagations, with the color and width on shared link 8-7 alternating. On the **top right** gesture to undo one propagation step on a node (left) and chained undo for backtracking multiple steps (right). On the **bottom right** design variations for displaying propagation distance using color intensity (top) and node-link size (bottom).

propagation continues until no more nodes can be reached from (are connected to) the origin node. Thus a propagated selection is a progressive query selection, that adds elements connected to the origin node at progressively increasing distances.

We note that the first step of propagation (selecting only the node) is not the same as *Basic* (selecting the node and its edges). We made this design choice as initial feedback indicated that the metaphor of transmission is better served if we consider that each tap opens the flow of transmission from the selected nodes to their neighborhood (both links and nodes).

To accommodate multiple users working in parallel, when users select a node that is not part of an existing propagated query, it becomes the origin of a new propagation selection (Figure 2-d). If they select a node already inside a propagation query (but not its origin), the query expands to also include propagations from this new origin. Thus one propagation query can have multiple origins.

As we designed the technique for touch surfaces, we chose a simple crossing zig-zag gesture to undo propagation steps. When performed on the origin, it backtracks the propagation by one step (Figure 2 left). The gesture can be chained to perform multiple backtracks without lifting a finger, undoing quickly several propagation steps in one interaction. When the selection is reduced to a single node (the origin), this gesture unselects the node.

A crossing gesture on an element (node or link) that is not the origin of a propagation, removes this node from the selection and blocks future propagation paths of this selection to go through it.

3.2.2 Visual design

Nodes and links in a propagated selection share a common color (as traditional color queries). Propagation origins stand out with a thicker border (Figure 2-a), and new propagations are assigned a different color, similar to *Basic* (Figure 2-d).

Due to the propagation of selections, a node can be selected by two or more colors. The node in question is divided visually into slices equal to the number of selections, and given the respective colors (Figure 2-e).

Links can similarly be part of several selections. Dividing them in segments equal to the number of selection colors (similar to nodes) could lead to few, but long, colored segments if links are long. Thus the multiple colors could be hard to see locally on a wall display. We decided instead to streak (dashing pattern) the links with the selection colors (Figure 2-f). We fixed the number of streaks to seven, as we observed that on our wall they were still visible locally, even on long links. Moreover, as the fixed number of streaks have different length depending on the total length of the link, they give locally an indication of its overall length.

We explored different design variations to emphasize the distance of elements (nodes and links) from the origin. This is of interest both within a single selection (to identify the farthest elements), but also for elements that are part of multiple selections to identify which origin is closest. As color is already used in selections, we considered other visual variables (Figure 2 right). Color intensity that drops with distance was considered, but rejected, as the perception of intensity may be affected by viewing distance and angle across the wall-display, and color intensity may vary across screens in tiled wall-displays [57]. We thus chose the size of elements, i.e., the thickness for the links and the radius for the node slices. While testing our prototype, we observed that as nodes have multiple incoming links, it is hard to identify which path and origin is responsible for the shortest distance that determines their size. Thus to avoid confusion and reduce clutter, we chose to only display distance information on the links.

As the thickness of selected links indicates their distance to the propagation origin node, the thicker the link the closer to the origin it is. We chose to display three visual levels of thickness: links with maximum thickness are linked to first-degree neighbors, ones of medium thickness link first and second-degree neighbors, and all remaining links selected through propagation have a similar minimum thickness. We found that more levels led to small variations in thickness that were hard to perceive in dense graphs. When a link is traversed multiple ways inside a selection (e.g., there are multiple origins in a selection, or the link belongs to multiple paths of different length), the link thickness is determined by the smallest distance to the closest origin in the selection.

3.2.3 Propagation Properties, Support for Graph Analysis

Multiple propagations allow multiple users to simultaneously explore different parts of the graph with their own color, examining connectivity relationships in different areas, as well as interactions between their selections made visible by the combined colors in nodes and links when propagations coincide. They also highlight relationships that may span large distances on wall displays, without the need for extensive physical movement.

Multiple propagations can also aid a single user to visually conduct basic set operations between selections. For example, the union of two or more propagation selections is the set of all the colored nodes. Their intersection are the nodes and edges colored by all respective colors simultaneously. And the difference of two selections (i.e. elements in one but not in the other), are all nodes and edges that colored by a single color.

Thus propagation from multiple nodes could be used to answer fairly complex topological questions, such as identifying all common neighbors of N-degree or less of multiple actors in a social network (union of N-level propagations), all the co-authors of one researcher that are not co-authors of her colleagues in a co-authorship network (difference of 1st level propagations), etc. We consider next topological tasks, such as the ones described by Lee et al. [37], that are well supported by propagation.

- Adjacency (direct connections): It is trivial to find and highlight the neighbors of a node by propagating one level. Nevertheless, there is no clear strategy for how to identify the node with most neighbors (highest degree) using the propagation technique.
- Accessibility (direct or indirect connections): This set of tasks are well supported by propagation. Nodes accessible from an origin are colored by the propagation. And the propagation level highlights nodes at distances less or equal to that level.
- Common Connections: To find the common neighbors of two or more nodes, we can propagate from each of these origin nodes and identify nodes that have both colors (i.e. belong to both propagation selections). And as before we control the distance of neighbors.
- Connected Components: To identify discrete connected components, i.e. subgraphs not connected to each other, we can choose a node and propagate until no more nodes are added, thus identifying a connected component. Repeating the process with uncolored nodes will identify the remaining connected components.
- Shortest distance between two nodes: The length of the shortest distance between two nodes can be found by propagating from one node and counting the number of propagation steps to reach the second. Nevertheless, determining the *actual shortest path* is more challenging: although the path is part of the propagated selection, it can be hard to identify it within all the selected elements, particularly in dense graphs.

This is a non exhaustive list of tasks well supported by propagation, and tested later on. More complex strategies could be devised for other tasks, to find for example articulation points or bridges (a node or link that is the only connection between two subgraphs).

3.2.4 Summary

Our adapted *Propagation* technique for interactive surfaces uses fast taps to expand, and a crossing gesture to backtrack. We support multiple propagations that can aid with several graph topology tasks. By design, propagation can select several nodes quickly, based on the connectivity structure of the graph, without requiring extensive moving around the wall-display. Nevertheless, it may cause visual disturbance in well connected graphs, as it will quickly span the entire graph, and it may disrupt the work of colleagues if links cross their work space.

4 Experiment 1: Propagation vs. Basic

It is unclear how *Propagation* and *Basic* selection will affect multiple users working on a wall-display. As there is little work on graph analysis on wall-displays in general, we also studied an individual user context, to tease out effects due to collaboration and ones due to the techniques.

As an instrument for this exploration we chose a well-defined topology task, the identification of the shortest path between two nodes, for several reasons. First, identifying the shortest path, or variations thereof, is a task used often in controlled graph evaluation studies (e.g. [59, 15]) and can be fairly involved in complex graphs. It requires an understanding of both the local context of nodes (identifying neighbors), as well as more global structure information, as a shortest path is not necessarily small in absolute distance. And as it is a well-defined, closed task, with an objective solution, it is well suited for controlled experiments. Second, the task is not clearly divisible, as a more global understanding of the graph structure is required. Thus it is unclear if multiple users working together would fare better than single users. As it can be performed individually, it gives us the opportunity to compare individual vs. multiple user work. Finally, and very importantly for our purposes, the task does not bias against *Basic* as it is not trivial to do with *Propagation*. As *Propagation* highlights a large number of possible paths (explained in section 3.2.3), this task could reveal issues with visual clutter caused by *Propagation*.

Based on the design and properties of the two techniques, we formulate the following general hypotheses:

H1 In both *Individual* and *Multi-user* contexts, performance (time & accuracy) will be better with *Propagation* than *Basic*.
H2 With both techniques, performance will be better in the *Multi-user* context than in the *Individual* context.
H3 *Propagation* will result in less participant movement, but will cause higher visual disturbance.

4.1 Experimental Design

4.1.1 Participants

We recruited 16 participants in pairs (6 females, 10 males), aged 23 to 39, with normal or corrected-to-normal vision. Pairs knew each other beforehand. Participants were HCI and visualization researchers or graduate students, with experience in reading graphs. Most (15/16) reported using at least once a day a device with touch interaction, and having already used a wall-display (13/16).

4.1.2 Apparatus

We used an interactive wall made of 75 LCD displays (21.6 inches, 3mm bezels each), composing a 5.9m $\times$ 1.96m wide wall, with a resolution of 14 400 $\times$ 4800 pixels (Figure 1). The wall was driven by a rendering cluster of 10 computers. A PQ labs[2] multi-touch layer allowed for direct touch over the wall. Participants' positions were tracked by a VICON motion-capture system[3].

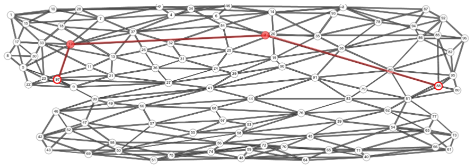
The experimental software ran on a master machine connected to the cluster through 1Gbit ethernet, and was implemented in Java using the ZVTM[4] Cluster toolkit [49]. The operator controlled the experiment progression using a smartphone running an android application implemented with the Smarties[5] toolkit [11].

4.1.3 Graph Types

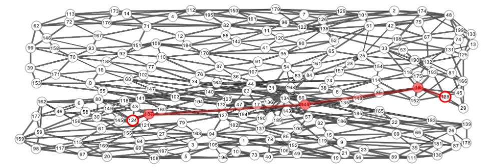
We considered two different Graph types:

- *Planar*: These graphs can be drawn without edge crossings. Transport networks (e.g. subway or air-routing networks) are often planar. We generated them using an algorithm inspired by Mehadhebi [43] to design air route networks.
- *SmallWorld*: These illustrate the small-world phenomenon identified by Milgram [45] in social networks, where most actors are linked by short chains of acquaintances. Social networks, communication networks, and airline networks are often Small-world graphs. We generated them using Kleinberg's algorithm in the JUNG toolkit [48].

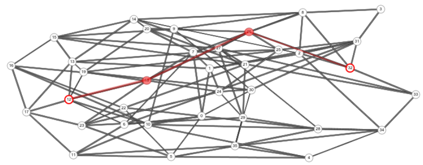
[2] pqlabs.com [3] vicon.com [4] zvtm.sourceforge.net [5] smarties.lri.fr



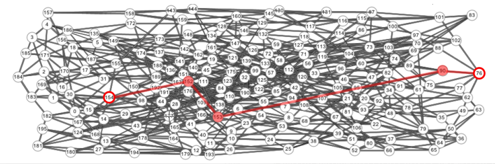
Planar-Low (N: 100 and L: 288)



Planar-High (N: 200 and L: 582)



SmallWorld-Low (N: 36 and L: 103)



SmallWorld-High (N: 196 and L: 588)

Figure 3. Graph examples used in Experiment 1 with their number of nodes (N) and links (L). Colored paths are for illustration purposes only, and highlight the shortest path between the two target nodes. During the experiment participants were only shown the first and last node (target nodes).

In a pilot study (2 pairs) we tested three types of generated graphs: *Planar* and *SmallWorld* ones, as well as randomly generated ones inspired by Ware and Mitchell's [66] algorithm. Participants' performance with the random graphs was very similar (time, errors, subjective comments) to *SmallWorld* ones, and we thus removed them from the experiment.

4.1.4 Complexity

To explore graphs of different complexity, we created two variations for each graph type, *Low* and *High* COMPLEXITY. We generated them by varying structural characteristics, such as number of nodes and edges and mean shortest path, and visual aesthetic criteria that can affect readability, such as visual density and number of edge crossings [50]. Visual density is calculated as the ratio of pixels occupied by nodes and links, over the entire surface used to calculate the layout (discussed later).

GRAPH	COMPLEXITY	#Nodes	#Edges	Shortest Path	Visual Density	#Crossings
Planar	*Low*	100	288	4.27	0.06	179
	High	200	582	5.69	0.10	627
SmallWorld	*Low*	36	103	2.27	0.02	249
	High	196	588	3.55	0.12	4879

Table 1. Mean metrics of the graphs used in the experiments.

Table 1 reports mean values for the metrics of graphs used in the experiment. We note that our purpose was not to equate all metrics across graph types, but rather to create "difficult" and "easy" variations for each type (Figure 3). For high complexity graphs of all types (*Planar* and *SmallWorld*), we chose high complexity graphs with similar visual density, i.e. the amount of ink or clutter, and number of nodes and edges. For low complexity graphs we found in a pilot (1 pair) that tasks on *Planar* graphs with less than 100 nodes were trivial and did not require interaction. Thus for the low complexity variation of *Planar* we chose higher visual density than for *SmallWorld* ones.[6]

Density and crossings depend on the layout used to draw the graph. To ensure consistent drawing across graph types, we used for all graphs the ISOM layout [44]. We tested several layout algorithms, such as classic force directed ones [16, 18], that position neighboring nodes close together and minimize edge crossings. Nevertheless, the tested force directed layouts [18] generated larger number of edge crossings compared to ISOM, a metric associated with readability [50], and did not uniformly fill our wall space. We thus moved to the ISOM layout that optimizes similar quantities to force directed layouts, while ensuring best coverage of our wall surface, and resulting into a smaller number of crossings. The ISOM layout is well adapted to planar graphs, but as other layout algorithms, it can lead to layout calculations that break somewhat the visual planarity of structurally planar graphs, as can be seen in Table 1. The same graphs and layouts were seen in both techniques (see Procedure), to keep this experimental factor consistent across techniques.

4.1.5 Task

Participants were asked to identify the shortest path between two target nodes. Target nodes were positioned in height at the middle 60% of the wall, thus not too high or too low to reach; and were spaced by a distance of at least 50% and 75% of the width of the wall to ensure paths were not too localized.

For each graph type and complexity we generated three variations to be used as "replications". In each of the three variations, we selected a path of LENGTH 3,4 and 5 respectively[7]. Paths of the given length were chosen automatically (using exhaustive search) to fulfill the following criteria: (i) the first and last node, that would become the "target nodes", met the above placement criteria; and (ii) all nodes in the path similarly fell into the middle 60% of the wall to ensure they were easily selectable.

4.1.6 Procedure and Design

The experiment was divided in two sessions, an *Individual* and a *Multi-user* one. To counterbalance these conditions, half of the participants did the *Individual* session first and half the *Multi-user* session first. In the *Multi-user* session, pairs saw both techniques (within-subject design), and the order of presentation was counterbalanced across groups. To end-up with an equal sample of group and individual sessions, in the *Individual* sessions each participant only saw one technique (between-subject design), chosen at random. *Individual* sessions lasted approximately 25 min, and *Multi-user* ones 40 min.

6 In our pilot we considered a no-interaction condition, but found that for our graphs (both *Low* and *High* complexities), tasks were respectively either very hard (double the time) or impossible to do without interaction to help trace one's process. Thus we did not test the "no interaction" condition further.

7 The use of LENGTH as a replication factor was justified, as there was no interaction between LENGTH and TECH, CONTEXT, GRAPH (see Results).

Overall our mixed experiment design consisted of: 8 sessions (pairs or individuals) $\times$ 2 CONTEXTS (*Individual*, *Multi-user*) $\times$ 2 TECHS (*Basic*, *Propagation*) $\times$ 2 GRAPHS (*SmallWorld*, *Planar*) $\times$ 2 COMPLEXITIES (*Low*, *High*) $\times$ 3 LENGTHS (3, 4 and 5) = 384 measured trials.

For each TECH in both contexts, participants conducted 7 training trials before proceeding to the main experiment. Trials began with a screen indicating the position of the two target nodes, to ensure visual search was not required. Participants were then shown the actual graph with the target nodes highlighted. They then interacted with the wall display to find a shortest path, and when they had an answer they verbally indicated to the experimenter to stop the timer, and showed their solution. An experimenter followed the discussion to ensure they did not "cheat", i.e. report done before finding all nodes. No such cases were observed. If their answer was correct, they would proceed to the next trial. If their answer was wrong, the trial was marked as an error. Nevertheless, the task resumed and participants had to continue the trial until they found the correct answer. This ensured participants did not rush to give partially formed answers. At the end of the sessions participants filled a questionnaire on the perceived load and visual disturbance, and provided general preferences and subjective comments.

We chose a verbal indication of when pairs had reached a consensus, because in a third pilot (1 pair) we found that other procedures did not always ensure a consensus. We first provided each participant with a mobile device with a "done" button. We observed that choosing as a trial completion the first time one of the two participants pressed "done" was problematic, as they often did so while the other was still working. We also considered the time both participants had pressed "done", but found that some would occasionally forget to press their button while discussing with their partner. We next provided a single mobile device to only one participant. Although in most cases a very clear verbal agreement would take place before they pressed "done", occasionally the participant holding the mobile would forget getting verbal agreement and would press the button too soon. Thus we decided to enforce verbal agreement between participants, by asking them to instead tell the experimenter together when they were done, a process they practiced during training. When the two verbal indications were given the experimenter would log the time.

For each technique and context, participants were shown the *Low* complexity graphs first to ease them into the task, while the order of graph type and path length was randomized, but consistent, across participants. The same graphs were seen across techniques and collaboration contexts, but to avoid learning we used mirrored versions of the graphs on the x and/or y axis (resulting in 4 variations per graph).

Participants were instructed to be as fast as possible while avoiding errors. We recorded the time to the first given answer as our task completion time (*Time*), and the count of incorrect answers. We logged kinematic data of participants' movements using a motion tracking system, and video recorded the sessions.

4.2 Results

We report on the measures: (i) *Time* taken by participants to state for the first time that they completed the task, approximating expert behavior. When the first answer was wrong trials were marked as errors and the task would resume to discourage participants from rushing through trials (but the extra timing was not logged). (ii) *ErrorRate*, i.e. the percent of trials where participants provided incorrect answers. (iii) *TraveledDistance* by participants in front of the wall. (iv) Subjective *rating* of visual disruption.

Statistical Method – Following recommendations from the APA [3], our analysis and discussion on continuous measures (*Time*, *TraveledDistance*) are based on estimation, i.e., effect sizes with 95% confidence intervals (CI). Our confidence intervals were computed using BCa bootstrapping. Error bars in our images reporting means, are computed using all data for a given condition.

When comparing means, we average the data by participants/groups (random variable) and compare the two conditions globally using a $(-1, 1)$ contrast (between-subject case), or by computing the CI of the set of differences by participants/groups (within-subject case). In our images we display the computed CI of the differences, and report the corresponding Cohen's d effect size, that roughly expresses the difference in standard deviation units. Finally, for completeness, we also report p values. These are computed as an approximation of the smallest $p \geq 0.001$ such that the $100.(1-p)\%$ CI interval does not contain 0 (i.e., we compute the "largest" I-levels that lead to a "significant" result)[8].

To compare errors and Likert results we use non-parametric tests (Wilcoxon rank sum), which are more adapted to bi-valued and ordinal measures.

As mentioned, LENGTH was used as a replication factor, and as such is not considered as part of the analysis. Nevertheless, we conducted a-posteriori tests and verified that although there was a difference between the 3 length variations in time and errors, there were *no interaction* effects between length and interaction technique, context, or graph type. We also did not find any learning effects due to technique presentation order.

4.2.1 Time

Individual: When working individually, participants were faster with *Propagation* (29.3 s) than *Basic* (54.1 s). To better understand the nature of this difference, we looked separately at each COMPLEXITY and GRAPH. Our analysis (Figure 4) shows *Propagation* consistently outperforming *Basic*, with the effect being stronger in *SmallWorld-High* (most complex graphs).

Multi-user: Similarly, *Propagation* (22 s) was measurably faster than *Basic* (30 s) for pairs, even though the difference was not as pronounced. Looking at conditions in detail (Figure 5), the effect mainly exists in the *High* complexity graphs.

Individual vs. Multi-user: Individuals were slower with *Basic* (almost double the time) than with *Propagation*. This tendency was also visible in the *Multi-user* condition, although mainly for the larger graph sizes. This indicates that *Propagation* is more efficient, in particular for larger and complex graphs.

When we compare the *Individual* and *Multi-user* condition, mean times for both *Basic* and *Propagation* were better for pairs, but this difference was not measurable (Figure 6-left). However, examining the different complexities, we found a measurable time improvement for *Basic* when collaborating on *Low* complexity graphs, and a measurable improvement for *Propagation* when collaborating on *High* complexity graphs (Figure 6-right). This indicates that collaboration does not compensate for the weakness of *Basic* for complex graphs (in particular the *SmallWorld-High* ones). While with *Propagation*, one user is as effective as pairs for simple graphs, but that the collaboration benefit is seen in more complex graphs.

8 A CI of a difference that does not cross 0, can be read as "significant".

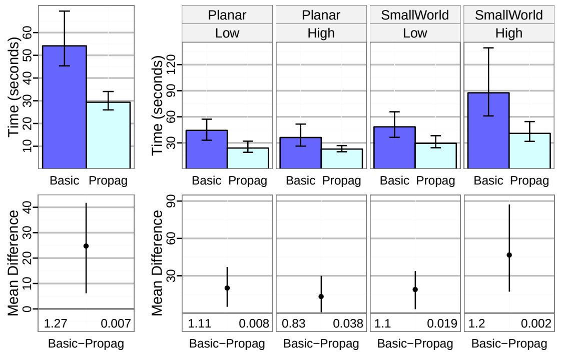


Figure 4. (Top) Average time to complet the task per TECH in the individual user case, aggregated on the left, and by GRAPH × COMPLEXITY conditions on the right. (Bottom) Corresponding 95% CIs for the mean differences *Basic − Propagation* used in analysis: bottom left numbers show the Cohen's d effect size and the right ones the p values. This convention is followed in all images.

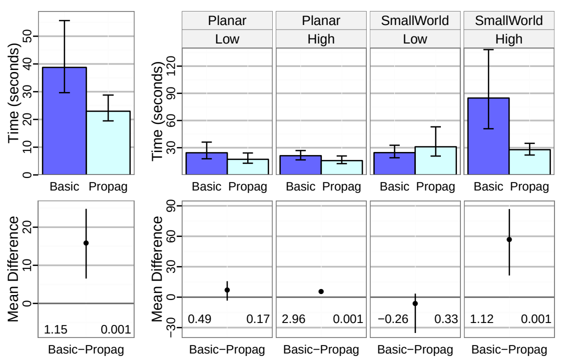


Figure 5. (Top) Average time to complet the task per TECH in the multi-user case, aggregated on the left, and by GRAPH × COMPLEXITY conditions on the right. (Bottom) Corresponding 95% CIs for the mean differences *Propagation − Basic*.

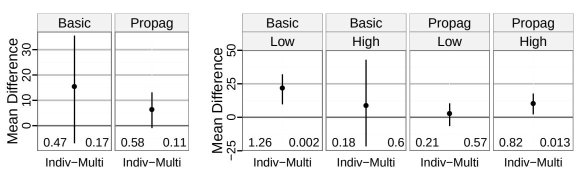


Figure 6. 95% CIs for mean time differences *Individual − Multi-user*, by TECH and by TECH × COMPLEXITY.

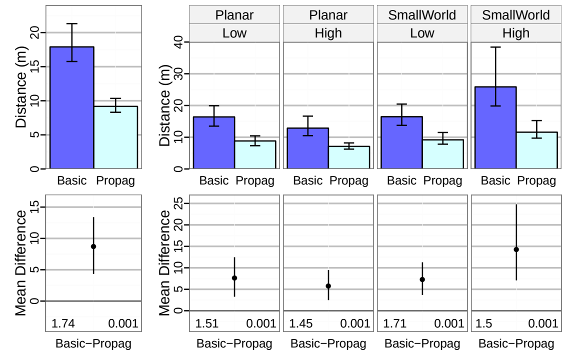


Figure 7. (Top) Average distance traveled by participant per TECH in the individual user case, aggregated on the left, and for each GRAPH × COMPLEXITY conditions on the right. (Bottom) Difference CIs for the analysis.

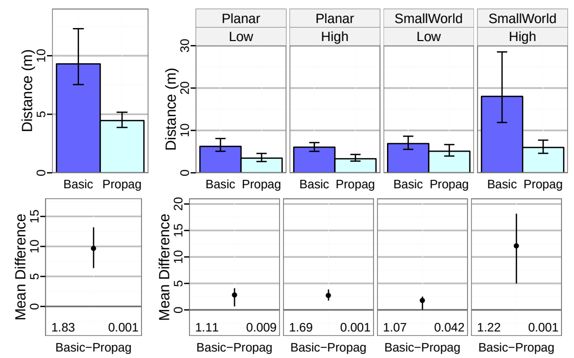


Figure 8. (Top) Average distance traveled by each participant per TECH in the multi-user case, aggregated on the left, and for each GRAPH × COMPLEXITY conditions on the right. (Bottom) Difference CI for the analysis.

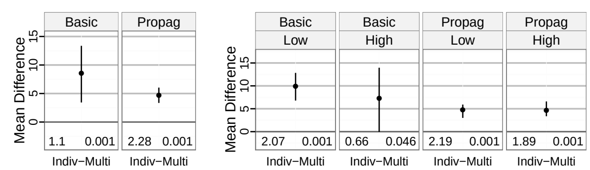


Figure 9. 95% CIs for mean differences of the traveled distance *Individual − Multi-user*, by TECH and TECH × COMPLEXITY.

4.2.2 Error Rate

Individual: We observed no measurable difference in *ErrorRate* between *Propagation* (9%) and *Basic* (13%), even if mean error rate was lower for *Propagation*. Table 2 shows the error rate for the different conditions. We can observe that almost all errors (95%) occurred with *SmallWorld* graphs irrespective of TECH.

Multi-user: On the contrary, we measured a difference in *ErrorRate* between *Propagation* (3.1%) and *Basic* (16.7%) in the collaborative case (p's $< .01$). We observed that *Propagation* led to less errors in all conditions (p's $< .05$), except in the *SmallWorld-Low*. Table 2 gives a break down for the different conditions.

	aggregated		*Planar*				*SmallWorld*			
			Low		*High*		*Low*		*High*	
	Basic	*Prop*	*Basic*	*Prop*	*Basic*	*Prop*	*Basic*	*Prop*	*Basic*	*Prop*
Indiv.	13.5%	9.4%	0%	4.2%	0%	0%	16.7%	8.3%	37.5%	25.0%
Collab.	16.7%	3.1%	8.3%	0%	12.5%	0%	4.2%	4.2%	41.7%	8.3%

Table 2. Error rate per TECH, aggregated and by GRAPH × COMPLEXITY conditions, in the individual user case and in the multi-user case.

Individual vs. Multi-user: Overall, the effect of *ErrorRate* was different for each technique across the individual and multi-user case. For *Propagation* there are marginally less errors when working in pairs (3.1%) compared to individuals (9.4%) ($p = 0.066$), with a very marked drop in error rate in the hardest graph *SmallWorld-High*, where pairs had an error rate of 8% compared to the 25% error rate for individuals.

We do not have such an effect for *Basic*, where error rate increased when pairs worked together (16.7%) compared to individuals (13.5%). When looking at different conditions, the trend was measurable for the *Planar* graphs ($p = 0.023$), but mean error rates were indeed higher for all conditions apart from *SmallWorld-Low*. We come back to this result in our discussion section.

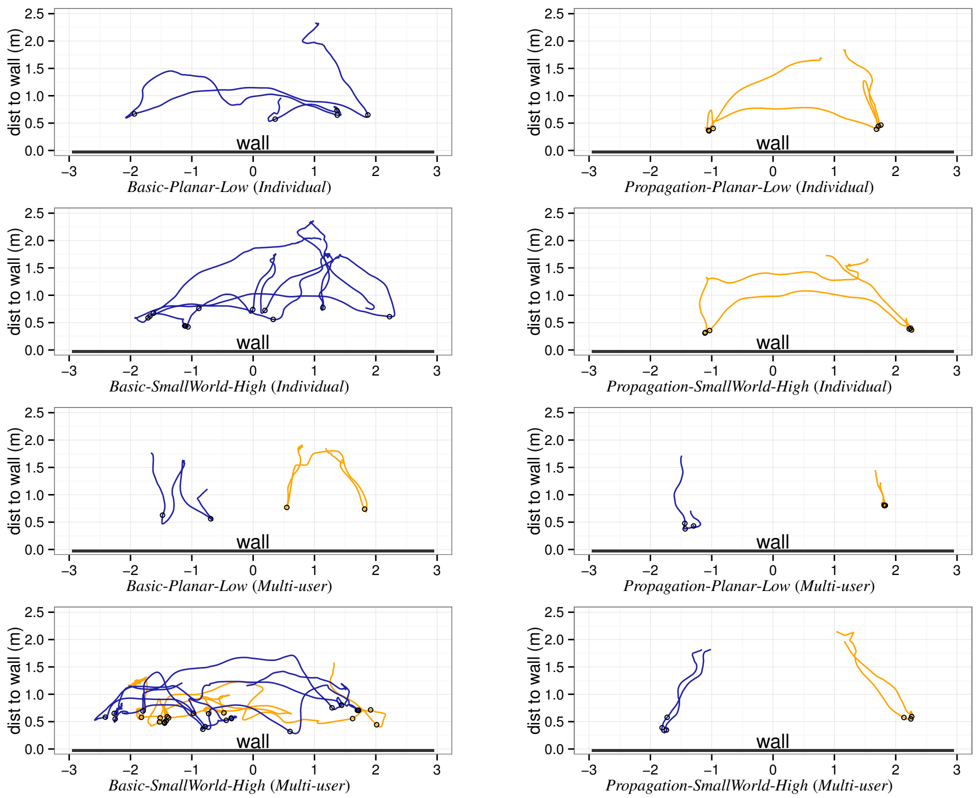


Figure 10. Bird's eye views of the movement of participants in trials for individual (2 top lines) and pairs (2 bottom lines), under the condition *Planar-Low* (easiest) and *SmallWorld-High* (hardest). *Basic* is seen in the left column, and *Propagation* in the right column. The wall is at the bottom of each graph, the unit is the meter, and the black little circles (o) indicate a touch interaction.

4.2.3 Distance Traveled

Individual: The amount of movement in the individual case was higher for *Basic* (17.9m) than for *Propagation* (9.2m), almost twice as much (three times in complex graphs), and the effect exists for all GRAPH × COMPLEXITY conditions (Figure 7).

Multi-user: Similarly, the distance covered by each participant when working in pairs was less with *Propagation* (4.6m) than *Basic* (9.3m), in all conditions (Figure 8).

Individual vs. Multi-user: As expected the distance traveled by participants in individual sessions is about twice that traveled by each participant in *Multi-user* sessions for both techniques (Figure 9-left). However, as shown in Figure 9-right, this effect is strong for *Propagation* for both *Low* and *High* complexity graphs, but only for *Low* complexity ones for *Basic*. This reinforces that the gain of working in pairs is less with *Basic* in complex graphs.

Figure 10 illustrates these results with examples of participant trajectories in front of the wall. Pairs tend to divide their work spatially, with the exception of using *Basic* in *SmallWorld-High*. Nevertheless, video recording indicates that even here participants start the task by dividing the space, but as they cannot reach a solution, they start moving more around the space to verify their work, stepping back likely to get an overview. Thus, these patterns are not just due to the need to reach nodes to interact with, but also due to the nature of collaboration using *Basic* in complex graphs.

4.2.4 Observed Strategies

Individual: Instead of propagating from a single node, all individuals using *Propagation* selected one node, propagated typically one time (sometimes two), and then moved to the second to propagate, alternating between the two until they saw an intersection (two-color node). This strategy reduced the number of selected nodes and visual clutter (less propagation steps), helping them identify the shortest path as intersection points are inside it.

The strategy used for *Basic* was different. Participants consistently selected a subset of neighbors that seem to be between the two nodes, trying to reconstruct short paths moving from one node to the other. This was successful for the smaller and less complex graphs, but did not work well for the hardest condition *SmallWorld-High*, where participants had to consider a large number of nodes, as seen by the high error rate in this condition.

Multi-user: When performing the task in pairs, participants were again consistent in their strategy. With *Propagation* it was similar to the individual sessions, but now each participant took charge of one of the two nodes, and propagated alternatively (but not concurrently) until they found intersecting nodes. They coordinated this asynchronous double-propagation using verbal communication. Then, both participants reconstructed together a shorter path candidate, each taking responsibility of their own end of the propagation. In more complex graphs, they occasionally checked each other's work (6 groups).

For *Basic*, participants again took charge of one node each, and tried to define paths using selections towards their partner, until they reached each other's work area. They worked more or less independently, and in parallel, until they started finding intersection nodes. After that, for the more complex *SmallWorld* graphs, they tried checking together candidate paths before making their choice (e.g. Figure 10 bottom-left graph, notice movement overlap). But, pairs did not double check each other's work in the easier *Planar* graphs (e.g. in Figure 10, 3rd row on the left, we see no movement overlap), which may explain the increased error rate. There was one notable strategy exception, one group decided to propagate systematically, simulating on their own the *Propagation* technique (which they had seen first).

4.2.5 Subjective Comments

Individual Users: Answers to the quantitative questions of the questionnaire (physical demand, visual disturbance, enjoyment) was very similar between the two TECH (p's = 1). This is not very surprising given that we used a between-subject design for TECH.

Multi-Users: After the collaborative session participants were able to directly compare the two techniques. All 16 preferred *Propagation*. On a 7-point Likert scale participants found that *Propagation* was less physically demanding (Avg=2.5, SD=1.2) than *Basic* (Avg=4, SD=1.6) since they were required to walk less (p's < 0.05). They also found *Propagation* more enjoyable (Avg=5.3, SD=1.1) than *Basic* (Avg=4.1, SD=1.2).

Surprisingly, they also found *Propagation* to be *less* visually disturbing (Avg=2.9, SD=1.6) than *Basic* (Avg=4.8, SD=1.8) (p's < 0.05), contrary to our hypothesis. When asked to explain why they found *Propagation* less visually disturbing, they explained that *Propagation* helped highlight paths of interest *"helps to see how many possible shortest paths there are, which is very convenient"*. Although four mentioned explicitly in their comments the existence of visual disturbance in *Propagation*, they commented that the visual footprint was desirable for tracking their work *"it gets visually disturbing very quickly after a few propagations, but it is good to be able to see the changes when we can go back and forth with the propagation easily."*.

When asked if they preferred conducting the task individually or collaborating with a partner, participants had mixed opinions. Six out of the eight that run the individual session with *Propagation* preferred to run the experiment in pairs with *Propagation*, instead of alone. As one explained *"having a partner is easier because there's someone to help check whether the answer is correct or not and I don't have to move around. However I'm not sure if doing it together is faster because sometimes communicating takes time"*. Five out of eight participants that run the individual session with *Basic* preferred to do the task in pairs with *Basic*. But, as one participant explained *"it happens that the other was exploring different solutions than me* [parallel work]*, so he was disturbing me"*. Thus, overall the multi-user context was been only slightly preferred than the individual context.

4.2.6 Discussion

Propagation was faster than *Basic* selection when identifying shortest paths, particularly in the more complex small-world graphs (confirming *H1* on time). This can be explained by participants moving more with *Basic*, twice as much overall and three times for complex graphs (confirming *H3* on movement). This is backed by subjective comments reporting less fatigue and higher preference for *Propagation*.

When moving from individuals to pairs, the mean time of both *Propagation* and *Basic* was faster, although this difference was not measurable overall. But there is a clear speed-up for complex graphs with *Propagation*, and for easy graphs with *Basic* (partially confirming *H2* on time). These differences are likely due to participant strategies. Individuals were fast with *Propagation* to begin with, and since pairs spent time coordinating and taking turns propagating, speedup due to collaboration is not visible. But as we move to more complex tasks, the cost of coordination drops compared to that of the task. On the other hand, individuals were slow with *Basic*, and as pairs worked in parallel first and combined their results later, this accelerated the work with simple graphs. But in more complex graphs this strategy was not effective, and collaboration did not compensate for the weakness of *Basic* when dealing with complex graphs.

Collaboration had an effect on accuracy. It increased when passing from individuals to pairs in *Propagation* (partially confirming *H2* on accuracy), particularly in the most complex graphs. Participants chose to closely coordinate their actions taking turns to avoid visual interference (supporting *H3* on visual disturbance). Thus it is possible they had increased workspace awareness [14], a fact supported by the ease with which they double checked each other's work. The colored propagation queries provided a filter to the interesting areas of the graph, that also helped participants focus more effectively on both their partner's and their own work, leading to the unexpected subjective feeling that propagation was less visually disturbing (subjective feel contrary to *H3* on visual disturbance). Surprisingly, accuracy decreased for *Basic* when moving to the collaborative setting. This can be explained by the adopted strategy of conducting part of the task independently, thus lacking a "big picture", that participants were forced to adopt in the individual case. This big picture is crucial for tasks such as shortest path identification, where dividing the task into spatial subtasks is not straightforward[9].

4.2.7 Summary

The two techniques, *Propagation* and *Basic*, support collaboration and wall display interaction differently:

- *Propagation* is promising for individual work for the shortest path finding task, requiring little physical movement. In group work it leads to increased accuracy, but no measurable increase in speed as there is an overhead related to coordination due to its visual footprint. Thus tight coordination, combined with the technique's highlighting of areas of interest, helped maintain an understanding of partners' work and increased accuracy.
- The *Basic* technique is as accurate when dealing with simple graphs for individuals, but considerably slower. And its performance degrades with more complex graphs. More importantly, when pairs divide tasks spatially, it can lead to loss of awareness of partners' work, resulting in loss of accuracy in collaborative work (compared to individual) when task division is not straightforward.

5 EXPERIMENT 2: OBSERVATIONAL STUDY

In the previous study we focused on a single controlled task that is not clearly divisible and parallelizable in its nature. Although pairs naturally took responsibility of one node, an overview of a larger area of the graph is required to correctly address the task. This is true for most low level graph analysis tasks suggested in the literature [37]. Nevertheless, studying them gives us insight as to how users can appropriate existing techniques in a collaborative manner. For example, *Propagation*, which quickly affected a large part of the graph, required explicit coordination. We examine, now, if this is true for other low level tasks.

[9] For example, when choosing among shortest path candidates, considering only the left half of paths is not enough to identify good candidates.

More specifically, we are interested in assessing *Propagation*, that proved more promising, as a general graph exploration technique, observing if pairs can ”discover” on their own how to perform new tasks without task specific training. And in whether they adopt similar coordination strategies as in Exp 1. Thus we are less interested in recording time, and more in observing if and how pairs collaborated.

5.1 Experimental Design

5.1.1 Participants & Apparatus

We recruited 8 volunteers (4 females, 4 males) in pairs, aged 23 to 39, with normal or corrected-to-normal vision. Pairs knew each other and had taken part in Exp 1. Sessions lasted 30min, using the same apparatus as in the first experiment.

5.1.2 Tasks

Groups performed the following topology tasks [37]:

T1 Find the shortest *distance* between two nodes (as opposed to the shortest path as in Exp 1).
T2 Find the common neighbors of degree 2 between two nodes.
T3 Find all connected components.
T4 Find an articulation point between connected components.
T5 Open exploration, reporting interesting observations.

5.1.3 Graph Types

In *T1* and *T2* we used high complexity small-world graphs similar to Exp 1. In *T1* the shortest distance was 6 and the two target nodes were separated by a physical distance of about 75% of the wall width. In *T2* the two target nodes were closer (about 50% of the wall width) and had 5 common neighbors.

In *T3* and *T4*, we combined unconnected small-world graphs (20 nodes each) of high complexity: three in *T3* (60 nodes in total) and two fin *T4* (40 nodes). To complicate the tasks, we tweaked the layout to get overlap between subgraphs. And in *T4* we hid the articulation point connecting the subgraphs inside one of them.

The graph used in the open task *T5* (similar to Figure 1) consisted of three subgraphs of different densities, and two unconnected nodes. Two subgraphs where connected through an articulation point, hidden within the third subgraph. These were the insights we wanted our participants to identify. The layout was tweaked so that subgraphs were not easy to separate visually.

5.1.4 Procedure

Participants were first reminded of the propagation technique, but no task specific training was given. Then the experimenter introduced the task without giving instructions on how to solve it, and participants performed the five tasks in order. Participants indicated they were done verbally, in a way similar to Experiment 1. If participants succeeded on their first trial within a timeout limit of 3000sec (5min), they moved on to the next task. If they failed, a strategy to solve the task was explained to them, and they were presented with another trial for that task. If they failed again, they were given a final trial, and then moved to the next task.

The experiment was recorded, and one experimenter took notes. A second experimenter gave instructions and logged the time (as in Exp 1). At the end, we asked participants if they had any suggestions for improving the technique, their thoughts on collaboration, and how confident they were in their answers.

Tasks	Discovered	Avg.Time (SD)	Correct
shortest distance	✓ (4/4)	63.5s (SD=21.9)	✓ (4/4)
2nd degree neighbors	✓ (4/4)	77.6s (SD=90.3)	✓ (4/4)
connected components	✓ (4/4)	47.6s (SD=22.4)	✓ (4/4)
articulation point	X (0/4)	timeout (3000s)	2nd try (3/4) 3rd try (1/4)

Table 3. Summary of findings for specific Tasks T1-4, indicating whether our pairs were able to discover how to perform a task, and the time it took them to do so (mean and SD). If they did not discover a strategy on their own within the timeout period, column Correct indicates on what try they succeeded.

5.2 Results

We report next participants' success in discovering a correct strategy and time averages logged during the experiment, as well as the strategies they adopted based on video log analysis and notes taken in the experiment.

5.2.1 Discovering

All pairs discovered without any training correct strategies for identifying the shortest distance between two points, the common neighbors of degree two, and the connected components (*T1-3*). No pair was able to develop a correct strategy for finding an articulation point (*T4*), but three pairs understood how to identify possible candidates. After instruction, three pairs were able to perform a new *T4* trial, and one pair on their third attempt.

All pairs conducted *T1-T3* within the time limit, with connected component completed faster 47.6s (SD=22.4), followed by shortest distance 63.5s (SD=21.9) and 2nd degree neighbors 77.6s (SD=90.3). The larger mean time and standard deviation of 2nd degree neighbors is due to one pair that did an extensive verification of their answer (described next in strategies). We note that the times reported here include both the discussion of strategy and the actual interaction to find the solution. Table 3 summarizes the discoverability of strategies and the time taken by our pairs.

In the open task, three pairs found four out of five possible insights, and one pair found all insights within the time limit. All pairs found two connected subgraphs and identified an articulation point between them. They also verified that the third subgraph was disconnected, and identified the extra disconnected nodes. One pair noticed the differences in the density of the subgraphs by calculating shortest paths.

5.2.2 Observed Strategies

We describe next the strategies adopted by participants, focusing on how they coordinated, and report their subjective comments.

Shortest Distance: In all pairs, each participant propagated from one of the two target nodes, until one or more nodes were selected by both their colors. They took turns propagating and observed each other's work so as not to loose count of the total propagation steps performed. One pair also used the thickness of edges to confirm that bi-selected nodes were at a distance of 3 from each target node.

Common Neighbors of degree 2: All pairs propagated two levels from both target nodes and then counted the number of nodes selected in both colors. Two pairs worked independently first (propagated in parallel) and checked later the bi-colored nodes together. Of these pairs, one backtracked their propagation to verify all bi-colored nodes were neighbors of degree two exactly, rather than neighbors of degree two or less for one of the nodes. The other two took turns propagating and looking at their partner's work, ensuring they considered neighbors of exactly degree two.

Connected Components: All pairs discovered that the best strategy was to start propagating from nodes that seem distant, and if one propagation no longer had an effect (no more nodes added) they had identified and fully selected a connected component. Two pairs worked in parallel, propagating in different areas simultaneously. While the other two took turns propagating and observing. One such pair had a discussion at the end of the task, noting they could have interacted in parallel to be more efficient.

Articulation Point: This task was more complex, even if the concept of articulation was easy to understand by all participants. No pair managed to find a correct strategy on their own. Nevertheless, three identified several possible candidates using propagation (including the actual one), although they were unsure how to proceed with proving it. The strategy of all pairs consisted of propagating from nodes in different areas in the graph and consider bi-colored nodes. But they did not verify that all following propagation steps between subgraphs passed through their candidates. After this strategy was explained, three pairs succeeded in their next try, while the last pair ran out of time and succeeded in its third attempt.

Open Exploration Task: Being inspired by the previous tasks, all pairs began by propagating from far away nodes and found the subgraphs connected by an articulation point, and the third disconnected subgraph. Pairs mixed their strategies, propagating in parallel at the very beginning of the exploration, and then coming together to discuss hypothesis and taking turns propagating and observing.

5.2.3 Subjective comments

All participants felt confident in their answers and strategies, especially for the first three tasks. Six commented that collaboration increased their confidence in their solutions. When prompted about their coordination strategy, four explained that taking turns helped them be more aware of each other's work, while two mentioned that sometimes they still lacked awareness of each other's work when working at distant locations. Three participants also commented on the visual footprint of propagation: occasionally the colored query of their partner would enter their work area, causing some visual disturbance, while rarely they also missed the effects of their own propagation when it was far away from their location. Nevertheless, these participants also mentioned that these colors helped them verify their partner's work.

They all felt the articulation point task was difficult, and three users independently suggested extending the propagation selection to better support this task, for example by being able to "block" a node and prevent propagation from going through it, or removing nodes temporarily. Four participants commented that it was sometimes hard to tell how many propagation steps they had performed, and suggested adding it as a small number close to the propagation origin. These last two features were implemented. Two participants requested the possibility to collapse and bookmark propagation queries for later use, and another two suggested the option to propagate using a different color within an existing propagation.

5.2.4 Summary

Participants were able to devise correct strategies for the majority of tested tasks, and in the articulation point task identify good candidates, demonstrating that the extended *Propagation* is an interesting general purpose technique for graph exploration. As in Exp1, participants divided the space and mostly took turns propagating (with few exceptions). We got several comments indicating that the reason for this turn taking was to coordinate and keep awareness of others' work, but also to avoid visual disruption due to the global footprint of the technique. Nevertheless, this global footprint also helped them check each other's work quickly.

6 Discussion and Design Implications

We examined how pairs and individuals work on wall-displays to solve low-level graph topology tasks. Our findings indicate that:

Exploring complex graphs individually requires interaction that highlights the structure of the graph, while basic interaction is enough for simple graphs. Wall-displays can comfortably display large graphs, nevertheless it is still challenging for individuals to explore complex graphs such as large small-world ones. Here we observed a significant benefit in using advanced interaction techniques, such as *Propagation* selection. For individuals, *Basic* selection did not scale well for complex graphs, nevertheless it performed reasonably well for simpler planar graphs.

Collaboration improves accuracy only *if techniques allow verification of partners' work.* Pairs were more confident in their responses than individuals with both techniques. Nevertheless, their actual accuracy improved only for *Propagation*. On the contrary, pairs using *Basic* were more error prone than individuals. Our observations and participants' comments indicate that this is because with basic selection it is difficult to acquire an overview of all choices considered by one's partner, and thus maintain a global view of the work and identify possible errors. On the contrary, with propagation selection it was easier to verify at a glance the work of one's partner and check for errors. In collaborative graph exploration, lack of workspace awareness [14, 19] can decrease accuracy, compared to individual work.

Even when tasks are not clearly divisible, pairs divide the wall spatially. For many topology tasks identified in the literature, and used in our experiments, there is no clear strategy to divide them in space, as they may require a global understanding of subgraphs that extend across the display. Nevertheless, irrespective of task and technique, pairs divided the wall spatially. Even when not optimal, they each took responsibility of one part of the wall and then combined their work, with mixed results. This division was observed in tasks that are clearly spatially divisible [32, 39, 62], but not in tasks that are not clearly spatially divisible, such as route planning tasks [59]. Designers should anticipate this division of space and encourage tighter collaboration (discussed next) when tasks are not spatially divisible.

If a technique has a global footprint, tight coordination is adopted. Although pairs occasionally worked in parallel with *Propagation*, they mostly took turns, working on different sections of the wall. They commented that this tight coordination was needed because the technique had a visual footprint that could reach all areas of the wall, risking disturbing the partner's work. Theoretical work on automated graph exploration using a variation of propagation [12] has shown that automated agents with full knowledge of others' exploration (i.e. high awareness) tend to explore the graph fully more quickly. Given our findings on propagation accuracy and the theoretical result on efficiency, designers could use techniques with large visual footprints to *encourage* close collaboration that can increase accuracy and efficiency. This is complementary to findings that when collaborating loosely, participants *chose* techniques with local visual footprints [59].

Consider awareness vs. disruption tradeoff in techniques. Participants' comments indicate there is a clear tradeoff between awareness and visual disruption. *Propagation* can be visually disrupting and affect the partner's work, but it also provides higher degree of workspace awareness [14, 19]. While *Basic* has a small visual footprint and is less disturbing, but pairs can loose track of their partner's work due to the wall size and graph complexity. Both types of techniques should be supported, and collaborators should be able to transition between them depending on how tight their work coupling is [59], and how divisible their task is.

Provide techniques that do not require extensive walking. Free walking is beneficial in wall displays [6, 10]. Nevertheless, techniques that require users to repeatedly walk to interact with different areas of visualizations (such as *Basic*) are fatiguing. Designers should provide interaction alternatives that can be activated locally but act globally, such as *Propagation* or ones proposed in the HCI literature for remote reaching [9, 56] and data manipulation [39]. Alternatively, designers could provide a combination of touch and distant interaction (e.g. using mobile devices) to ensure users can perform large scale or remote interactions across distances.

7 Conclusion and Future Work

In this work, we study two selection techniques for graph exploration on wall-displays, used by individuals and multiple users. We adapted two existing techniques for use by multiple users on a touch enabled display, a basic selection, and a propagation selection using the idea of transmission. We performed a user study that showed *Propagation* to be faster in both individual and multi-user contexts, to be more accurate for multiple users, and to require less movement than *Basic* in a shortest path identification task. It is also versatile enough to be used in a series of topology tasks, observed in a second study.

Nevertheless, as *Propagation* has a large visual footprint, it requires higher coordination when used by multiple users. When working in pairs, propagation selection increases accuracy overall, but due to a coordination cost it improves time only for complex graphs. When using basic selection, that has a small visual footprint, accuracy dropped for pairs, most noticeably in complex graphs. Indeed, we observed that using basic selections, participants tended to work independently and lose awareness of each other's work, which proved detrimental for the task we consider, that is not clearly divisible. We conclude with design implications, stressing the tradeoffs of techniques with global vs. local visual footprints, and the need to allow users to switch between such techniques depending on whether the task is spatially divisible, and on the nature of collaboration (loose or tight).

A possible future direction includes improving the propagation technique. As other multicolor query selections, it prevents the use of color for encoding other information on the graph. We plan to explore other visual encodings, such as motion [65], that nevertheless need to be considered carefully when applied to techniques that feature a large visual footprint in multi-user settings. More generally, we plan to investigate design variations for propagation that reduce this global footprint, for example re-layout the graph to move selected nodes closer together. Nevertheless, as we are dealing with multi-user settings, care must be taken to limit colleague disturbance. Finally, we plan to explore visualization techniques to better emphasize grouping of nodes belonging to one [35] or multiple selection groups.

As this is the first work to examine how multiple users that move freely to explore graphs on walls displays, we focused on fairly controlled topology tasks. We next plan to investigate more open ended exploration tasks, where we suspect task division will be different. Moreover, we plan to explore different types of input, for example combinations of touch and distant pointing, to better support user mobility.

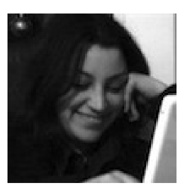
Although the idea of multiple propagations was used in the context of a collaborative vertical surface, we believe it has potential in horizontal tabletops, but also in desktop settings, and should be further studied. It can also be adapted to serve other graph representations such as directed graphs and matrices, and for dynamic graphs [5, 2].

References

[1] J. Abello, E. Koutsofios, E. Gansner, and S. North. Large networks present visualization challenge. *SIGGRAPH Comput. Graph.*, 33(3):13–15, 1999.
[2] J. Ahn, C. Plaisant, and B. Shneiderman. A task taxonomy for network evolution analysis. *IEEE TVCG*, 20(3):365–376, 2014.
[3] APA. *Publication manual of the American Psychological Association (6th edition).* APA, Washigton DC, USA, 2009.
[4] D. Archambault, T. Munzner, and D. Auber. TugGraph: Path-preserving hierarchies for browsing proximity and paths in graphs. In *PacificVis '09*, pp. 113–121. IEEE, 2009.
[5] B. Bach, E. Pietriga, and J.-D. Fekete. GraphDiaries: Animated transitions and temporal navigation for dynamic networks. *IEEE TVCG*, 20(5):740–754, 2014.
[6] R. Ball and C. North. The effects of peripheral vision and physical navigation on large scale visualization. In *GI '08*, pp. 9–16. CIPS, 2008.
[7] R. Ball, C. North, and D. A. Bowman. Move to improve: Promoting physical navigation to increase user performance with large displays. In *CHI '07*, pp. 191–200. ACM, 2007.
[8] M. Beaudouin-Lafon et al. Multi-surface interaction in the WILD room. *IEEE Computer*, 45(4):48–56, 2012.
[9] A. Bezerianos and R. Balakrishnan. The Vacuum: Facilitating the manipulation of distant objects. In *CHI '05*, pp. 361–370. ACM, 2005.
[10] A. Bezerianos and P. Isenberg. Perception of visual variables on tiled wall-sized displays for information visualization applications. *IEEE TVCG*, 18(12):2516–2525, 2012.
[11] O. Chapuis, A. Bezerianos, and S. Frantzeskakis. Smarties: An input system for wall display development. In *CHI '14*, pp. 2763–2772. ACM, 2014.
[12] D. Dereniowski, Y. Disser, A. Kosowski, D. Pajak, and P. Uznanski. Fast Collaborative Graph Exploration. In *ICALP '13*, pp. 520–532. Springer, 2013.
[13] M. Dork, N. Henry Riche, G. Ramos, and S. Dumais. PivotPaths: Strolling through faceted information spaces. *IEEE TVCG*, 18(12):2709–2718, 2012.
[14] P. Dourish and V. Bellotti. Awareness and coordination in shared workspaces. In *CSCW '92*, pp. 107–114. ACM, 1992.
[15] T. Dwyer, N. H. Riche, K. Marriott, and C. Mears. Edge compression techniques for visualization of dense directed graphs. *IEEE TVCG*, 19(12):2596–2605, 2013.
[16] P. Eades. A heuristic for graph drawing. *Congressus Numerantium*, 42:149–160, 1984.
[17] A. Endert, C. Andrews, Y. Lee, and C. North. Visual encodings that support physical navigation on large displays. In *GI '11*, pp. 103–110. CHCCS, 2011.
[18] T. M. J. Fruchterman and E. M. Reingold. Graph drawing by force-directed placement. *Software: Practice and Experience*, 21(11):1129–1164, 1991.
[19] C. Gutwin and S. Greenberg. A descriptive framework of workspace awareness for real-time groupware. *CSCW*, 11(3):411–446, 2002.
[20] J. Heer, M. Agrawala, and W. Willett. Generalized selection via interactive query relaxation. In *CHI '08*, pp. 959–968. ACM, 2008.
[21] J. Heer and D. Boyd. Vizster: Visualizing online social networks. In *InfoVis '05*, pp. 32–39. IEEE, 2005.
[22] J. Heer, F. Ham, S. Carpendale, C. Weaver, and P. Isenberg. Creation and collaboration: Engaging new audiences for information visualization. In *Information Visualization*, pp. 92–133. Springer-Verlag, 2008.
[23] N. Henry Riche, T. Dwyer, B. Lee, and S. Carpendale. Exploring the design space of interactive link curvature in network diagrams. In *AVI '12*, pp. 506–513. ACM, 2012.

[24] I. Herman, G. Melancon, and S. Marshall. Graph visualization and navigation in information visualization: A survey. *IEEE TVCG*, 6(1):24–43, 2000.
[25] J. D. Hincapié-Ramos, X. Guo, P. Moghadasian, and P. Irani. Consumed endurance: A metric to quantify arm fatigue of mid-air interactions. In *CHI '14*, pp. 1063–1072. ACM, 2014.
[26] P. Isenberg and S. Carpendale. Interactive tree comparison for co-located collaborative information visualization. *IEEE TVCG*, 13(6):1232–1239, 2007.
[27] P. Isenberg, S. Carpendale, A. Bezerianos, N. Henry, and J.-D. Fekete. CoCoNutTrix: Collaborative retrofitting for information visualization. *IEEE CGA*, 29(5):44–57, 2009.
[28] P. Isenberg, P. Dragicevic, W. Willett, A. Bezerianos, and J.-D. Fekete. Hybrid-image visualization for large viewing environments. *IEEE TVCG*, 19(12):2346–2355, 2013.
[29] P. Isenberg, N. Elmqvist, J. Scholtz, D. Cernea, K.-L. Ma, and H. Hagen. Collaborative visualization: Definition, challenges, and research agenda. *Information Visualization*, 10(4):310–326, 2011.
[30] P. Isenberg and D. Fisher. Collaborative brushing and linking for co-located visual analytics of document collections. In *EuroVis'09*, pp. 1031–1038. EA, 2009.
[31] M. R. Jakobsen and K. Hornbæk. Interactive visualizations on large and small displays: The interrelation of display size, information space, and scale. *IEEE TVCG*, 19(12):2336–2345, 2013.
[32] M. R. Jakobsen and K. Hornbæk. Up close and personal: Collaborative work on a high-resolution multitouch wall display. *ACM ToCHI*, 21(2):11:1–11:34, 2014.
[33] M. R. Jakobsen and K. Hornbæk. Is moving improving?: Some effects of locomotion in wall-display interaction. In *CHI '15*, pp. 4169–4178. ACM, 2015.
[34] Y. Jansen, P. Dragicevic, and J.-D. Fekete. Tangible remote controllers for wall-size displays. In *CHI '12*, pp. 2865–2874. ACM, 2012.
[35] R. Jianu, A. Rusu, Y. Hu, and D. Taggart. How to display group information on node-link diagrams: An evaluation. *IEEE TVCG*, 20(11):1530–1541, 2014.
[36] U. Kister, P. Reipschläger, F. Matulic, and R. Dachselt. BodyLenses: Embodied magic lenses and personal territories for wall displays. In *ITS '15*, pp. 117–126. ACM, 2015.
[37] B. Lee, C. Plaisant, C. S. Parr, J.-D. Fekete, and N. Henry. Task taxonomy for graph visualization. In *BELIV '06*, pp. 1–5. ACM, 2006.
[38] A. Lehmann, H. Schumann, O. Staadt, and C. Tominski. Physical navigation to support graph exploration on a large high-resolution display. In *ISVC '11*, pp. 496–507. Springer-Verlag, 2011.
[39] C. Liu, O. Chapuis, M. Beaudouin-Lafon, and E. Lecolinet. Shared interaction on a wall-sized display in a data manipulation task. In *CHI '16*, pp. 2075–2086. ACM, 2016.
[40] C. Liu, O. Chapuis, M. Beaudouin-Lafon, E. Lecolinet, and W. Mackay. Effects of display size and navigation type on a classification task. In *CHI '14*, pp. 4147–4156. ACM, 2014.
[41] M. Marner, R. Smith, B. Thomas, K. Klein, P. Eades, and S.-H. Hong. Gion: Interactively untangling large graphs on wall-sized displays. In *GD '14*, pp. 113–124. Springer-Verlag, 2014.
[42] M. McGuffin and I. Jurisica. Interaction techniques for selecting and manipulating subgraphs in network visualization. *IEEE TVCG*, 15(6):937–944, 2009.
[43] K. Mehadhebi. A methodology for the design of a route network. In *3rd Air Traffic Management R&D Seminar*. Eurocontrol & FAA, 2000.
[44] B. Meyer. Self-organizing graphs a neural network perspective of graph layout. In *GD '98*, pp. 246–262. Springer-Verlag, 1998.
[45] S. Milgram. The small world problem. *Psychol. Today*, 1(1):61–67, 1967.
[46] T. Moscovich, F. Chevalier, N. Henry, E. Pietriga, and J.-D. Fekete. Topology-aware navigation in large networks. In *CHI '09*, pp. 2319–2328. ACM, 2009.
[47] C. Mueller, D. Gregor, and A. Lumsdaine. Distributed force-directed graph layout and visualization. In *EGPGV '06*, pp. 83–90. EA, 2006.
[48] J. O'Madadhain, D. Fisher, P. Smyth, S. White, and Y.-B. Boey. Analysis and visualization of network data using JUNG. *J. Stat. Softw.*, 10(2):1–35, 2005.
[49] E. Pietriga, S. Huot, M. Nancel, and R. Primet. Rapid development of user interfaces on cluster-driven wall displays with jBricks. In *EICS '11*, pp. 185–190. ACM, 2011.
[50] H. Purchase. Metrics for graph drawing aesthetics. *J. Visual Lang. Comput.*, 13:501–516, 2002.
[51] F. Rajabiyazdi, J. Walny, C. Mah, J. Brosz, and S. Carpendale. Understanding researchers' use of a large, high-resolution display across disciplines. In *ITS '15*, pp. 107–116. ACM, 2015.
[52] K. Reda, A. E. Johnson, M. E. Papka, and J. Leigh. Effects of display size and resolution on user behavior and insight acquisition in visual exploration. In *CHI '15*, pp. 2759–2768. ACM, 2015.
[53] E. M. Rogers. *Diffusion of Innovations*. Free Press, 5 edition, 2003.
[54] C. Rooney, A. Endert, J.-D. Fekete, K. Hornbæk, and C. North. Powerwall: int. workshop on interactive, ultra-high-resolution displays. In *CHI EA '13*, pp. 3227–3230, 2013.
[55] S. Schmidt, M. Nacenta, R. Dachselt, and S. Carpendale. A set of multi-touch graph interaction techniques. In *ITS '10*, pp. 113–116. ACM, 2010.
[56] G. Shoemaker, A. Tang, and K. S. Booth. Shadow Reaching: A new perspective on interaction for large displays. In *UIST '07*, pp. 53–56. ACM, 2007.
[57] M. C. Stone. Color and brightness appearance issues in tiled displays. *IEEE CGA*, 21(5):58–66, 2001.
[58] D. Tan, D. Gergle, P. Scupelli, and R. Pausch. With similar visual angles, larger displays improve spatial performance. In *CHI '03*, pp. 217–224. ACM, 2003.
[59] A. Tang, M. Tory, B. Po, P. Neumann, and S. Carpendale. Collaborative coupling over tabletop displays. In *CHI '06*, pp. 1181–1190. ACM, 2006.
[60] J. J. Thomas and K. A. Cook, editors. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. National Visualization and Analytics Ctr, 2005.
[61] T. Tsandilas, A. Bezerianos, and T. Jacobs. SketchSliders: Sketching widgets for visual exploration on wall displays. In *CHI '15*. ACM, 2015.
[62] E. Tse, J. Histon, S. D. Scott, and S. Greenberg. Avoiding interference: How people use spatial separation and partitioning in SDG workspaces. In *CSCW '04*, pp. 252–261. ACM, 2004.
[63] F. van Ham and A. Perer. "Search, show context, expand on demand": supporting large graph exploration with degree-of-interest. *IEEE TVCG*, 15(6):953–60, 2009.
[64] T. Von Landesberger, A. Kuiper, T. Schreck, J. Kohlhammer, J. Van Wijk, J.-D. Fekete, and D. Fellner. Visual analysis of large graphs: State-of-the-art and future research challenges. *Comput. Graph. Forum*, 30(6):1719–1749, 2011.
[65] C. Ware and R. Bobrow. Supporting visual queries on medium-sized node-link diagrams. *Information Visualization*, 4(1):49–58, 2005.
[66] C. Ware and P. Mitchell. Visualizing graphs in three dimensions. *ACM TAP*, 5(1):2:1–2:15, 2008.
[67] K.-P. Yee, D. Fisher, R. Dhamija, and M. Hearst. Animated exploration of dynamic graphs with radial layout. In *InfoVis '01*, pp. 43–50, 2001.
[68] B. Yost and C. North. The perceptual scalability of visualization. *IEEE TVCG*, 12(5):837–844, 2005.
[69] B. Zimmer and A. Kerren. Sensemaking and provenance in distributed collaborative node-link visualization. In *IEEE VIS '14 Workshop on Provenance for Sensemaking*, pp. 1–2, 2014.

Arnaud Prouzeau is a PhD student at Univ Paris-Sud, France, and a member of the HCC (LRI) and ILDA (INRIA) teams. He received an MSc in HCI from University of Toulouse Paul Sabatier, France, and an engineering degree from the Ecole Nationale de l'Aviation Civile (ENAC), France. His interests include interaction and visualization on wall-displays, collaboration and command and control systems.



Anastasia Bezerianos is an assistant professor at Univ Paris-Sud, France, and a member of the HCC (LRI) and ILDA (INRIA) teams. She received her PhD from U. of Toronto, Canada. Her interests include interaction and visualization designs for large displays, visual perception, user evaluation, and collaborative work. She has served in the program committees of IEEE VIS, IEEE PACIFICVIS, ACM CHI, ACM UIST.

Olivier Chapuis is a research scientist at CNRS working at Univ Paris-Sud, France, and is deputy-head of the HCC (LRI) and ILDA (INRIA) teams. He received a PhD in Mathematics in 1994 from University Paris Diderot. Since 2004, he works in the domain of Human-Computer Interaction. His current interests include interaction techniques in general, and, in particular, multi-user interaction for wall-displays.

Storytelling in Visual Analytics tools for Business Intelligence

Micheline Elias[1,2] , Marie-Aude Aufaure [1,3] and Anastasia Bezerianos[4]

[1] MAS Laboratory, Ecole Centrale Paris, France
[2.] SAP Research, Business Intelligence Practice, France
[3] AxIS Research Group, INRIA Paris, France [4] Univ Paris-Sud & CNRS, INRIA, France
{michellelias}@gmail.com,
{Marie-Aude.Aufaure}@ecp.fr,{anab}@lri.fr

Abstract. Stories help us communicate knowledge, share and interpret experiences. In this paper we discuss the use of storytelling in Business Intelligence (BI) analysis. We derive the actual practices in creating and sharing BI stories from in-depth interviews with expert BI analysts (both story "creators" and "readers"). These interviews revealed the need to extend current BI visual analysis applications to enable storytelling, as well as new requirements related to BI visual storytelling. Based on these requirements we designed and implemented a storytelling prototype tool that is integrated in an analysis tool used by our experts, and allows easy transition from analysis to story creation and sharing. We report experts' recommendations and reactions to the use of the prototype to create stories, as well as novices' reactions to reading these stories.
Keywords: Visual Storytelling, Business Intelligence.

1 Introduction

Stories are one of humanity's communication structures and storytelling a means of passing on wisdom and culture. Individually and collectively, stories help us make sense of our past and reason about the future. Johnson 17] and Maclntyre [21] argue that story narrative also goes beyond communication, it is also a process of extracting meaning from events, that is central to human experience and conduct.

It is thus not surprising that intelligence analysts, who make sense of data, identify links between disparate pieces of intelligence, and communicate their findings to decision makers, are interested in storytelling. Their analysis process is supported by the construction of stories and narratives, both during sense-making and during presentation of results. Bier et al. [3] point out that a story is a powerful abstraction to conceptualize threats and understand patterns as part of the analytical process, and story structures and storytelling is the means to present the analysis results. As analysts continue to work with increasingly large data sets, data visualization has become an incredibly important asset both during sensemaking analysis, and when communicating findings to other analysts, decision makers or to a broader public [15, 27].

Given the importance of storytelling in different steps of the analysis process it is clear there is a need to enhance visual analysis tools with storytelling support. Nevertheless this process is not simple [33, 20], as analysts need to work within very large data resources and highlight and explain items or events of importance and their connections to their audiences. Despite the growing use and research work on storytelling and narrative visualization in the visualization domain [26, 9, 15], related research on the domain of BI has not equally progressed. Our work attempts to rectify this.

In Business Intelligence (BI) analysis, the most popular visualization tools are dashboards. Dashboards [8] are collections of visual components (such as charts or tables) on a single view [12], that permit analysts to explore their data and quickly view different aspects of complex datasets. Nevertheless, simple collections of visual representations cannot be interpreted by untrained audiences; to become meaningful they require interpretation and explanation, often presented in a story narrative.

We attempt to answer the following research questions: What are the actual practices of BI experts in creating and communicating visual stories to their audiences, and do current BI visualization tools support well this story creation and storytelling process? How can we enhance BI visual analysis tools with narrative capabilities, and are these capabilities effective in communicating analysis stories to others?

Our work makes the following contributions:

(1) Interviews with expert BI analysts (story ”creators” and ”readers”), provide a better understanding of current practices in creating BI stories. BI stories are an asynchronous, visual, and interactive means of transmitting and sharing information on data between analysts and their audience. They include visualized data (dashboards, charts or tables) and detailed explanations on the story structure in the form of presentation(s), detailed textual annotations and external resources such as wikis. (2) Current tools fail to support the storytelling process, which remains cumbersome and requires frequent switching between software (analysis tools, screen captures, text editors, presentation tools, etc.). We emphasize the need for storytelling support, and extract requirements for enhancing BI tools with visual storytelling capabilities. (3) Following these requirements and a user-centered design approach, we implement a storytelling prototype incorporated in an existing visual analysis dashboard, to fluidly transition from analysis to story creation and reading. (4) We report on feedback from BI experts on the usefulness of the prototype as a communication and teaching medium, and from BI novices reading a story on its effectiveness for story communication.

2 Related Work

Stories are series of ordered events and facts, and their connections (Oxford English Dictionary). In intelligence analysis, it is furthermore an abstraction used by analysts to conceptualize threats and understand patterns in the analytical process [3], and communicate their findings to others [2].

Stories in Business. In recent years, organizations and their leaders have identified the importance and value of narrative and anecdotal information conveyed in the form of stories [29], to present either complex ideas, the context and details of crucial information, or personal interpretation of that information [28].

Research conducted to date has demonstrated the value of storytelling to improve organizational structure [24] and collaboration quality [23, 6], socialization and adaptation of new employees [7, 18, 19], organizational and financial success [4, 5], innovation and new product development [22], and teaching and learning [16].

The majority of this work is a meta-analysis of the effect of storytelling within an organization, rather than identifying the storytellers' needs in terms of supporting and enhancing the storytelling process as is our case. Moreover, the stories themselves discussed in this work, relate to the transmission of information and knowledge within an organization, mostly in textual or verbal form, rather than in visual form. The widespread use of *visualization* dashboards in the domain is a more recent development [8], and so is the transmission of knowledge *between organizations* (dedicated BI analysis organizations and their clients). Thus storytelling needs in the domain have evolved. In Section 3 we explain current practices in visual knowledge transmission and we take a step at characterizing current problems and needs more precisely.

Stories in sense making. Baber et al.[2] point out that contemporary theories of sense making rely on the idea of 'schema', of a structure to organize and represent factual information, as well as the knowledge, beliefs and expectations of the people who are engaged in sense making. They can thus be considered as a collection of narratives. They further discuss the formalism of stories in sense making, and how the most effective stories are organized around the actors in the stories, their actions and rationale, events, their context, and most importantly the relationships between these. As argued by Bier et al.[3], for effective collaboration and communication we need to use less text, and organize knowledge around entities (people, places, things, times etc.) rather than free form text. Similarly, Pirolli & Russell [25] propose the mapping of intelligence facts and insights into frames, that can be expressed in a variety of forms including stories, maps, organizational diagrams or scripts. It is thus clear that conducting intelligence analysis, communicating findings, and organizing knowledge in stories, has a strong visual component that represents entities and their connections.

Stories in data visualization. Often text or audio transmit the main story, while visualizations support the story or provide details. Comics and flowcharts are special narratives relaying mostly on visuals rather than text. Recently, we have seen an increase in integrating complex visualizations into narratives in many news organizations [11], journalism reports (e.g. New York Times, Washington Post, the Guardian), and television reports (e.g. Gapminder[1]). Segel et al.[26] explore different aspects of narratives from a variety of sources and identify distinct genres of narrative visualization.

In the business intelligence domain the main visualization tools are dashboards [8], collections of multiple visual components (e.g. charts, tables) on a single view [12]. BI dashboards permit users to interpret data at a glance, and are very popular (e.g. [Dundas[2], Oracle bi 10g[3], Xcelsius, Spotfire[4], Tableau[5]]). But as Wojtkowski and

[1] http://www.gapminder.org, 2010
[2] http://www.dundas.com/microsite/dashboard
[3] http://www.oracle.com/technetwork/middleware/bi-enterprise-edition/
[4] http://spotfire.tibco.com/
[5] http://www.tableausoftware.com/

Wojtkowski [33] point out, dashboards and other visualization tools used to analyze complex data cannot simply tell stories. They need to be ”tailored” to accommodate storytelling by better highlighting items of importance within very large data resources [14], in a way that is efficient for the storyteller and clear for the audience.

Some visualization systems began to integrate storytelling [20]. GeoTime [9], a geo-temporal event visualization system, integrates a story system that shows events in space and time, hypertext linked visualizations, and visual annotations to create an environment for both analytic exploration, and story creation and communication. The design of the tool was based on real user needs and has been evaluated with intelligence analysts. Wohlfart and Hauser [32] present a system for demonstrating to audiences the path followed in analyzing 3D volumetric data, while providing audience with limited interaction with the 3D data. Both these systems deal with a single visualization/chart seen over time or from different views. Thus their designs are not necessarily applicable to the multi-dimensional and multi-chart BI visualizations where the connections and links between data are not as clear.

Storytelling tools in BI are not yet as advanced. Systems like Sense.us [15] and Tableau[5] allow analysts to visualize data, conduct analysis, and store a history of the exploration, that can serve as a step towards creating a story. Many Eyes [31], Tableau Public[5] and Sense.us [15] publish interactive visualizations online, and permit collaborative analysis through comments on a single visualization, creating an evolving analysis. This collaborative annotation can be seen as a step towards a collaborative knowledge narrative, where an analysis story could be extracted from the visualizations and comments. Nevertheless, these tools do not provide explicit means to indicate story progression and highlight relationships between multiple visualizations (seen in dashboards), that are key to intelligence analysis and communication [2].

3 Interviews

To support visual storytelling in BI, we first investigate current practices in transmitting BI analysis results, and identify challenges experts face currently when creating and sharing their stories, or reading and interpreting stories of others. We interviewed 5 BI experts in a leading business intelligence development company, with experience from 6 months to 12 years. Participants' job descriptions included development project manager, project manager, development director, data warehouse engineer and delivery manager. Three experts used dashboards and communicated their analysis or read analysis from others daily, while two several times a week. Interviews were held in person or by phone and lasted over 1 hour (Table 1). We report next our major findings. Note that they hold true for all BI reporting tools available today, as indicated by our experts' experience with multiple tools, and our own investigation.

Table 1. Experts', experience, dashboard use and interview duration.

Experts	P1	P2	P3	P4	P5
Experience	7 years	6 months	6 years	3 years	3 years
Dashboard use	daily	weekly	weekly	daily	daily
Duration	90 min	120 min	60 min	60 min	60 min

3.1 Current Practices: BI Reports

All experts communicate their analysis or read analysis from others in the form of BI reports. These contain an entire dashboard, often accompanied by several single charts and tables. Their purpose is to help monitor aspects of business performance, by highlighting with charts key performance indicators (KPIs) that indicate success of project management and progression of development teams. Our experts were report creators, but occasionally acted as readers of others' reports.

A BI report can be a single page, with only a title and a dashboard (Fig. 1.). This one-page report summarizes the most important data, and can serve as a starting point for longer reports, up to 50 pages in length, that give more details. Details can be additional visualizations, tables, annotations, links to the data used in the visualizations, and finally block text, although all participants preferred limiting text to 5% of the space in a report. When possible, creators want to make their reports "live", with interactive visualizations for the audience to explore. Thus BI reports have a very strong visual component, with little text added for explanations. The one page report is preferred by clients like company managers, that want an "at-a-glance view" of a project status, and that see many such reports during a day. Longer detailed reports are used to communicate findings to other analysts, project teams and product owners.

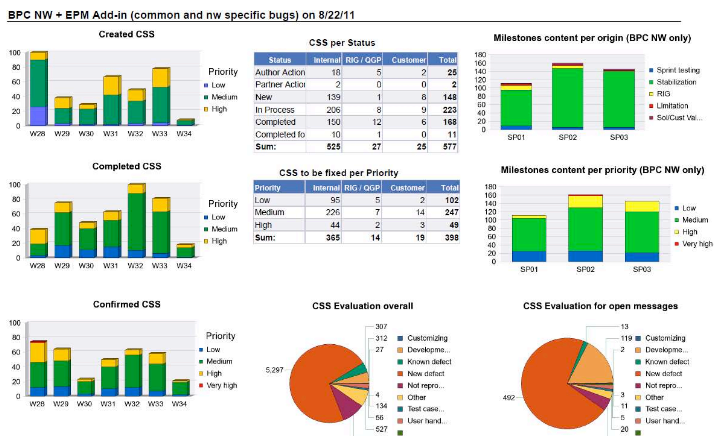


Fig. 1. One page BI report from an expert: dashboard presenting bug status in a project. Note the lack of annotations/explanations, aside from chart titles and data dimensions.

Experts mentioned that BI reports are used to (i) answer specific questions (e.g. What is the code development progress in project X the last 3 weeks?); (ii) investigate specific data (e.g. Investigate the increase of bugs in section Y of project X); (iii) manage conflicts and highlight problems (e.g. Team X completed less use cases than planned, because test team Y did not complete testing); and (iv) interpret past data and predict future trends (e.g. Given last year's sales, what is this year's projection?).

BI reports are nevertheless only one of the tools used to communicate analysis findings, and are not a complete BI story.

3.2 Current Practices: Supporting material for BI Storytelling

Experts explained that reports are difficult to understand without detailed explanations from the creator. Thus they don't represent a complete BI story on their own. Before publishing a new report, the creator provides an introduction session to report readers. During this (usually one hour long) session, she explains with a presentation the entire story and goal of the report, the meaning of each chart, the relations between different KPIs in different charts, as well as in what sequence to read it. Thus their verbal explanations and presentation slides are ways for creators to explain their analysis path. These presentations often show relational diagrams, text explanations and interpretations, highlight specific visualizations or parts of visualizations using colored highlights, arrows and other symbols. The report creator often draws by hand on a chart the ideal data values, to help compare the current situation to target goals.

The audience can ask questions during or after the training sessions, through emails or arranged meetings. Content similar to the presentation, is often put on a wiki page to answer follow-up questions. The session recording, wiki page, and presentation slides, are made available to the audience as reminders and reference material. Experts explained that this material is not a complete story either, as it does not include the visualization and data details shown on the report.

Thus a *complete BI story* is a collection of visual representations of the most important data followed by further data details (BI report), accompanied by instructions on how to read the visualizations (order, connections, importance) in the form of presentations and verbal or textual instructions on a wiki. Although creators present a desired way to view data, this structure is not enforced: the audience can pursue the story in a different sequence and dig for data details in the report. Thus a BI story differs from a simple fixed sequence presentation that prohibits exploration. Its goal is to communicate analysis findings and supporting evidence.

3.3 Current Practices: Teaching BI storytelling

According to our experts, all the material making up a BI story is also a tool for training analysts. As a recently trained creator explained (P2), this material taught her the key aspects of BI reports and how to interpret them. When analysts start out they read in detail older reports and their supporting material to understand how to analyze visualizations, see relations, and identify important points and their link to KPIs.

Experts mentioned that designing a new report is hard and requires a lot of experience, thus they often use a template that they modify according to their needs. Three (P2,P3,P4) stressed the influence of a senior creator (P1), and their reliance on her BI report templates to create their own. They still occasionally contact the senior creator when facing difficulties or are unsure of the clarity of their message in a report.

The supporting material is also often reused. All experts explained that data (and thus reports) change, but often the structure of BI stories repeats itself. Moreover, they sometimes create a sequence of reports on dynamically evolving data (e.g. sales over several years). These stories change very little. Thus material on how to read reports can be re-used by adapting it to new stories, nevertheless, experts pointed out it can be a tedious and repetitive process.

3.4 Current Practices: Collaboration in BI Stories

Experts explained that report creation follows several iterations of communication with a client (usually a decision maker like a manager or executive), who is one of many possible report readers. Thus the focus and structure of the story changes and evolves through the communication between the story teller and the story reader over time, and stories become the collaboration product of story tellers and readers. This process does not appear in most other visual analysis domains and visualization storytelling, where the reader's needs do not directly factor into the creation of the story.

This interactive report evolution process can be heavyweight (through phone calls, emails, etc.) and take several iterations to iron out, while the communication details are often lost between different versions of the report. At each iteration, creators explore the desired data using interactive visualization dashboard tools, create initial visualizations, provide additional information and details about the analytical techniques used, and finalize reading paths (i.e. how to read a report).

Once the story is finalized, it serves to clarify connections in the data and answer specific questions analyzed by BI analysts (see section on BI Reports). These answers often result in more detailed or tangent questions from the audiences, and the need for analysts to conduct side data investigations and generate new stories. As in other visualization domains, BI stories are thus a visual communication medium between storytellers and their audiences, but contrary to other domains, there is an open communication loop, where audiences can continue to ask for new stories.

3.5 BI Storytelling Challenges

All experts use an in-house reporting tool, that gives access to different data sources and can extract interactive dashboards, charts and tables from their dashboard analysis system. It also provides the capability of adding text. Three (P1,P3,P5) also used Excel for creating less complex reports. Our experts identified several issues with this process, which are similar to other BI analysis and reporting tools.

When extracting interactive visualizations from their analysis tools, metadata and annotations from their analysis are not extracted and have to be recreated. Moreover, the annotation capabilities of the report creating software are very limited. They cannot annotate specific data points, while sequences and connections cannot be displayed graphically but have to be explained in one of the supporting material. Finally the report creation tool can be in-house software. Thus to share reports, creators often extract static snapshots of the tool output and save it in PDF format for their readers.

Some report creation tools give access to interactive visualizations and dashboards through hyperlinks. Our experts like this option and add these links to their non-interactive reports (PDF, Excel) and presentations. Readers within the organization may have access to the in-house reporting tool, and can then interact with individual visualizations or dashboards (e.g. query and filter data, or drill down/up). This interactivity is lost when reports are communicated outside the company.

To overcome the shortcomings of reporting tools, creators are forced to provide the supplementary material (presentations, wiki pages) and when possible links to interactive visualizations for in-company clients. They described this process as limiting and requiring a lot of work duplication.

We requested clarification on two points regarding the possible integration of analysis and storytelling: (i) Given that analysis dashboards are exploratory environments that allow users to interact fully with data, should such interactions be allowed in the storytelling? (ii) If creators had the opportunity to create their entire story with all support material in one place, how would they prefer the story to be visualized?

Interactivity. Due to the constraints of current reporting tools, shared reports are often non-interactive when accessed outside the organization. When creators were asked if they want the visualizations in the reports to be completely interactive and encourage readers to interact with them (e.g. using drill down/up, filter, link & brush), four (P1,P2,P3,P5) of our experts prefer to have interactive visualizations that permit linking and brushing (i.e. data selection). But they would limit the more advanced interactions such as drill down/up or filtering. They felt that all the data needed to tell the story should be displayed clearly in the report without the need to explore the data further. The fifth (P4) would not be opposed to fully interactive visualizations. Thus authors feel business stories should be mostly author-driven and constraint, known to work best when the goal is storytelling or efficient communication [26].

Story templates. Going from the current practices of storytelling (BI report, presentations, wiki) to a dedicated storytelling tool is not straightforward. We thus showed our experts a group of story templates identified by Segel et al. [26] to see if they met their needs. All chose the ”Annotated chart” as the preferred template (Fig. 2), with the modification that it should have multiple charts on the same page (dashboard) that they can annotate. Four experts (P1,P2,P3,P5) identified ”Partitioned poster” as a potential template, where the side list of charts display details that support the main chart in the central region. Three (P1,P2,P5) mentioned that the ”Slide Show” template is useful both as a means to focus on attention on each chart and zoom to details, and as a step-by-step presentation that clarifies the analysis path and the ideal reading sequence. One expert (P5) found the ”Flow Chart” useful for showing some business scenarios, like following bugs during development (discovered, tested, fixed etc.). Another (P3) found the ”Comic strip” useful template, but with added annotations.

So besides the templates identified by Segel et al.[26], for BI stories we need a new template that consists of an annotated dashboard. Our users attempt to do this with their current reporting tools, but they are limited (cannot annotate detailed points and relations), or they create it manually in their presentations.

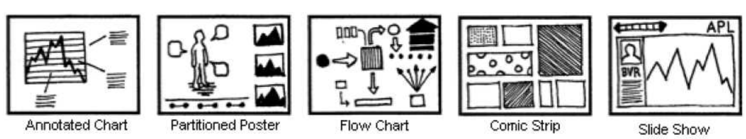


Fig. 2. Genres of visual narratives templates from Segel et al. selected by our BI experts.

4 Participatory Design Session

We brought back the senior BI expert that trained other analysts (P1), to identify the best design for a BI storytelling tool. She provided us beforehand with one of her old reports and supplementary material. During this session we gave the expert a copy of her material, and several cut-outs of the entities in the report (charts, titles or tables). She was also given ”narrative aids”, such as arrows, lines, numbered items, grouping containers of different colors, and annotation bubbles. We asked her to construct a stand-alone story of the report on an A3 paper, with accompanying audio if necessary, in a way that it can be interpreted by a general audience with no training. The session lasted 2 hours and was recorded.

Fig. 3 shows the final version of the manually created story by the expert, where we can see the intense usage of annotations and explanations, connected by arrows and lines. Other structures used by the participant were chart grouping rectangles (enclosures for placing charts that need to be read together), sequence markers (what needs to be read first), and highlight markers (to draw attention to part of the story).

We asked our expert how she would prefer the readers to see the story, as a static image or an animated presentation. She explained that both are needed, the static representation shows the ”entire story and gives context”, whereas the animation ”focuses the audience where I want”. She then played out for us how she wanted the animation to be presented, explaining when and how to zoom to specific areas of the story. This play-back was recorded using static shots.

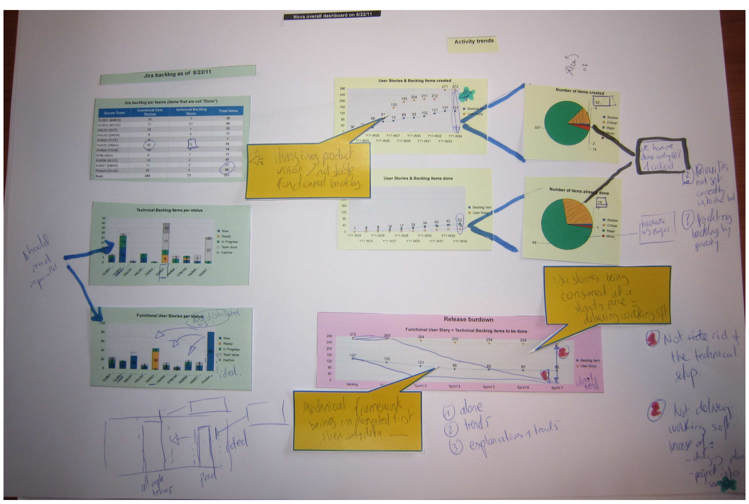


Fig. 3. Story created by an expert during the design session.

5 Requirements For Creating BI Stories

Based on expert self-reporting, BI reports, the main communication medium of BI analysis, consist of snapshots of visualizations and textual descriptions, made in a different environment than the one used for analysis. These reports are fact-based and can be interpreted only by an expert audience. To increase their audience, report creators use supplementary material, like wiki pages, presentations, and when possible links to interactive visualizations. Based on our experts' interviews and a participatory design session, we identified a set of requirements for enhancing BI analysis tools with storytelling capabilities:

R1. Fluid transition. Analysis tools used to explore data and create visualizations are different from report creation tools. Exporting visualizations from the first to the second to create a BI story costs time and effort, and limits the possibility of embedding meta data or annotations created during the analysis. To ensure that creators do not recreate information, they require a fluid and integrated way to transition from their analysis and meta data associated with it, to report creation.

R2. Integration. To tell their stories, BI creators need tools that combine all materials used currently in their story creation: BI reports, interactive visualizations, ways to indicate story structure, highlighting capabilities, presentation of the story in sequence, and textual or audio explanations.

R3. Narrative visual aids. Report creators need to add focus expressions to draw attention to specific visualization data, such as highlighting, coloring, annotating and zooming. They also require ways to indicate reading sequence (e.g. vectorial references, like arrows [13][30]). These are not available in reporting or BI analysis tools.

R4. Interactive visualizations. Visualizations on shared reports are often non-interactive when read outside the organization. A storytelling tool should have completely interactive visualizations, although the way that readers interact with the data should be limited (by default to brushing and linking) and be controlled by the creator. This balance has been identified as a challenging aspect of storytelling [20].

R5. Appropriate BI Story templates. BI stories have specific structure not necessarily shared by other story narratives identified by Segel et al.[26]. Our experts identified templates of interest (Fig. 2) and highlighted the need for a new template that consists of an annotated dashboard.

R6. Reuse. Although BI reports and data changes from analysis to analysis, often the underlying structure of BI stories remains the same. It is thus important to be able to easily reuse the structure of stories created within the tool both for stories of evolving data and similar future stories.

R7. Optional playback. Readers should be provided with a static representation of the entire story to get an overview and explore it on their own pace, as well as a guided playback to help them follow analysis paths that not easy to discover and understand the important points according to the creator.

Such storytelling support would facilitate the creation and learning to read BI stories. While the scientific visualization storytelling tool 32] supports R3,R4,R7, and GeoTime [9] supports R1 and R3, we are not aware of any BI dashboard or other multi-chart visualization system that provides *all* the above functionalities.

6 BI Narrative Prototype

Following these requirements, we extended an existing dashboard exploration tool to support the creation of BI stories.

6.1 Exploration/annotation dashboard (Fig. 4)

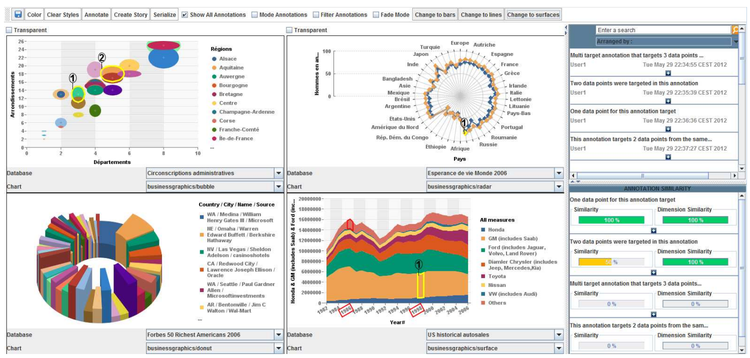


Fig. 4. Dashboard with visualizations created during analysis on the left, and list of annotations added to different data contexts while exploring analysis visualizations on the right.

The user starts from a traditional analysis dashboard, a collection of coordinated (synchronized) charts connecting one or more data sets. It provides advanced exploration capabilities such as data selection and filtering, and OLAP functionality (e.g. drill-up/down) [1]. We build our narrative tool on top of an existing visualization dashboard system [10] that supports annotations on ”data targets”, such as charts or tables (e.g. a bar-chart) or parts of them (e.g. a specific bar in the chart). Annotated data targets are highlighted in the dashboard and an icon indicates the number of attached annotations. A list of all annotations is also available on the right.

After or while conducting her analysis on the dashboard, the analyst can create a story. The menu option ”Create story”, opens up the narrative board window. All visualizations and their annotations (annotation text + data targets highlighted) are placed in the narrative board. Thus the analyst can transition fluidly (**R1**) from analysis to story creation. Because dashboards can present evolving data, but a story can be an instance in time, by default each visualization is placed on the narrative board with a time-stamp to indicate when the visualization was taken from the data.

6.2 Narrative board (Fig. 5)

Here users create a BI story. Visualizations and annotations from the dashboard appear on the main window, arranged by default as in the dashboard (to support the dashboard template identified in our interviews **R5**). Users can resize and relocate all story entities freely, or choose to organize them using other appropriate narrative templates (**R5**). They can edit visualization titles, annotations of entities, and add new text entities. Users can add visual entities that indicate relations, grouping and emphasis (**R3**) to help create stories, available at a tool palette on the right. The story can be seen in: a static representation, where all entities remain one screen; and a playback representation, where entities are highlighted sequentially according to an author-defined sequence (**R7**). We categorize the entities available on the narrative board as:

1. Information entities: visualizations, text, annotations. These are mainly imported visualizations and annotations created during the analysis. For annotated visualizations, the annotation text, the visualization, as well as the annotated data target (specific part of the visualization that is annotated), are all defined as story entries. Users can also place text anywhere on the narrative board to further explain their story.

2. Relational entities: arrows, lines, html links. A story can include relationships between entities, such as causality. We provide visual arrows, lines, html links and other vectorial references to define relations between entities (Fig. 5.c).

3. Organization entities: visual grouping and sequences. Previous work [25] and our interviews emphasized the need of visual grouping of entities that are to be seen together in a story. We support this with entity grouping borders. Our experts also indicated the need to define a reading sequence for each entity, to help readers move through the story. We achieve this by allowing users to define the order of appearance of all entities through a sequence list (Fig. 5.c). Users can change the entity's sequence in the list, delete, or rename any entity. They can also define a playback time for each entity, a time for the entity to be in focus in the playback presentation mode. Audio can also be recorded for any entity, to be played during the story playback.

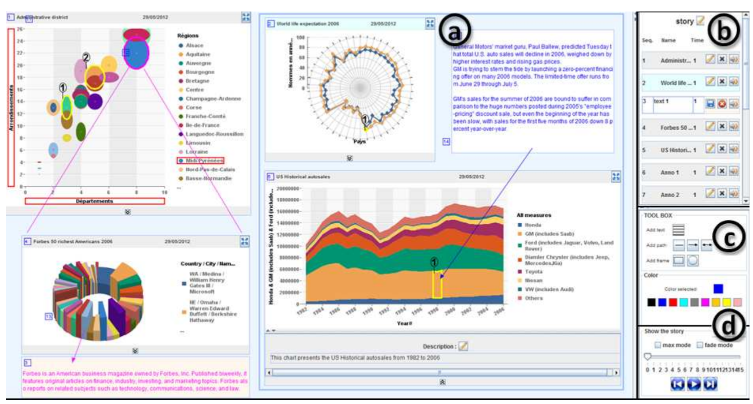


Fig. 5. (a) Narrative board containing all story entities arranged by the story teller: information, relational (arrows), organization (groupings) and emphasis entities, and numbers indicating author reading sequence. This sequence appears in a list (b), where authors can define playback properties and add audio commentary. Below is a pallet of relational and organization entities (c), and the Playback panel (d) to control playback through the story time line.

4. Emphasis entities: highlighting and zooming. To focus reader's attention to a specific story entity, users can add color highlights to any entity (e.g. a visualization, an arrow, or text entity), or select and highlight parts of visualizations (e.g. a bar in a bar chart visualization). This color highlighting can be present in the static story presentation (always visible), or during playback (the color highlight appears at a specific point in the story timeline). Besides color highlighting, users can add zooming highlighting to any story entity, to take place in a specific point in time during playback.

Playback. As discussed in (**R7**) our experts desire 2 ways of showing their stories to their readers, a static overview version that provides context and allows free exploration, and a playback that shows a recording of the author suggested path. Our narrative board acts as the static version. We provide three options of animated playback, giving focus to entities in sequence for a given duration. In the "color highlight", entities in focus change color to grab attention (default playback). In the "max playback", entities in focus are zoomed-in to the maximum possible size, taking up the entire narrative board. Finally, in the "fade mode", during playback, all entities except the entities in focus fade out (Fig. 6). As mentioned, authors can record audio to accompany the playback. Readers can pause the playback at any point to explore the story.

Fig. 6. Two playback modes: Max (a), and Fade (b).

Interactive visualizations and exploration. Our visualizations are imported with time-stamps referring to a "snapshot" of the dashboard data at the time of the story creation. Nevertheless, they are still connected to that version of the data and are still live. Thus users can still interact with them to perform brushing and linking actions (**R4**) and explore them further. By default we deactivate advanced interactions like drill-down/up, but authors can reactivate them through the visualization's properties.

The creators may also decide to take a new snapshot of the data (i.e. change the timestamp for the visualizations, and so the values of the data). Thus they can reuse the story structure for a new version of data (**R6**), which is particularly important for evolving dynamic data. Finally, users can select and replace any visualization with another, even from another dashboard loaded in the system. Thus they can reuse the story structure not only for dynamic data, but also for completely different datasets.

Thus our system integrates all the material used currently in BI story creation (**R2**): BI reports (visualizations and annotations), interactive visualizations, ways to indicate story structure, highlighting, optional story presentation in sequence, and text or audio explanations. It can then be shared with readers with access to the prototype. With respect to other BI systems, ours is one based on a user-centered design based and expert user requirements, and is unique in supporting annotations that link multiple points in the story, different story playbacks, and providing templates for BI stories.

7 User Feedback

We conducted two user feedback sessions to assess the usability and effectiveness of the system, both from the creator's and the reader's perspective.

7.1 Story creation by experts

We invited back two of the interviewed experts (P1,P2), and showed them the system in individual one hour sessions. Expert P1 was the experienced analyst (7 years) that has trained others. Expert P2 has been an analyst for 6 months and was trained by P1. They explored the prototype, saw story examples, and created stories themselves.

General feedback and recommendations. Both experts were very enthusiastic with the prospect of having access to such a system for their work. They found it easy to use and helpful, especially as it is integrated in the exploration/annotation dashboard tool. As P1 mentioned it "saves me from recreating any charts or annotations to present to others". It also saves time and effort not just in report creation, but also in communicating reports: as P1 explained, by using the prototype, there is no need for an explanation presentation to clients, for wiki pages or recreating interactive charts.

Both creators reaffirmed that story reading must be guided by the story creator, else the goal of the story may be lost. Currently they enforce this by the order of explanations used in their BI reports. They commented that the prototype supports this well with the numbered sequence for entities, while the playback (with fade out and zoom) guides readers through the story and keeps their focus on one entity at a time. But they appreciated that they can create and show a story on one page (even with scrolling), as opposed to the current multi-page reports, because as P1 mentioned "people tend to read only the first page, and explore less the following pages". And because it gives readers flexibility "the reader can always understand the position and link between parts and the overall story" and "look for more details when she wants".

Participants liked the ability to use the annotated dashboard template, saying that "this is how we want to present our analysis story, similarly to our dashboard". They both commented that annotations attached to data points are very important in pointing out to readers important data values "I add a lot of manual arrows to point to annotations that refer to specific data points on a chart, and their ideal value".
Both experts suggested our system should support two types of BI narrative stories: *(i)Fixed stories*, that present snapshots of datasets at specific points in time, yet are interactive (e.g. for filtering), to be shared as tutorials, explanations, or reports. This is the default case in our system. And *(ii) Online stories,* that present dynamically evolving data, and can have the same analytic scenario regardless of data values. Thus stories may have the same chart descriptions (e.g. what data is shown), the same KPI relations, and the same reading sequence. Here visualizations in stories are no longer snapshots, but are updated with data changes. We have implemented this extension.

Collaboration and Communication with BI stories. Expert P2 commented on how this storytelling prototype can also be used as a means to evolve stories. Multiple analysts can integrate their own comments and knowledge in the story, encouraging peer learning, but also collaboratively creating more complete and detailed stories.

Expert P1 mentioned that she could envision using the system to iterate the definition of the story directly with her clients (a process conducted now by email or phonecalls). She envisioned clients adding themselves on the story further explanations on specific entities and their relations (to explain patterns), possible summaries of decisions they took based on the report, or highlighting what information is missing. This goes beyond collaborative analysis and storytelling: it directly empowers readers and becomes a medium to communicate what they want from the story.

In both cases, the prototype moves from a one-way communication to a collaboration medium, where the authoring of a story opens-up and evolves with the contributions of many users, and acts as an archive of knowledge and different points of view. Such a system, our experts explained, needs to clearly differentiate between the contributions of individual authors. We are currently exploring this extension.

Finally the story can be archived and used by new analysts that learn how to create stories (comment from our trainer analyst). P1 stressed how important such an archiving is for knowledge passing between analysts. The recently trained P2 stated the tool can help him further improve his reports by looking at the story structure of others.

7.2 Story reading by novices

We then ran a second session to evaluate the prototype from the reader's perspective, and thus close the story communication cycle. We conducted 40 to 50 minute sessions, with 5 BI novices. All were IT professionals, knew what a dashboard is but had never worked with one. Two had heard of BI reports but never used one. Participants were asked to (1) Read a BI report created by one of our experts, (2) Read a BI story (created by an expert in the previous session), and (3) Explore our prototype.

The story presented the progress of a development project from different perspectives: General (how is the project evolving in terms of finished code in a sprints time line, how many code components are added to a waiting list, and how many critical, major or minor code components are done each week); and Detailed (the progress of each development team in coding and testing each code component). The development is not progressing according to plan because many bugs fixed are not critical, whereas new bug reports coming in add critical bugs to the waiting list. The bottom right chart in the report highlighted the problem and the rest provide details.

Reading the report. When given the report, all readers read the first page dashboard from left to right and up to down (which is not the ”author” suggested order). They understood the goal (progress of a project) and what the charts displayed (e.g. bugs in waiting list) with the aid of chart titles and legends. But they all struggled to find the problem illustrated in the report. Only one participant noticed that the project development is not progressing over time, but she could not understand why. This supports our experts' comment that reports cannot be read without the supporting material.

Reading the story. Participants found that reading a story was easier ”it showed the facts in an understandable manner”. All 5 readers found the system easy to use, understood the story, and were able to answer correctly comprehension question related to the story content. We report here our main observations and readers' comments:

Memorization: All users remembered even detailed aspects of the story and only went back to the system when answering detailed quantitative questions (e.g. to retrieve numbers from charts), while all general comprehension questions were answered without going back to the story visualizations and comments. They tended to answer using the same terms used by the expert analyst in the story. When asked similar questions on the report, they went back to it each time to search for answers.

Confidence: Readers were not confident in their understanding of the report as they had to draw their own conclusions. They expressed worry that they may have misunderstood or not noticed important points. While when reading the story, all felt confident, as their interpretations are confirmed by the analyst's comments.

Guidance: All readers appreciated the guidance in reading the story, both static and playback. They commented on the importance of both modes: "the static mode permits to understand the whole story" and "dig for facts", while in playback it "is easier to follow" the story sequence and focus on important data. Four participants preferred the fade playback, and three the max mode when focusing on an entity.

Understanding: All users understood the story using our tool, and found it easy to read and interpret. They commented on how it was easy to find answers both to qualitative and quantitative questions. The story structure showed clearly what is the problem, how to analyze it and how to find the cause. While they described the report as ambiguous, as they couldn't identify the relation between charts or KPIs.

Transmitting knowledge: All readers found annotations very helpful in explaining the relationships between charts and KPIs, and in teaching them the analysis logic they should follow. Four mentioned that the system can "aid in transmitting different knowledge in the company between different users", and two would like to use the system to communicate with their team their own data (even if they are not analysts).

Engagement: Three readers got very engaged with the story, as "stories are more encouraging than static reports". They began searching in charts to find how this problem can be solved, exploring the story outside the suggested structure.

The comparison of BI reports without supplementary material to our story prototype is clearly unfair and our goal is not to prove its superiority in terms of communication value. It serves to illustrate how our stand-alone prototype could be an effective means of communicating BI stories without additional material. A comparison of our tool against the collective BI story material used today is future work, but our experts already identified its superiority in terms of saving time on report creation, the sharing of interactive stories, and being a medium for collaborative story evolution.

8 Discussion and Conclusions

We identify for the first time the current practices and needs of BI storytelling. BI analysts often organize visually and communicate their analysis story to others. Nevertheless their tools do not allow easy transition from visual analysis to storytelling. They often use multiple tools, replicate work, and train their audience to understand their analysis. This reinforces the need for explicit storytelling support, missing in most existing visual analytics systems even in other domains, and provides insights on how to address this need for dashboards and other coordinated view systems.

Based on interviews and a paper prototyping session with expert BI analysts, we derived requirements for extending BI visual analysis dashboards to support storytelling. These can be adopted as-is by designers of BI systems, or inspire and inform visual storytelling research in other domains. For example providing an author suggested story structure, seeing the story in either a playback or it in a static overview, and imposing exploration constraints to interactive visualizations, can apply to other domains. Others, like the need to reuse story structure, may be BI specific and need to be reexamined for other domains. Using these requirements we implemented an extension to a BI dashboard, allowing transition from analysis to storytelling.

We evaluated our prototype with story creators and found the requirements and prototype meets their needs. More importantly, they highlighted the potential of storytelling tools as (possibly asynchronous) two-way communication channels, were authors communicate their findings, and readers also pose questions directly on the story. This empowers readers, informs authors of limitations of their story, and captures the evolution of the story. In BI it also accelerates story creation, that is highly client driven, and whose details are often lost in emails and phone calls. Tools can also act as collaborative platforms, where multiple authors refine, create variations of, and archive stories. Our findings open questions regarding the archiving and navigation mechanisms of story versions, trust in authoring changes, and maintaining a clear story focus. These questions open up exciting new research avenues.

We then presented a story from an expert to novice readers, to test the understandability of stories in our system, and its potential as a stand-alone tool for communicating BI stories. Readers understood the stories without training, and answered complex comprehension questions. Few previous works (e.g. [9]) evaluate visual storytelling systems from the reader's perspective, a crucial step in the communication cycle.

Our system was identified by experts as having great potential for reaching a broader audience, as little to no training is required to read stories, and for training analysis. Although designed as a stand-alone tool, it does not aim to replace expert analysts. As an expert mentioned, ours and other storytelling tools can aid novice analysts or readers to quickly read analysis results and focus their questions to experts on more complex aspects, such as methodology, goals, content details, etc.

We are finalizing our prototype to give to BI analysts. We aim to verify requirements gathered through expert self-reported with more direct observations of the use of the tool and its impact on current analysis and communication practices.

9 References

1. OLAP and OLAP server definitions. The OLAP Council, 1995.
2. C. Baber, D. Andrews, T. Duffy, and R. McMaster. Sensemaking as narrative: Visualization for collaboration. VAW2011, U. London College, 2011.
3. E. A. Bier, S. K. Card, and J. W. Bodnar. Entity-based collaboration tools for intelligence analysis. IEEE VAST, 99–106, 2008.
4. D. M Boje. The storytelling organization: A study of story performance in an office-supply firm. Administrative Science Quarterly, 106–126, 1991.
5. D. M Boje. Stories of the storytelling organization: A postmodern analysis of disney as ”tamara-land”. Academy of Management Journal, 997–1035, 1995.

6. M. E. Boyce. Collective centring and collective sense-making in the stories and storytelling of one organization. Organization Studies, pages 107–137, 1995.
7. M. H. Brown. That reminds me of a story: Speech action on organizational socialization. University of Texas at Austin, 1982.
8. J. Cutroni. Google analytics. OReilly Media, Inc, 2007.
9. R. Eccles, T. Kapler, R. Harper, W. Wright. Stories in geotime. IEEE VAST, 19-26, 2007.
10. M. Elias and A. Bezerianos. Annotating bi visualization dashboards: Needs & challenges. ACM CHI'12, 1641–1650, 2012.
11. P. Farhi. CNN hits the wall for the election, http://bit.ly/cnn-wall, 2008.
12. S. Few. Show me the numbers: Designing tables & graphs to enlighten. Analytics Press.
13. C. Forceville. Educating the eye? kress and van leeuwens reading images: The grammar of visual design (1996). Language and Literature, 8(2):163–178, 1999.
14. Nahum Gershon and Ward Page. What storytelling can do for information visualization. Commun. ACM, 44(8):31–37, 2001.
15. J. Heer, F. B. Vi´egas, and M. Wattenberg. Voyagers and voyeurs: Supporting asynchronous collaborative visualization. Commun. ACM, 52:87–97, 2009.
16. J. E. Helmer. Between horses: An ethnographic study of communication and organizational culture at a harness track. U. of Illinois at Urbana-Champaign, 1989.
17. M. Johnson. The moral imagination. Chicago: University of Chicago Press, 1993.
18. M. R. Louis. Surprise and sense making: What newcomers experience in entering unfamiliar organizational setting. Administrative Science Quarterly, 1980.
19. M. R. Louis. Organizations as culture-bearing milieu. Organizational symbolism: Monographs in organizational and industrial relations, 39–54, 1983.
20. K.-L. Ma, I. Liao, J. Frazier, H. Hauser, and H.-N. Kostis. Scientific storytelling using visualization. IEEE Computer Graphics and Applications, 32(1):12–19, 2012.
21. A. Maclntyre. After virtue. Notre Dame, IN: University of Notre Dame Press, 1981.
22. W. McWhinney. A health confluence. Unpublished paper, 1995.
23. W. McWhinney and J. Battista. How remythologizing can revitalize organizations. Organizational Dynamics, 46–58, 1988.
24. I. Mitroff and R. H Kilmann. Stories managers tell: A new tool for organizational problem-solving. Management Review, 1975.
25. P. Pirolli and D.M. Russell. Introduction to this special issue on sensemaking. Human-Computer Interaction, 2011.
26. E. Segel and J. Heer. Narrative visualization: Telling stories with data. IEEE InfoVis 2010.
27. Y. B. Shrinivasan and J.J. van Wijk. Supporting the analytical reasoning process in information visualization. ACM CHI '08, 1237–1246, 2008.
28. D. Snowden. The art and science of story or 'are you sitting uncomfortably?'. Business Information Review, 215–226, 2000.
29. D. Sole and D. G. Wilson. Storytelling in organizations: The power and traps of using stories to share knowledge in organizations. Harvard Graduate School of Education, 2002.
30. B. Tversky, J. Heiser, S. Lozano, R. MacKenzie, and J. Morrison. Enriching Animations, volume Learning with animation. Cambridge University Press, 2007.
31. F. B. Viegas, M. Wattenberg, F. van Ham, J. Kriss, and M. McKeon. Manyeyes: a site for visualization at internet scale. IEEE TVCG, 13:1121–1128, 2007.
32. M. Wohlfart and H. Hauser. Story telling for presentation in volume visualization. IEEE EuroVis 2007, 91–81, 2007.
33. W. Wojtkowski and W. G. Wojtkowski. Storytelling: its role in information visualization. European Systems Science Congress, 2002.

The Attraction Effect in Information Visualization

Evanthia Dimara, Anastasia Bezerianos, and Pierre Dragicevic

Abstract—The attraction effect is a well-studied cognitive bias in decision making research, where one's choice between two alternatives is influenced by the presence of an irrelevant (dominated) third alternative. We examine whether this cognitive bias, so far only tested with three alternatives and simple presentation formats such as numerical tables, text and pictures, also appears in visualizations. Since visualizations can be used to support decision making — e.g., when choosing a house to buy or an employee to hire — a systematic bias could have important implications. In a first crowdsource experiment, we indeed partially replicated the attraction effect with three alternatives presented as a numerical table, and observed similar effects when they were presented as a scatterplot. In a second experiment, we investigated if the effect extends to larger sets of alternatives, where the number of alternatives is too large for numerical tables to be practical. Our findings indicate that the bias persists for larger sets of alternatives presented as scatterplots. We discuss implications for future research on how to further study and possibly alleviate the attraction effect.

Index Terms—Information visualization, decision-making, decoy effect, attraction effect, asymmetric dominance effect, cognitive bias.

1 Introduction

Suppose you are voting for primary elections and need to choose between candidates Bob and Alice (Table 1). Bob has a solid education plan, but not much concern for crime control. In contrast, Alice's education plan is weak but she has an excellent strategy for crime control. If both education and safety are important to you, this can be a difficult choice. Now suppose there is a third candidate, Eve. Like Alice, Eve focuses more on crime control than education, but her crime control strategy is not as good as Alice's. O'Curry and Pitts [38] used a similar decision task in a study, and showed that adding Eve as an option shifted participants' preference towards Alice.

Table 1: Three hypothetical candidates in political elections

	Bob	Alice	(Eve)
education	★★★★★	★★	★★
crime control	★★	★★★★★	★★★★

This shift in preference called the *attraction effect* (also known as the *decoy effect* and the *asymmetric dominance effect*), is a cognitive bias whereby people tend to favor the option for which there exists a similar, but slightly inferior, alternative. Like other cognitive biases, the attraction effect leads to irrational decisions and has important implications in many areas such as politics and advertising. Our goal in this article is to find out whether the attraction effect also has implications for information visualization design. In our example, voters' decision is influenced by the presence of Eve, which is inferior in all respects and therefore irrelevant to the choice. If, in the same way, if someone uses a visualization to choose among several options (e.g., when buying an apartment [55]), will the presence of inferior choices affect their decision? In other words, does the attraction effect transfer to visualizations?

The current information visualization literature does not offer much empirical data to help us answer this question. Although biases and misjudgments have been studied, the focus has been on perceptual biases such as in color perception or magnitude estimation [53]. Cognitive biases differ from perceptual biases in that they persist even if the information has been correctly processed at a perceptual level. There is a growing interest in cognitive biases in information visualization, but studies have so far focused on probabilistic reasoning and judgment under uncertainty [34, 42].

While there has been little work studying the role of cognitive biases in information visualization, visualization systems are increasingly used to support decision making. Large companies switch to visualization solutions to improve their human strategic decisions for profitable drug trials [46], or use visualizations to choose which features of a software they should release and when [3]. In addition, many visualization tools previously introduced in research explicitly or implicitly claim to help people make decisions such as choosing a house to buy [55], finding a nursing home [57], selecting healthy cereals [58], choosing a digital camera [16, 32], finding a profitable investment [42, 13], or selecting a site for a new factory branch [2].

A visualization is generally considered effective if it helps people extract accurate information [9, 59]. Nevertheless, we know from decision making research that full access to information does not necessarily yield good decisions [28]. Generally, the more complex a decision, the more we resort to heuristics, i.e., *"simple procedures that help find adequate, though often imperfect, answers to difficult questions"* [28]. While heuristics can be very effective [18], they can also lead to cognitive biases [28]. Therefore, in order to fully understand how information visualizations can support decision making, we need to study how they interact with cognitive biases.

We focus on the attraction effect for two reasons. First, it is one of the most studied cognitive biases in fields such as psychology, consumer research and behavioral economics. Second, these studies generally employ very small sets of alternatives (typically three) and numerical presentation formats, so it is still unknown whether the bias generalizes to data visualizations. Although some visual representations have been considered, there is conflicting evidence and a heated debate on whether the effect generalizes [17, 26, 48, 56]. Some argue that the effect occurs only in numerical stimuli [17], e.g., when attributes are presented in tables. Whereas others argue that it is generic and robust, and can be observed in many contexts such as visual judgments in shapes [52], oral instructions [44], or even among animals when they choose their food [31]. This debate suggests that the attraction effect is far from being fully understood and needs to be investigated from a variety of perspectives.

We study the attraction effect from an information visualization perspective in two crowdsourcing experiments. In the first, we test and verify that the attraction effect indeed persists when alternatives are presented in a scatterplot rather than in a numerical table. We then generalize the attraction effect procedure to more than three alternatives, and verify that the effect can persist when participants are presented with more realistic scatterplot visualizations involving about 20 data points. We finally discuss our findings and conclude with implications for future research.

- *E. Dimara is with Inria and Université Paris-Saclay. E-mail: evanthia.dimara@gmail.com*
- *A. Bezerianos is with Univ Paris-Sud & CNRS (LRI), Inria, and Université Paris-Saclay. E-mail: Anastasia.Bezerianos@lri.fr*
- *P. Dragicevic is with Inria and Université Paris-Saclay. E-mail: Pierre.Dragicevic@inria.fr*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx/

2 Background

We review work on decision-making and cognitive biases in information visualization, and on the attraction effect in other fields.

2.1 Information Visualization

Compared to numerical and textual formats, it is known that data visualizations can highlight relationships in the data, facilitate the recognition of patterns, and reduce cognitive load [9, 43, 50]. As they aid data exploration and understanding, it is generally assumed that data visualizations can support better decision making [2]. Based on this intuition, several decision-support systems that rely on interactive data visualization have been developed [59].

2.1.1 Use of Visualizations in Decision Support

A range of interactive data visualization tools have been proposed to help people make decisions. Sometimes decision making is simply used to demonstrate a new visualization or interaction technique. For example, HomeFinder [55] helps people find a house to buy using scatterplot visualizations and dynamic queries. The Dust & Magnet tool [58] is illustrated with a scenario for choosing cereals by moving magnets that attract or repel cereals according to attributes such as calories. Other tools target explicitly decision support. For example, Asahi et al. [2] visualize hierarchical decision criteria using treemaps, augmented with interactions to make decisions such as whether or not to construct a dam, file a patent, or choose a factory's location. ValueCharts [5] let consumers choose a TV set or a hotel by providing a set of domain-independent visualizations.

Domain-specific decision-support visualization systems have also been proposed. For example Decision Map / Table & Box [57] helps people find an appropriate nursing home by combining several coordinated views. Stratos [3] helps software project managers select which features to include in each production stage, by simultaneously visualizing all possible software release plans. VisIDM [13] helps people choose a financial investment through uncertainty visualizations and support for personalized risk preferences.

Although most of these tools come from research, similar ones are used in industry. For example, after losing millions of dollars in late drug trial failures, a large pharmaceutical company decided to use interactive visualizations to better track and facilitate decisions of "cut or go" projects in their early stages [46].

2.1.2 Limitations of a Pure Informational Approach

Interactive data visualizations facilitate data exploration and sensemaking, making data accessible and promoting informed decisions. Furthermore, the use of interactive systems, rather than automatic analysis, leaves room for human judgment, which is crucial where expert knowledge or subjective preferences cannot be fully formalized (e.g., importance of education vs. crime control). However, most visualization tools for decision support appear designed under the assumption that decisions are made by rational people who only need to be given complete information to be able to make good decisions. And thus, that good decisions should be the natural outcome of reliable data conveyed with well-designed visualizations.

It is by now widely recognized that even perfectly informed people are not perfect decision makers [28]. The imperfections of heuristics we routinely use manifest themselves as cognitive biases, like the attraction effect. Cognitive biases are far from trivial to overcome: they occur even when all relative information is available and well perceived, and they persist even when we inform or train people on how to overcome them [19]. Thus we need to investigate further if visualization designs are likely to suffer from cognitive biases, and whether we could improve our designs to alleviate these biases.

2.1.3 Cognitive Biases and Visualizations

Information visualization has studied perceptual biases [53, 60], but cognitive biases have comparatively received little attention.

Zuk and Carpendale [61] discuss cases where visualizations may aid to remediate uncertainty biases. Researchers have studied how visualizations, such as Euler diagrams and frequency grids, can reduce the base rate bias in probabilistic reasoning [34, 29]. FinVis [42] is a tool that shows investment options using tables and visualizations to help investors overcome the uncertainty aversion and diversification bias. Miller et al. [35] used scatterplots and histograms to help fantasy baseball experts overcome regression bias in their predictions. Although many of these previous studies try to examine how visualizations can help overcome cognitive biases, some studies found that visualization-based remediation can be challenging [34, 30], or that cognitive biases can co-occur with visualizations. For example, Zhang et al. [59] showed that startup companies presented with tabular visualizations were subject to conservatism and loss aversion biases in their probability judgments. Some biases, like the within-the-bar bias, only appear with visualizations [11].

Most of these previous studies focused on judgment under uncertainty. Although reasoning based on uncertain information is hard and pervades our everyday lives, uncertainty is not the only cause of irrationality in decision making. In the attraction effect, irrationality stems instead from the fact that decisions are influenced by irrelevant information (the presence of a decoy). Our work is thus significantly different from previous work about reasoning under uncertain information.

2.2 The Attraction Effect

We next define the attraction effect and the terminology used in this article. We present theories on why the effect exists, and discuss recent studies investigating the effect on visual stimuli.

2.2.1 Terminology

A **decision task** involves choosing one among several **alternatives** (i.e., Alice, Bob or Eve in our example). Alternatives are characterized by **attributes** (e.g. their support for education and crime control), which take values that are unambiguously ordered in terms of preference (e.g. more crime control or education is better than less).

An alternative **dominates** another if it is strictly superior in one attribute and superior or equal in all others. An alternative is **dominated** within a set of alternatives if there is at least one alternative that dominates it. In our example Eve is dominated by Alice, because she is equal in education and worst in crime control. In this decision task Eve would be formally a "wrong" answer.

An alternative is **asymmetrically dominated** within a set of alternatives if it is dominated by at least one alternative, but is not dominated by at least one other [25]. Eve is asymmetrically dominated because she is dominated by Alice but not Bob, since Eve offers better crime control than Bob. We call two alternatives **formally uncomparable** if neither dominates the other, as is the case for Alice and Bob. The best candidate is a matter of personal choice.

A typical attraction effect experiment involves a decision task with three alternatives, two that are formally uncomparable, and one that is asymmetrically dominated. They are referred to as: the **decoy**, the asymmetrically dominated alternative (Eve); the **target**, the alternative that dominates the decoy (Alice); the **competitor**, the alternative that does not dominate the decoy (Bob). This decision task is typically compared with a task where the decoy is absent, i.e. that involves only the two formally uncomparable alternatives.

The **attraction effect** is a cognitive bias where the addition of a decoy (Eve) in a set of two formally uncomparable alternatives increases people's preference for the target (Alice) [25, 27]. In experimental settings this preference switch is observed not for any single individual but between groups, where a higher percentage of people generally choose the target when the decoy is present. This switch in preference is irrational because it violates a basic axiom of rational choice theory, the principle of *regularity*, according to which the preference for an alternative cannot be increased by adding a new alternative to the choice set [25]. Attraction effect experiments assume that decision makers behave rationally in all other respects, and that they are able to perceive dominance relations. As a consequence, they are expected to never choose the decoy.

Later on, we will generalize the attraction effect to more than three alternatives. For now, we discuss previous work on the attraction effect, which always involves two alternatives plus a decoy.

2.2.2 Why does the Attraction Effect Occur?

Two types of explanatory theories have been offered for the attraction effect: strategic ones and perceptual ones [33].

Strategic Explanations: According to strategic theories, people use the dominance over the decoy as a heuristic to simplify an otherwise difficult decision. Choosing the target is also easier to justify to others [47] — in our example, someone who chooses Alice could argue that she is at least better than Eve. Neuroimaging studies have additionally shown that the presence of a decoy tends to reduce negative emotions associated with the decision task [21].

Perceptual Explanations: So-called "perceptual" theories assume that the addition of a decoy changes how people perceive the relative importance of the attributes involved, giving more weight to the attribute on which the target is strong [1, 22]. By analogy with perceptual contrast effects (e.g., an object appears larger when surrounded by small objects), the target appears more attractive when surrounded by unattractive alternatives [49]. In our example, if Eve is present, crime control may appear more important as two candidates perform relatively well on this criterion. Since this is the strength of Alice, it may raise her perceived value compared to Bob.

All explanations agree that for the attraction effect to occur, a perceptible dominance relation between the target and the decoy is key.

2.2.3 Can the Attraction Effect Occur with Visualizations?

Studies suggest the attraction effect is quite general and robust, e.g., it occurs when people choose consumer products like beers, cars, or films [25], when they gamble [54], select candidates to hire [24], decide which suspect committed a crime [51], or vote [38]. Even animals like hummingbirds [4], bees [45], and amoebae [31] appear to be subject to the same bias when selecting their food.

The attraction effect has been observed under a variety of experimental conditions, the majority of which present decision tasks as numerical tables. A few studies have shown that the effect generalizes to non-tabular representations, such as pictures of consumer products [49], verbal instructions [44], and physical objects (i.e., people choosing between cash and a pen, or between tissues and towels) [49]. Studies have further suggested that the effect occurs when carrying out visual judgment tasks, such as finding the largest rectangle [52] or finding similarities in circle and line pairs [10].

Nevertheless, several authors [17, 56] have recently argued that the attraction effect only occurs when attributes are presented in numerical format, and reported failures to replicate the previous studies involving the representations mentioned above. Others subsequently questioned the validity of these replications [26, 48]. This debate on whether the effect generalizes to non-numerical presentations opposes *(i) numeric* displays of *quantitative* information with *(ii)* displays of *qualitative* information such as photos, verbal descriptions, or physical objects. As most data visualizations are *pictorial* displays of *quantitative* information, the debate does not provide evidence on whether the effect occurs in visualizations.

Frederick et al. [17] however studied a gambling task with two or three bets presented either as a table, or as a diagram. Each bet had a prize in dollars and a probability to win. In the diagram condition, the probability of each ticket was shown as a "probability wheel" (analogous to a pie chart), and the prize was shown underneath, as a number. When gambles were presented as numeric tables, the decoy nearly doubled the share of the target, but when pie charts were used, the effect disappeared. To the best of our knowledge, this is the study that comes closest to a test of the attraction effect on visualizations. Nevertheless the diagram design was very domain-specific, and only one of the two attributes (probability, but not price) was encoded visually. We address this by using 2D scatterplots.

Although why the attraction effect occurs is still not fully understood, the possibility that it persists in visualizations is consistent with both the strategic and the perceptual explanatory theories. Both assume that the effect requires the ability to make attribute-to-attribute comparisons and to recognize the dominance relation between target and decoy. If anything, visualizations could make these tasks easier and could perhaps even amplify the effect.

3 Gym Experiment: Table/Scatterplot, 3 Choices

The purpose of this first experiment is to replicate the design of a standard attraction effect experiment (two alternatives plus a decoy presented in a numerical table), and then to test if the effect persists when alternatives are shown using a scatterplot visualization.

Similar to Frederick et al. [17] who successfully replicated the attraction effect with tables but not with non-numerical formats, our study was conducted using crowdsourcing. Crowdsourced experiments are now commonly used in information visualization [23], including in studies involving judgment and decision making [34, 29]. We used Crowdflower[1] as the crowdsourcing platform.

3.1 Design Rationale

Although the attraction effect is thought to be robust, a replication can fail if not enough attention is paid to the details of the experimental design [26, 48]. We therefore based our design choices on lessons and recommendations from the attraction effect literature.

3.1.1 Scenario and Attribute Values

By *scenario* we refer to the semantic and narrative context of the decision task. In our introduction example, alternatives are candidates, attributes are support for education and crime control, and the decision consists of voting for a candidate.

Many different scenarios and attribute values have been employed since the original studies of the attraction effect [25, 27]. We reasoned that a recent study is more likely to employ an optimal design, since it has more accumulated knowledge to build on. We therefore chose to replicate the scenario from the first experiment of recent work by Malkoc et al. [33], that involved choosing a fitness club (or gym), and found a clear attraction effect.

In Malkoc et al.'s study, each gym was defined by its variety and its cleanliness, both rated from -10 to +10. A positive rating meant better than average, and a negative rating meant worse. The study investigated whether undesirable options (all negative ratings) eliminate the attraction effect. But as the effect was strong for their control condition (all positive ratings), we chose it for our replication.

The experiment employed four gyms g_C, g_V, g_C^*, g_V^*, where g_C was cleaner, g_V had more variety, and g_C^* and g_V^* were slightly less attractive than g_C and g_V respectively. The attribute values were $g_C(\textit{variety}=1, \textit{cleanliness}=4)$, $g_V(4,1)$, $g_C^*(0,4)$, and $g_V^*(4,0)$. Three decision tasks were tested: $\{g_C, g_V\}$ (no decoy), $\{g_C, g_V, g_C^*\}$ (decoy on g_C), and $\{g_C, g_V, g_V^*\}$ (decoy on g_V). These attribute values however cause the data points g_C^* and g_V^* to overlap with scatterplot axes, possibly creating visual anchoring effects that could affect participant responses. Since such effects were outside the scope of our study, we incremented all values by one. Thus we used as attribute values $g_C(2,5)$, $g_V(5,2)$, $g_C^*(1,5)$, and $g_V^*(5,1)$. These values preserve all dominance and similarity relationships between alternatives.

3.1.2 Stimuli: Tables and Scatterplots

We used a numerical table as a control condition, to test our experiment design and compare our results with previous studies. Figure 1a shows the 2×2 table representation for the decision task $\{g_C, g_V\}$, and Figure 1b shows the 3×2 table for the decision task $\{g_C, g_V, g_V^*\}$. Attributes were presented in rows and alternatives in columns, as in Malkoc et al. [33]. Alternatives were labeled A, B or A, B, C from left to right. The ordering of rows and columns in the table will be discussed in the next subsection.

In the visualization condition, alternatives were conveyed with scatterplots (see Figure 1c,d) and similarly labeled A, B or A, B, C from left to right and from top to bottom.

There are four main reasons behind the choice of scatterplots for the visualization condition. First, 2D scatterplots are a standard information visualization technique [16, 37]. Second, they are suited for visualizing any tabular dataset with two quantitative dimensions, which captures the decision tasks used here and most decision tasks used in previous studies on the attraction effect. Third, a

[1] http://www.crowdflower.com/

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TVCG.2016.2598594

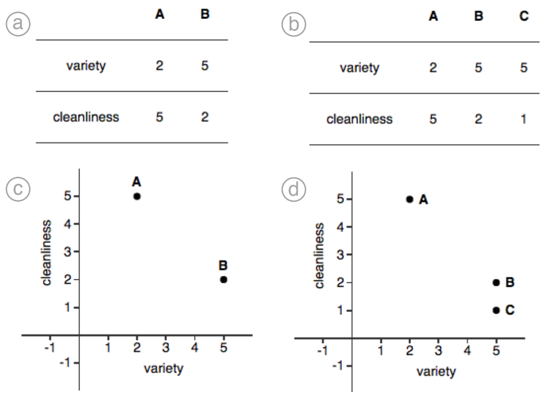


Fig. 1: Examples of experimental stimuli for the table (a,b) and the scatterplot (c,d) conditions. The left decision task (a,c) has no decoy, while the right decision task (b,d) has a decoy on B.

scatterplot shows all data cases within the same frame of reference, thus providing a rapid overview of all alternatives. A unified frame of reference also likely supports comparisons better than side-by-side views such as Frederick et al.'s [17] pie charts discussed in the background section. In fact, scatterplots are used as figures in most articles on the attraction effect for conveying the alternatives used in the experiments [4, 20, 24, 25, 27, 31, 33, 36, 38, 40, 45, 47]. Finally, scatterplots scale up to more than three items, which is an important requirement for our follow-up experiment.

The appearance of tables and scatterplots was kept as similar as possible to avoid experimental confounds due to choices in visual design. Both presentation formats took approximately the same amount of screen real estate, and graphical attributes (colors, line thickness and font sizes) were kept consistent. In both conditions, participants indicated their choice through separate radio buttons.

3.1.3 Ordering of Alternatives and Attributes

Although Malkoc et al. [33] used a fixed order of presentation for attributes and alternatives, the choice of ordering may affect participant responses, in particular in our experiment where different presentation formats are used. For example, participants may give more weight to variety if it is shown first on a table, but on a scatterplot, it is not clear whether the choice of horizontal vs. vertical axis would have a similar effect. In addition, alternatives can be presented in any order within a table, while on a scatterplot the way alternatives are laid out is dictated by attribute values.

To balance out any possible order effect, we thus randomized the order of presentation of attributes and alternatives across participants. In the scatterplot condition, axes can be flipped, leading to 2 different attribute orderings (variety on x and cleanliness on y, or vice versa). In a 2×2 table, there are 2 ways to order rows and 2 ways to order columns, yielding 4 different tables. Similarly, a 2×3 table can be presented in 12 different ways. Since the decoy is typically placed next to the target in attraction effect experiments (e.g., [17, 20, 21, 22]), we removed cases where the target was not next to the decoy (4 tables out of 12). Since the decoy cannot appear between the target and the competitor in the scatterplot, we also removed cases where the decoy was in the middle (4 tables out of 12). In summary, we used 4 different table stimuli and 2 different scatterplot stimuli for each of the three decision tasks $\{g_C, g_V\}$, $\{g_C, g_V, g_C^*\}$ and $\{g_C, g_V, g_V^*\}$, for a total of 18 different experimental stimuli.

3.1.4 Crowdsource Quality Control

Quality control is important in any crowdsource experiment [23], and in attraction effect studies in particular [48]. Quality was ensured by recruiting highly-rated crowdsource contributors (level 3 on the Crowdflower platform), by including test questions, and by devising a job assessment scheme prior to running the experiment. Four criteria were used for job assessment:

Completion time. A job completion time of less than 1 minute or more than 30 minutes was considered abnormal. Our pilots indicated an average task completion time of 6 minutes.

Justification. Participants had to provide a textual justification for their choice. Justifications were classified by one investigator as either proper or improper, depending on whether it made a reference – direct or indirect – to either cleanliness or variety. Participants were informed in advance that they would have to justify their choice, as this has been linked to a stronger attraction effect [47].

Prior preferences. After the experimental task, participants were asked if they suffered from an abnormal fear of dirt (or bacteriophobia), with "no", "yes", or "unsure" as answers. This identified participants with a strong prior preference for cleanliness, as strong prior preferences are known to reduce the effect [36, 41].

Table and scatterplot tests. After carrying out the task, participants were subjected to two screening tests: a *numerical table test*, and a *scatterplot test*, irrespective of the condition they saw. Both tests involved choosing between three laptops based on their RAM and CPU, with one laptop clearly dominating the other two (i.e., had both higher RAM and higher CPU). The tests were designed to be trivial, with a single correct answer, using a presentation format similar to the experimental task (see Figure 1). The purpose of the table test was to screen for contributors who did not pay attention to the tasks. The purpose of the scatterplot test was to control for visualization literacy [8], and make sure that participants were able to read scatterplots and to perceive dominance relations [26].

We classified jobs in three categories: the *Red*, where the job is rejected (and the contributor not paid); the *Orange*, where the job is accepted but the data discarded from our analysis; and the *Green*, where the job is accepted and the data kept in our analysis. Due to limitations in the Crowdflower platform we had to pay all contributors, but we report here on the three categories nonetheless.

A total of 437 jobs were submitted, after removing invalid completion codes and duplicate worker IDs. A job was marked *Red* if: the completion time was abnormal (1 % of all submitted jobs), the gym choice was not properly justified (14%), or the contributor failed the table test (12%). A job was marked *Orange* if: the response to the bacteriophobia question was "yes" (12% of all submitted jobs), or the contributor failed the scatterplot test (13%). In total, 16% of all submitted jobs were marked *Red* and 14% were marked *Orange*. These jobs were discarded from all our analyses.

3.2 Experiment Design

The experiment followed a 3×2 between-subjects design. The first independent variable was the decision task, which involved three different datasets: $\{g_C, g_V\}$, referred to as the *no decoy* condition; $\{g_C, g_V, g_C^*\}$, referred to as *decoy on cleanliness*; and $\{g_C, g_V, g_V^*\}$, referred to as *decoy on variety*. The second independent variable was the presentation format, with two conditions: *table* and *scatterplot*.

3.2.1 Procedure

We conducted a first pilot study to ensure the clarity of the instructions, and we then uploaded the experiment as a Crowdflower job.

Participants had to open an external 8-page Web form. They were told they would have to choose a fitness club based on two attributes: variety of the machines and cleanliness of the club. They had to assume that they had done some preliminary research, and had narrowed down their choices to two (in the no-decoy condition) or three (in the decoy conditions) clubs. They were then shown the gyms as a table or a scatterplot (Figure 1) and asked to choose one.

Once finished, participants rated their confidence on a 7-point scale and provided an open text justification for their choice. They also rated their enthusiasm towards fitness clubs on a 7-point scale and reported on whether they suffered from bacteriophobia. Finally, they were given the table and scatterplot tests (Section 3.1.4), and filled a short questionnaire with demographic information.

At the end of the experiment, participants copied the provided completion code and pasted it in the crowdflower platform to receive payment. The entire job took on average 6 minutes to complete, and participants were paid $0.60 upon completion.

3.2.2 Participants

Our population sample consisted of 305 crowdsource contributors who submitted valid responses, i.e., jobs classified as *Green* (Section 3.1.4). Job assignments were left on the crowdsourcing server until the planned sample size of n=50 per condition was approximately reached. We obtained n=54, 51, 50 for the *table* decision tasks, and n=47, 53, 50 for the *scatterplot* tasks.

A summary of our participants' self-reported demographics is shown in Figure 2 (map and bar charts labeled "Gyms"). As can be seen, participants tended to be educated young male adults.

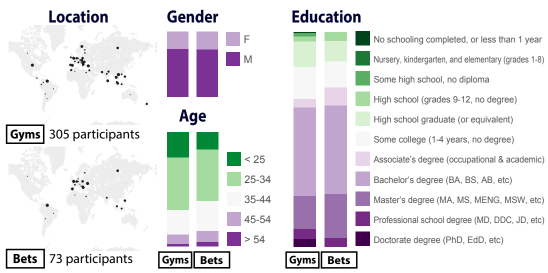


Fig. 2: Participant demographics for both experiments.

3.2.3 Hypotheses

Our statistical hypotheses were:

H1 A larger proportion of participants will choose the target in the *table* × *decoy on cleanliness* and the *table* × *decoy on variety* conditions than in the *table* × *no decoy* condition.

H2 A larger proportion will choose the target in the *scatterplot* × *decoy on cleanliness* and the *scatterplot* × *decoy on variety* conditions than in the *scatterplot* × *no decoy* condition.

3.3 Results

We analyze, report and interpret all our inferential statistics using interval estimation [15]. The experimental stimuli, data and analysis scripts are available at `http://www.aviz.fr/decoy`.

3.3.1 Planned Analyses

All analyses reported in this section were planned before data was collected. One planned analysis (an analysis of differences between attraction effects) was not conducted because it required equal sample sizes across all conditions.

Only one participant out of 306 chose a decoy, which is low compared to previous studies, where decoy selection rates can be as high as 13% [17]. This shows that participants carried out the tasks seriously and could perceive dominance relationships. The decoy choice is removed from the rest of this analysis.

Participant choices are shown in the top of Figure 3 marked "Gyms" ("Bets" refers to our second experiment) .The top three bars are for the *table* format, in the conditions *no decoy*, *decoy on cleanliness* and *decoy on variety*. Adding a decoy is expected to increase the proportion of choices of the target, in the direction indicated by the arrow. This was indeed the case for the *decoy on variety* condition (a 20% increase), but not for *decoy on cleanliness* (a 6% decrease). The next three bars refer to the *scatterplot* format. Here the expected increase was observed for both *decoy on cleanliness* (a 18% increase) and *decoy on variety* (a 3% increase). We now turn to inferential statistics to determine to what extent these effects are reliable.

The previously reported effects are shown in Figure 4 — the four black dots under the category "Gyms". Effects are expressed in percentage points, where a positive value (i.e., to the right of the vertical dashed line) indicates an attraction effect. Dots are sample

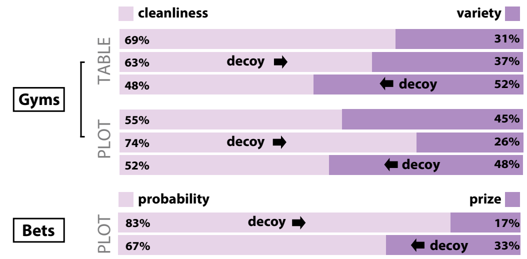


Fig. 3: Proportions of participant choices in both experiments.

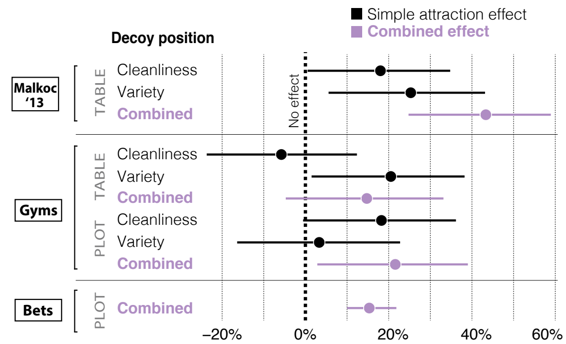


Fig. 4: Point estimates and 95% confidence intervals for the attraction effects in Malkoc et al. [33], and in our two experiments.

statistics, while error bars are 95% confidence intervals indicating the range of plausible population effects [12]. Confidence intervals were computed using score intervals for difference of proportions and independent samples.

Figure 4 shows that the unexpected reversal observed in the *table* × *decoy on cleanliness* is too unreliable for any conclusion to be drawn. The same is true for the small effect found for *scatterplot* × *decoy on variety*. However, we have good evidence for an attraction effect in the other two conditions. The magnitude of the effect is comparable to Malkoc et al. [33], shown on the top of Figure 4.

Thus, our results partially confirm **H1** and **H2**, but are less "clean" than in Malkoc et al.'s [33] original study.

3.3.2 Additional Analyses

Participants reported similar confidence in their answers across all conditions (Figure 5). They were overall highly confident, with a mean rating of 5.9 to 6.1 on a 7-point Likert scale, depending on the condition. Participants' reported familiarity with fitness clubs varied, but they were overall rather familiar (Figure 5).

We computed *combined attraction effects*, shown as purple dots and error bars in Figure 4. A combined attraction effect is the sum of the attraction effects obtained in both decoy conditions, or equivalently, the difference in choice proportions between these two conditions (i.e., the differences between the bars marked "decoy" in Figure 3). This combined measure generally yields more statistical power and facilitates comparisons of results since some experiments (e.g., [54] and our next experiment) do not include a no-decoy condition and thus only report combined attraction effects.

The two purple error bars in Figure 4-Gyms show that the data overall speaks in favour of an attraction effect, both for the table and the scatterplot. To better quantify the strength of evidence, we conducted a Bayesian analysis using the Jeffreys prior for proportions [7]. Ignoring previous studies and considering our data only, the presence of a combined attraction effect in the *table* condition is 34 times more likely than a practically null effect (set to ±1%), and 11 times more likely than a "repulsion" effect. In the *scatterplot*, a combined attraction effect is 150 times more likely than a practically null effect, and 66 times more likely than a repulsion effect.

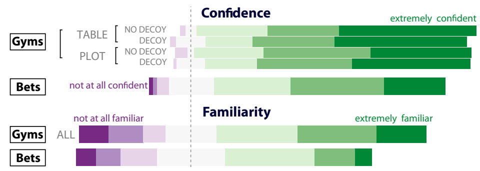


Fig. 5: Self-reported confidence and familiarity in both experiments.

3.3.3 Discussion

We found evidence for an attraction effect on *table* for the *decoy on variety* condition, but not for the *decoy on cleanliness* condition, where the effect may be smaller or even possibly negative (see Figure 4). We do not have an explanation for this asymmetry, but the wide confidence intervals and their large overlap suggests that the difference may be due to a large extent to statistical noise [12].

Based on the combined attraction effect which is a more holistic measure with more statistical power, we replicated the attraction effect on tables (**H1**) but the results are less strong than in the initial study [33] (i.e., about half of the original study, as shown by the purple CIs in Figure 4). It is common for a replication to yield smaller effect sizes [39], but the differences in results could also be due to modifications we made to the original experiment design.

We produced four different stimuli for each decision task in order to eliminate possible presentation order effects for alternatives and attributes, whereas Malkoc et al. [33] used a unique table. The use of different stimuli could have yielded a higher variability in responses.

Our study was also a crowdsource experiment, whereas Malkoc et al. conducted theirs with students in a lab, where participants are less diverse and generally more focused [34]. Perhaps the feeling of being evaluated was also stronger for students, which we know can amplify the attraction effect [47]. Our rejection criteria (e.g., textual justification for the answer, table and scatterplot test, attention test) could have also filtered subsets of the population that are more vulnerable to the effect. Finally, our participants were on average rather familiar with gyms (Figure 5), and 11% were unsure if they suffered from bacteriophobia, and we know that familiarity with the subject matter and strong prior preferences can reduce the effect [36, 41]. Malkoc et al. [33] do not report on familiarity and prior preferences.

Despite mixed results for the table condition, we obtained good evidence for an attraction effect in the *scatterplot* condition. There still appears to be an asymmetry between the two decoy conditions (this time, in the opposite direction), but CIs show no evidence for a difference. The combined attraction effect provides compelling evidence that the attraction effect can generalize to scatterplots (**H2**). This observed shift in preference after adding an irrelevant option to a two-point scatterplot gives credence to the idea that people may make irrational decisions even when they use visualizations as decision making aids. Thus we decided to explore the effect further, using scatterplots with larger sets of alternatives.

4 Extending the Attraction Effect

Our gym experiment confirmed that the attraction effect can extend to scatterplot formats. However, we have so far only considered three data points, which does not capture most real-word decision tasks where visualizations would be used.

Previous work has focused on only three alternatives because in numeric tables, it is hard to perform rapid attribute-to-attribute comparisons and recognize dominance relationships between many points. Bettman et al. [6] point out that the attraction effect requires asymmetric dominance relationships to be "perceptual in nature" and "easy to access". They expect that the bias will be eliminated with multiple alternatives, as the number of pairwise comparisons increases and these relationships become harder to understand. This may be true for numerical tables, but not necessarily for visualizations such as scatterplots, that are designed to aid viewers read and understand complex data, and support comparison of many data points at once [37]. It is thus plausible that visualizations of many alternatives can also elicit attraction effects.

4.1 Ways of Adding More Alternatives

There are three ways the classical attraction effect procedure can be extended to include more than three alternatives:

1. By adding more non-dominated options. In our introduction example, the only non-dominated alternatives were Bob and Alice. We could add more candidates that neither dominate nor are dominated by Bob and Alice. The set of formally uncomparable or non-dominated alternatives is also called the *Pareto front*.
2. By adding more decoys. In our example the only decoy is Eve. We could however add more decoys similar to Eve.
3. By adding "distractors", i.e., irrelevant options that play neither the role of target, of competitor, or of decoy. An example would be a dominated candidate that appears both in the baseline condition and in the decoy condition.

The first approach is problematic in at least two respects. One is that since it breaks the dichotomy between target and competitor, it would require a major change in the way the attraction effect is measured in experiments. A second problem is that it would cause the attraction effect to interfere with other cognitive biases. For example, the *compromise effect* is a bias by which if presented with several formally uncomparable alternatives, people tend to avoid extremes and choose options in the middle [47]. Even though it could be informative to study how the two effects may combine, we decided here to focus on the attraction effect only.

Adding an arbitrary number of distractors (option 3) is however possible. With many distractors a single decoy is unlikely to produce a measurable effect, but more decoys can be added (option 2). The Pareto front however still needs to consist of only two alternatives – a target and a competitor. We present an extension of the attraction effect procedure using this approach.

4.2 Extended Procedure

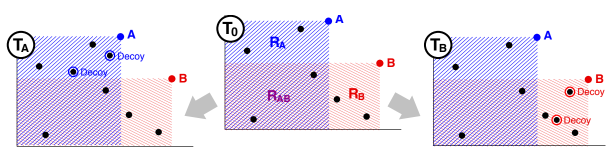


Fig. 6: A baseline decision task T_0 and two possible test decision tasks: T_A, where A is the target, and T_B, where B is the target.

The procedure consists of starting with a **baseline decision task** T_0 (see middle of Figure 6 for an example). This baseline decision task has two non-dominated alternatives, A and B. All other alternatives are dominated by A and/or B, and are called **distractors**.

For convenience, we divide the space of all possible alternatives into three dominance regions, shown in Figure 6. If d_A is the region dominated by A (blue hatches in the Figure) and d_B is the region dominated by B (red hatches), then $R_A = d_A \setminus d_B$ (region dominated by A but not by B), $R_B = d_B \setminus d_A$ (region dominated by B but not by A), and $R_{AB} = d_A \cap d_B$ (region dominated by both). In the figure, the baseline decision task contains two distractors per region.

From the baseline decision task one can derive two types of test decision tasks, labeled T_A and T_B in the figure. The decision task T_A is created by adding extra alternatives to the region R_A. Thus T_A only differs from T_0 in that it contains more alternatives that are dominated by A but not by B (twice as many, in this example). These *extra* asymmetrically dominated alternatives are referred to as **decoys**, while A is called the **target** and B the **competitor**. Similarly, the task T_B is created by adding extra decoys to the region R_B, and this time B is the target and A is the competitor.

In case no distractor is included in T_0 and a single decoy is added to T_A and to T_B, we obtain a classical attraction effect experiment. Thus our new definitions for decoy, target and competitor are consistent with the definitions from Section 2.2.1 and generalize them to more complex cases. However, decoys, targets and competitors are always defined with respect to a baseline decision task.

5 Bet Experiment: Scatterplot, Many Choices

We expand our study of the attraction effect to situations where participants are presented with scatterplots with multiple alternatives. We conducted another experiment prior to the one reported here, with a different design and inconclusive results. This inconclusive experiment is reported in a separate research report [14].

5.1 Design Rationale

Here we describe and motivate the design of this new experiment, highlighting the differences with the first (gym) experiment.

5.1.1 Replicated Study

Most attraction effect studies (including our previous experiment) follow a between-subjects design. However, these designs typically suffer from low statistical power. The width of confidence intervals in our gym experiment indicates this was the case there.

We therefore decided to adopt a within-subjects design. Wedell [54] was able to measure a clear attraction effect with numerical tables using a within-subjects procedure, where participants were given multiple decision tasks. He further tried to increase statistical power by *i)* excluding no-decoy conditions and only measuring the combined decoy effect, and *ii)* choosing a scenario with which people were less familiar in an attempt to amplify the effect [36, 41]. We therefore decided to replicate Wedell's design.

5.1.2 Scenario and Attribute Values

Wedell's scenario involved choosing among three lottery tickets, each defined by two attributes: the probability of winning (*probability*), and the amount that can be won (*prize*). Participants were presented with twenty decision tasks in sequence. Each time, three lottery tickets were presented and participants had to choose one. Wedell thought that the abstract nature of the task and of the attributes would reduce possible carry-over effects, such as participants building up strategies based on past choices.

Table 2: The non-dominated alternatives used in our tasks.

	A	B	C	D	E	(F)
probability	0.83	0.67	0.5	0.4	0.3	0.25
prize	$12	$15	$20	$25	$33	$40

The non-dominated alternatives (targets and competitors) used in all Wedell's tasks were taken from a pool of five alternatives (A to E in Table 2). All had the same expected value of ~$10. Thus, though a rational decision maker would only need to compare alternatives along a single dimension (expected value), the decision tasks had the same dominance structure as tasks involving two independent attributes such as in the previous gym experiment.

For each possible pair of alternatives in (A,B,C,D,E) Wedell generated two decision tasks, one with a decoy on probability, and one with a decoy on prize. We use the notation XY to refer to a task where X is the target and Y is the competitor, and refer to the two decision tasks XY and YX as *matched*. For example, the pair of alternatives (A,C) yields the two matched tasks AC (where the decoy is on A) and CA (where the decoy is on C). Wedell's design resulted in 10 pairs of matched decisions tasks (20 tasks in total).

Although we planned to reuse the same targets and competitors, it appeared that the distance between the target and the competitor was visually very small in some scatterplots compared to others. Thus we added an alternative with the same expected value (F in Table 2) and excluded all tasks that involved adjacent target/competitor pairs (e.g., AB, or DE). This new design also resulted in 10 pairs of matched decisions tasks, and 20 tasks in total.

5.1.3 Adding Distractors and Decoys

While Wedell only added one decoy to each of the decision tasks, our goal was to present many alternatives as explained in the previous section. For each pair of matched decision tasks, the procedure consisted of two steps. We explain the procedure for AC and CA (see results in Figure 7), but it is the same for all other pairs:

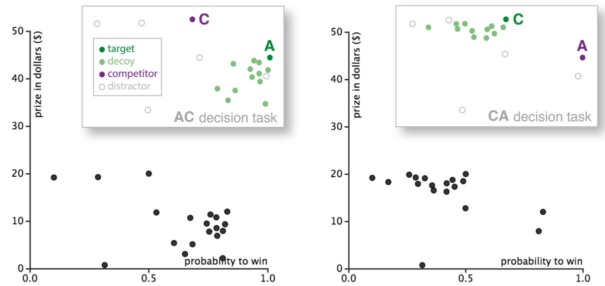


Fig. 7: Experimental stimuli for the two matched decision tasks AC and CA (black-and-white background images), and explanatory annotations (box overlays). See Section 4.2 for the full details.

Step 1. A baseline decision task analogous to T_0 in Figure 6 was created by adding distractors dominated by A and/or C. One or two distractors (number randomly drawn) were added in each of the regions R_A, R_C and R_{AC}, following a uniform spatial distribution.

Step 2. Two separate decision tasks AC and CA were then created by adding decoys as shown in Figure 6. For the task AC (decoys on A), 10 to 20 decoys (number randomly drawn) were added to the region R_A following a bivariate half-normal probability distribution. On each axis, the mode of the half-normal was A's value on this axis, and the mean was this value multiplied by 0.7. The use of half-normals yielded decoys that tend to cluster near A, but whose density smoothly decreases with distance to A for a more natural look. The same was done for the decision task CA.

In both steps, overlaps were eliminated by *i)* defining overlap between two alternatives as a distance less than 0.025 in normalized coordinates (prize divided by 40, probability left unchanged) and *ii)* whenever a new alternative is randomly drawn, iterating until there is no overlap. The reason why the number of alternatives to draw was randomized (i.e., 1–2 for each region in Step 1 and 10–20 in Step 2) was to create more variation across scatterplots and make it more difficult for participants to infer patterns in the experiment.

5.1.4 Ordering of Decision Tasks

Our presentation order for the 20 decision tasks was inspired from Wedell [54], but modified to account for our different set of tasks and for the fact that we present each task on a separate Web page, while Wedell used a four-page paper-and-pencil test.

We created a task ordering such that *i)* a decision task and its matched task (e.g., AC and CA) are always at least 5 pages apart; and *ii)* the role of an alternative alternates over time. For example, if D appears as a target in a task, it will be a competitor the next time it appears. To reduce further possible ordering effects, we created a second ordering where each task is replaced with its matched task. Participants were randomly assigned to each ordering.

To make it more difficult for participants to infer patterns in the sequence of decision tasks, we additionally inserted seven irrelevant decision tasks at various positions, which were not used in our analyses. These tasks differed in that they had either one or three non-dominated alternatives (instead of two), and they did not exhibit an imbalance in the number of asymmetrically dominated alternatives.

5.1.5 Stimuli: Interactive Scatterplots

In this experiment, we added minimal interaction to the scatterplot visualizations. In the first experiment, the scatterplots were static and each data point was labeled with a letter (Figure 1), so that participants could specify their choice through separate radio buttons. As we are now dealing with more data points, labels were removed to prevent clutter (Figure 7), and participants were asked to specify their choice by selecting the data point. Points were highlighted when hovered. Hovering a point also displayed horizontal and vertical projection lines, and the data point's X and Y values were overlaid on the axes. Such interactions help examine the data and are not uncommon in scatterplot visualizations. After a point was clicked, its color changed and the participant was asked to confirm her choice by clicking on a button at the bottom of the page.

We added a short flicker during task transitions in order to elicit change blindness and prevent participants from easily detecting similarities and differences between two successive scatterplots.

5.1.6 Crowdsource Quality Control

We made two major modifications to the previous procedure: *i)* we added a preliminary tutorial, *ii)* we used a real decision making task where choices affected subsequent monetary gains.

The tutorial simultaneously explained the scenario (the lottery tickets, and what their probability and prize meant), and how to read scatterplots. Although Wedell [54] did not provide similar training, crowdsource contributors do not necessarily have the same qualifications as university students, and the notion of probability in particular is known to be challenging [34]. In order to prime participants to use their intuition rather than doing calculations, probability was explained qualitatively rather than quantitatively.

After the tutorial, participants were given a test question consisting of choosing one among 13 lottery tickets presented as a scatterplot. Three tickets were non-dominated (and thus formally uncomparable), and the remaining 10 were considered wrong answers.

In order to better approximate real-life decisions and motivate our participants, we informed them that a computer will run the lottery after the experiment is completed, and for every winning ticket they picked, they will be payed a bonus proportional to the ticket's prize. The use of a real decision task with consequences is common in behavioral economics and is occasionally used when studying the attraction effect (e.g., choosing between objects or money [49]).

Similarly to our previous experiment, we defined our rejection criteria in advance and categorized jobs as *Red* (rejected and not payed), *Orange* (payed but not analyzed) and *Green* (analyzed).

A total of 120 jobs were submitted with a valid completion code. A job was marked as *Red* (12%) if its completion time was abnormal (0.8%), if the contributor failed the tutorial test (11%), or if during the experimental trials, the contributor selected a dominated option more than half of the time (12%). A job was marked as *Orange* (27%) if the contributor always chose the highest probability (27%) or the highest prize (0%). These contributors had a too strong prior preference (in this case, risk aversion) to be sensitive to the attraction effect. The remaining 61% (*N*=73) were marked as *Green*.

5.2 Experiment Design

The design consisted of two within-subjects factors: *task pair* (10 pairs of matched tasks), and *decoy position* (on *probability* or *prize*).

5.2.1 Procedure

We first briefed our crowdsource contributors that they will have to choose lottery tickets and will receive a bonus for each winning ticket, for a total of $0.60 on average. They then opened an external link to the 10-page tutorial. Contributors who chose a valid ticket on the test were told that the ticket won, and that they would get a $0.10 bonus for the ticket if they proceed and complete the job.

Participants then opened a second external link to the main study, a 31-page form, where they saw the twenty decision tasks, mixed with the seven distractor tasks. After completing all decision tasks, participants rated their overall confidence, their perceived difficulty of the job, their familiarity with gambling games, and whether they knew of the notion of "expected value" in probability. They then filled a short demographic questionnaire.

Finally, participants were presented again with one of the study's scatterplots with the target and competitor labeled *A* and *B*, and were asked whether the higher number of tickets near *A* affected their choices, why they thought there were more tickets, and whether they had this explanation in mind during the study.

All participants received a baseline payment of $0.20, while *Orange* and *Green* received a bonus of $0.10 plus a lottery bonus. The expected lottery bonus was $0.50 if no dominated alternative was chosen, based on a conversion rate of 0.0025 between the scenario's "virtual dollars" and USD. After the experiment was over, we determined each lottery bonus by *i)* running Bernouilli random draws to determine the winning status of each chosen ticket, *ii)* summing up the prizes of winning tickets *iii)* multiplying by the conversion rate.

5.2.2 Participants

Our participants were 73 crowdflower contributors whose job was marked *Green*. Their demographics, shown in Figure 2, were similar to the first experiment.

5.2.3 Hypotheses

Our statistical hypothesis was **H3**: the mean *attraction score* (as defined in the next section) will be strictly positive.

5.3 Results

One planned analysis for assessing the consistency of participants' responses within matched tasks is not reported for space reasons. We report on all other planned analyses.

5.3.1 Planned Analyses

We first report descriptive statistics of participant choices in a similar way to Wedell [54]. We recorded a total of 1460 choices (73 participants × 20 decision tasks). We pair choices according to matched tasks (e.g., tasks *AC* and *CA* in Figure 7), yielding 73 × 10 = 730 choice pairs. Of all these choice pairs, only 24 (3.3%) included a dominated alternative. Wedell reports similar results (2%), even though his tasks only involved a single dominated alternative.

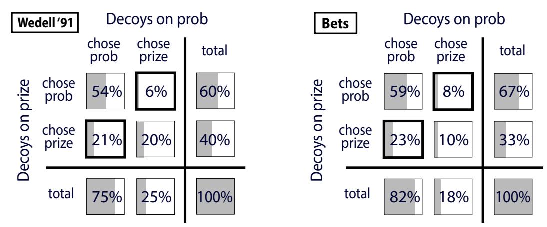


Fig. 8: Contingency tables showing choice pairs for all matched tasks

Figure 8 (right) summarizes the remaining 706 choice pairs as a contingency table, shown next to Wedell's (on the left). Choice pairs fall into four categories. One is choosing the ticket with highest probability in both tasks (i.e., ticket *A* in Figure 7). This represents 59% of all choice pairs, and is reported in the top-left cell in Figure 8. A second possibility is choosing the ticket with highest prize twice, which represents 10% of all cases. The remaining two possibilities, shown in bold cells, consist in always choosing the target (23%), or always choosing the competitor (8%).

The patterns in our contingency table follow Wedell's closely [54]: participants favoured higher probability overall (reflecting again risk aversion), but when their choice was inconsistent across two matched tasks, they chose the targets more often than they chose the competitors. We now turn to inferential statistics.

Similarly to Wedell, we used as dependent variable an *attraction score*, calculated on a per-participant basis as follows. Each of the 20 decision tasks was assigned a score of 1 when the ticket with highest probability was chosen, a score of 0 when the ticket with highest prize was chosen, and a score of 0.5 when another (dominated) ticket was chosen. Then, we averaged all scores for the 10 decision tasks where the decoys were on probability (yielding a score S_{prob}) and did the same for the 10 tasks where the decoys were on prize (yielding a score S_{prize}). The difference between the two scores $S = S_{\text{prob}} - S_{\text{prize}}$ was the *attraction score*.

A participant who is not subject to the attraction effect should exhibit the same preference for high probability irrespective of the position of the decoys, thus her attraction score should be close to zero. We multiplied the attraction score by 100 to obtain a percentage analogous to the combined decoy effect reported in the gym experiment. The difference here is that the percent difference is computed within-subjects instead of between-subjects, and it incorporates choices of dominated options as "neutral" observations.

The mean attraction score was 15%, with a 95% bootstrap confidence interval of [10%, 22%] (see Figure 4). Thus we have very solid evidence for **H3**, even if the effect is smaller than in Malkoc's gym study [33]. We cannot directly compare our effects with Wedell's [54] due to the use of different statistical methods, but Figure 8 suggests the effect sizes are comparable.

5.3.2 Additional Analyses

As shown in Figure 5, participants reported various levels of familiarity with gambling and were confident in their choice overall, although slightly less than in the gym experiment. Data on participants' knowledge of expected values was missing due to a bug.

Concerning the final questionnaire on how participants interpreted the presence of decoys (see Section 5.2.1), 8 participants reported not being able to see the scatterplot image, leaving data from 65 participants. When asked whether the uneven distribution of tickets affected their choices, 41% replied "never" or "rarely", 46% replied "sometimes", 12% replied "often", and none replied "always". When asked why they thought there were more tickets in one region than the other, most (86%) gave responses that were irrelevant or unintelligible based on an informal content analysis of open text responses. Out of the 9 remaining responses, 5 referred to a strategy employed by the lottery organizer (e.g., *"To tempt people to choose tickets of high prize but with low probability, increasing the profitability of lottery owner"*; *"To distract from choosing the higher chances of winning"*), and 4 referred to tickets as past choices from other players (e.g., *"Customers want to win a higher prize"*; *"Maybe more people played the same"*). Only 4 participants (quoted here) reported that they had their explanation in mind while performing the task, while the other 5 reported that it was prompted by our question. Thus there is little evidence that participants' preference for the target was motivated by deliberate, reasoned strategies.

6 General Discusion and Conclusions

Taken together, our two experiments suggest that the attraction effect generalizes to data visualizations. While the first experiment focuses on a traditional procedure with only two or three alternatives, the second experiment shows that the effect can persist with more alternatives. Bettman et al [6] expected that the effect would disappear as more alternatives are added, since pairwise comparisons and dominance recognition becomes hard if numerical tables are used. Our findings suggest that this may not be the case when using visualizations, as visualizations such as scatterplots support fast comparisons and dominance recognition. Overall, our study indicates that when people visualize choice alternatives using scatterplots, the number and position of irrelevant (dominated) alternatives may influence their choice. This shift in preferences violates basic axioms of rational choice theory [25]. In addition to being the first infovis study on the attraction effect, our work contributes to the ongoing debate in decision-making research on whether the effect generalizes to non-numerical formats [17, 26, 48, 56].

6.1 Implications for Design

On a general level, our study indicates that cognitive biases can affect decisions even if the data is well visualized and fully understood, thus traditional visualization design rules may not apply when the goal is to support decision making. This article has not considered debiasing techniques for the attraction effect, but a simple way to eliminate the bias would be to only show the Pareto front, i.e, to hide all dominated options. However, this approach assumes that the system has full knowledge of the user's choice criteria, which may not be the case in practice. In addition, dominated options can help understand dataset trends, and may in some cases provide useful context when making decisons. Thus, debiasing techniques should only be available as options, and activated on demand. Alternatively, one could consider techniques such as de-emphasizing dominated options or highlighting the Pareto front, but the effectiveness of such techniques remains to be experimentally tested.

6.2 Limitations

There are several potential limitations to our study. One stems from a general criticism of cognitive bias research, namely, that heuristics that appear irrational may not be so upon deeper examination [18]. Concerning the attraction effect, the way dominated alternatives are distributed could in some cases provide relevant information. For example, a real estate investor may infer from a region with many dominated alternatives that a certain type of house is more common, and therefore represents a larger market. At the same time, situations also exist where the number and position of dominated alternatives is clearly irrelevant and where a preference for the target would be irrational. This was the case for our experiment involving real bets, and our data does indicate that the vast majority of our participants were unable to rationalize their choices based on where the dominated alternatives were located.

Although we have observed attraction effects, we did not investigate why they occur. In particular, we do not know how much of the effect has cognitive vs. perceptual causes. Since in the bet experiment regions with many decoys were visually more salient, it is possible that they drew participants' attention towards the target, or similarly, that participants sometimes failed to see the competitor because it was an isolated point. This possibility does not invalidate the existence of an attraction effect (as defined in Sections 2.2.1 and 4.2), but it does raise the possibility that part of the effect with scatterplots (but not with numerical tables) has perceptual origins.

Finally, we tested very specific datasets, i.e., synthetically-generated datasets with only two non-dominated options and a large number of decoys. More realistic datasets need to be tested, although our inconclusive results with real datasets suggest that the effects may be small and hard to measure [14].

6.3 Future work

While our work is a first step in investigating the attraction effect in visualizations, much more work is needed. More realistic datasets and decision making situations remain to be tested. We also focused on scatterplots, but clearly other commonly used visualizations need to be evaluated to assess whether the effect persists across visual encodings. Other cognitive biases [19] remain to be studied, both in isolation (as we did here), and in combination. How cognitive biases interact with visual perception is also an important and difficult question that has remained largely unexplored.

Acknowledgments

We thank S.A. Malkoc for sharing their material and D.H. Wedell for their study advice, our colleagues J.-D. Fekete, F. Vernier, T. Isenberg, P. Goffin and L. Besancon for feedback on the paper, and our reviewers for their insightful comments and suggestions.

References

[1] D. Ariely and T. S. Wallsten. Seeking subjective dominance in multidimensional space: An explanation of the asymmetric dominance effect. *Organizational Behavior and Human Decision Processes*, 63(3):223 – 232, 1995.

[2] T. Asahi, D. Turo, and B. Shneiderman. Using treemaps to visualize the analytic hierarchy process. *Information Systems Research*, 6(4):357–375, 1995.

[3] B. A. Aseniero, T. Wun, D. Ledo, G. Ruhe, A. Tang, and S. Carpendale. Stratos: Using visualization to support decisions in strategic software release planning. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 1479–1488. ACM, 2015.

[4] M. Bateson, S. D. Healy, and Hurly. Irrational choices in hummingbird foraging behaviour. *Animal Behaviour*, 63(3):587–596, Mar. 2002.

[5] J. Bautista and G. Carenini. An empirical evaluation of interactive visualizations for preferential choice. In *Proceedings of the working conference on Advanced visual interfaces*, pages 207–214. ACM, 2008.

[6] J. R. Bettman, M. F. Luce, and J. W. Payne. Constructive Consumer Choice Processes. *J. of Consumer Research*, 25(3):187–217, Dec. 1998.

[7] W. M. Bolstad. *Introduction to Bayesian statistics*. J. Wiley & Sons, 2013.

[8] J. Boy, R. Rensink, E. Bertini, and J.-D. Fekete. A principled way of assessing visualization literacy. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):1963–1972, Dec 2014.

[9] S. K. Card, J. D. Mackinlay, and B. Shneiderman, editors. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999.

[10] J. M. Choplin and J. E. Hummel. Comparison-induced decoy effects. *Memory & Cognition*, 33(2):332–343, 2005.

[11] M. Correll and M. Gleicher. Error bars considered harmful: Exploring alternate encodings for mean and error. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):2142–2151, 2014.

[12] G. Cumming and S. Finch. Inference by eye: confidence intervals and how to read pictures of data. *American Psychologist*, 60(2):170, 2005.

[13] M. Daradkeh, C. Churcher, and A. McKinnon. Supporting informed decision-making under uncertainty and risk through interactive visualisation. In *Proc. Fourteenth Australasian User Interface Conference-Vol 139*, pages 23–32. Australian Computer Society., 2013.

[14] E. Dimara, A. Bezerianos, and P. Dragicevic. Testing the Attraction Effect on Two Information Visualization Datasets. Research Report RR-8895, Inria, Mar. 2016. `http://hal.inria.fr/hal-01295624`.

[15] P. Dragicevic. Fair statistical communication in HCI. In *Modern Statistical Methods for HCI*, pages 291–330. Springer, 2016.

[16] N. Elmqvist, P. Dragicevic, and J.-D. Fekete. Rolling the Dice: Multidimensional Visual Exploration using Scatterplot Matrix Navigation. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1141–1148, Nov. 2008. Best Paper Award.

[17] S. Frederick, L. Lee, and E. Baskin. The limits of attraction. *Journal of Marketing Research*, 51(4):487–507, 2014.

[18] G. Gigerenzer. Why heuristics work. *Perspectives on psychological science*, 3(1):20–29, 2008.

[19] M. L. Graber, S. Kissam, V. L. Payne, A. N. Meyer, A. Sorensen, N. Lenfestey, E. Tant, K. Henriksen, K. LaBresh, and H. Singh. Cognitive interventions to reduce diagnostic error: a narrative review. *BMJ Quality & Safety*, pages bmjqs–2011, 2012.

[20] T. B. Heath and S. Chatterjee. Asymmetric decoy effects on lower-quality versus higher-quality brands: Meta-analytic and experimental evidence. *Journal of Consumer Research*, 22(3):268–284, 1995.

[21] W. Hedgcock and A. R. Rao. Trade-off aversion as an explanation for the attraction effect: A functional magnetic resonance imaging study. *Journal of Marketing Research*, 46(1):1–13, 2009.

[22] W. Hedgcock, A. R. Rao, and H. A. Chen. Could ralph nader's entrance and exit have helped al gore? the impact of decoy dynamics on consumer choice. *Journal of Marketing Research*, 46(3):330–343, 2009.

[23] J. Heer and M. Bostock. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *Proceedings of CHI 2010*, pages 203–212. ACM, 2010.

[24] S. Highhouse. Context-dependent selection: The effects of decoy and phantom job candidates. *Organizational Behavior and Human Decision Processes*, 65(1):68 – 76, 1996.

[25] J. Huber, J. W. Payne, and C. Puto. Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, 9(1):90–98, 1982.

[26] J. Huber, J. W. Payne, and C. P. Puto. Let's be honest about the attraction effect. *Journal of Marketing Research*, 51(4):520–525, 2014.

[27] J. Huber and C. Puto. Market boundaries and product choice: Illustrating attraction and substitution effects. *Journal of Consumer Research*, pages 31–44, 1983.

[28] D. Kahneman. *Thinking, fast and slow*. Macmillan, 2011.

[29] A. Khan, S. Breslav, M. Glueck, and K. Hornbæk. Benefits of visualization in the mammography problem. *International Journal of Human-Computer Studies*, 83:94–113, 2015.

[30] U. Khan, M. Zhu, and A. Kalra. When trade-offs matter: The effect of choice construal on context effects. *Journal of Marketing Research*, 48(1):62–71, 2011.

[31] T. Latty and M. Beekman. Irrational decision-making in an amoeboid organism: transitivity and context-dependent preferences. *Proceedings of the Royal Society B: Biological Sciences*, 278(1703):307, 2011.

[32] S. Lelis and A. Howes. Informing decisions: how people use online rating information to make choices. In *Proceedings of CHI 2011*, pages 2285–2294. ACM, 2011.

[33] S. A. Malkoc, W. Hedgcock, and S. Hoeffler. Between a rock and a hard place: The failure of the attraction effect among unattractive alternatives. *Journal of Consumer Psychology*, 23(3):317 – 329, 2013.

[34] L. Micallef, P. Dragicevic, and J.-D. Fekete. Assessing the effect of visualizations on bayesian reasoning through crowdsourcing. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2536–2545, 2012.

[35] S. Miller, A. Kirlik, A. Kosorukoff, and J. Tsai. Supporting joint human-computer judgment under uncertainty. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 52, pages 408–412. SAGE Publications, 2008.

[36] S. Mishra, U. Umesh, and D. E. Stem Jr. Antecedents of the attraction effect: An information-processing approach. *Journal of Marketing Research*, pages 331–349, 1993.

[37] T. Munzner. *Visualization Analysis and Design*. CRC Press, 2014.

[38] Y. P. S. O'Curry and R. Pitts. The attraction effect and political choice in two elections. *Journal of Consumer Psychology*, 4(1):85–101, 1995.

[39] Open Science Collaboration et al. Estimating the reproducibility of psychological science. *Science*, 349(6251), 2015.

[40] J. C. Pettibone. Testing the effect of time pressure on asymmetric dominance and compromise decoys in choice. *Judgment and Decision Making*, 7(4):513–523, 2012.

[41] S. Ratneshwar, A. D. Shocker, and D. W. Stewart. Toward understanding the attraction effect: The implications of product stimulus meaningfulness and familiarity. *J. of Consumer Research*, pages 520–533, 1987.

[42] S. Rudolph, A. Savikhin, and D. S. Ebert. Finvis: Applied visual analytics for personal financial planning. In *IEEE Symposium on Visual Analytics Science and Technology*, pages 195–202. IEEE, 2009.

[43] A. Savikhin, R. Maciejewski, and D. S. Ebert. Applied visual analytics for economic decision-making. In *Visual Analytics Science and Technology, 2008. VAST'08. IEEE Symposium on*, pages 107–114. IEEE, 2008.

[44] S. Sen. Knowledge, information mode, and the attraction effect. *Journal of Consumer Research*, 25(1):64–77, 1998.

[45] S. Shafir, T. A. Waite, and B. H. Smith. Context-Dependent Violations of Rational Choice in Honeybees (Apis mellifera) and Gray Jays (Perisoreus canadensis). *Behavioral Ecology and Sociobiology*, 51(2):180–187, 2002.

[46] A. Shen-Hsieh and M. Schindl. Data visualization for strategic decision making. In *Case Studies of the CHI2002*, pages 1–17. ACM, 2002.

[47] I. Simonson. Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, 16(2):158–174, 1989.

[48] I. Simonson. Vices and virtues of misguided replications: The case of asymmetric dominance. *J. of Marketing Research*, 51(4):514–519, 2014.

[49] I. Simonson and A. Tversky. Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research*, 29:281–295, 1992.

[50] J. J. Thomas and K. A. Cook. *Illuminating the path:the research and development agenda for visual analytics*. IEEE Computer Society, 2005.

[51] J. Trueblood. Multialternative context effects obtained using an inference task. *Psychonomic Bulletin and Review*, 19(5):962–968, 2012.

[52] J. S. Trueblood, S. D. Brown, A. Heathcote, and J. R. Busemeyer. Not just for consumers context effects are fundamental to decision making. *Psychological science*, 24(6):901–908, 2013.

[53] C. Ware. *Information visualization: perception for design*. Elsevier, 2012.

[54] D. H. Wedell. Distinguishing among models of contextually induced preference reversals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(4):767, 1991.

[55] C. Williamson and B. Shneiderman. The dynamic homefinder: Evaluating dynamic queries in a real-estate information exploration system. In *Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '92, pages 338–346, New York, NY, USA, 1992. ACM.

[56] S. Yang and M. Lynn. More evidence challenging the robustness and usefulness of the attraction effect. *Journal of Marketing Research*, 51(4):508–513, 2014.

[57] J. S. Yi. *Visualized decision making: development and application of information visualization techniques to improve decision quality of nursing home choice*. PhD thesis, Georgia Institute of Technology, 2008.

[58] J. S. Yi, R. Melton, J. Stasko, and J. A. Jacko. Dust & magnet: multivariate information visualization using a magnet metaphor. *Information Visualization*, 4(4):239–256, 2005.

[59] Y. Zhang, R. K. Bellamy, and W. A. Kellogg. Designing information for remediating cognitive biases in decision-making. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2211–2220. ACM, 2015.

[60] C. Ziemkiewicz and R. Kosara. Laws of attraction: From perceptual forces to conceptual similarity. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6):1009–1016, 2010.

[61] T. Zuk and S. Carpendale. Visualization of uncertainty and reasoning. In *Smart Graphics*, pages 164–177. Springer, 2007.

Eurographics Conference on Visualization (EuroVis) 2013
B. Preim, P. Rheingans, and H. Theisel
(Guest Editors)
Volume 32 (2013), Number 3

Evolutionary Visual Exploration: Evaluation With Expert Users

N. Boukhelifa[1], W. Cancino[1], A. Bezerianos[2,1] and E. Lutton[1,3]

[1]INRIA Saclay - Île-de-France, France [2]Univ Paris-Sud & CNRS, Orsay, France [3]INRA, Grignon, France

Abstract

We present an Evolutionary Visual Exploration (EVE) *system that combines visual analytics with stochastic optimisation to aid the exploration of multidimensional datasets characterised by a large number of possible views or projections. Starting from dimensions whose values are automatically calculated by a PCA, an interactive evolutionary algorithm progressively builds (or evolves) non-trivial viewpoints in the form of linear and non-linear dimension combinations, to help users discover new interesting views and relationships in their data. The criteria for evolving new dimensions is not known a priori and are partially specified by the user via an interactive interface: (i) The user selects views with meaningful or interesting visual patterns and provides a satisfaction score. (ii) The system calibrates a fitness function (optimised by the evolutionary algorithm) to take into account the user input, and then calculates new views. Our method leverages automatic tools to detect interesting visual features and human interpretation to derive meaning, validate the findings and guide the exploration without having to grasp advanced statistical concepts. To validate our method, we built a prototype tool (EvoGraphDice) as an extension of an existing scatterplot matrix inspection tool, and conducted an observational study with five domain experts. Our results show that EvoGraphDice can help users quantify qualitative hypotheses and try out different scenarios to dynamically transform their data. Importantly, it allowed our experts to think laterally, better formulate their research questions and build new hypotheses for further investigation.*

1. Introduction

The purpose of visual exploration is to find meaningful patterns in the data which can lead to insight. In a high-dimensionality context, this task becomes rather challenging as viewers may be faced with a large space of alternative views on the data. One way to help navigate such a space is the "grand tour" method [Asi85] which offers a complete view of the search space in a smooth sequence of projections showing various viewpoints of the data. However, the time required to inspect all these views may be prohibitive [Hub85]. A related approach that improves on this is "projection pursuit" [Fri87] where the aim is to visit only the most interesting views; interesting referring to projections that deviate more from a normal distribution. The criteria for deciding whether a projection is interesting have mostly been defined prior to user exploration, using objective measures such as the quality metrics surveyed in [BTK11].

We present a novel visual analysis tool to explore multi-dimensional datasets where the system proposes interesting views based on both objective measures, such as different visual patterns in the two-dimensional projections of the data, and subjective measures corresponding to user satisfaction with the presented view. These subjective measures are not known prior to user exploration. To demonstrate our ideas, we built a prototype (*EvoGraphDice*) as an extension of an existing scatterplot matrix inspection tool. We use low dimension projection to handle data multi-dimensionality, and linear and non-linear combinations of dimensions for an axis of the projection plane to propose alternative views. User exploration is guided by an Interactive Evolutionary Algorithm (IEA) which can both generate new views and adapt to user interest. Below, we provide background for the topic of evolutionary computation before listing our contributions.

Evolutionary Algorithms (EAs) are stochastic optimisation heuristics that copy, in a very abstract manner, the principles of natural evolution that let a population of individuals be adapted to its environment [Gol89]. They have the major advantage over other optimisation techniques of making only few assumptions on the function to be optimised. In short, an EA considers populations of potential so-

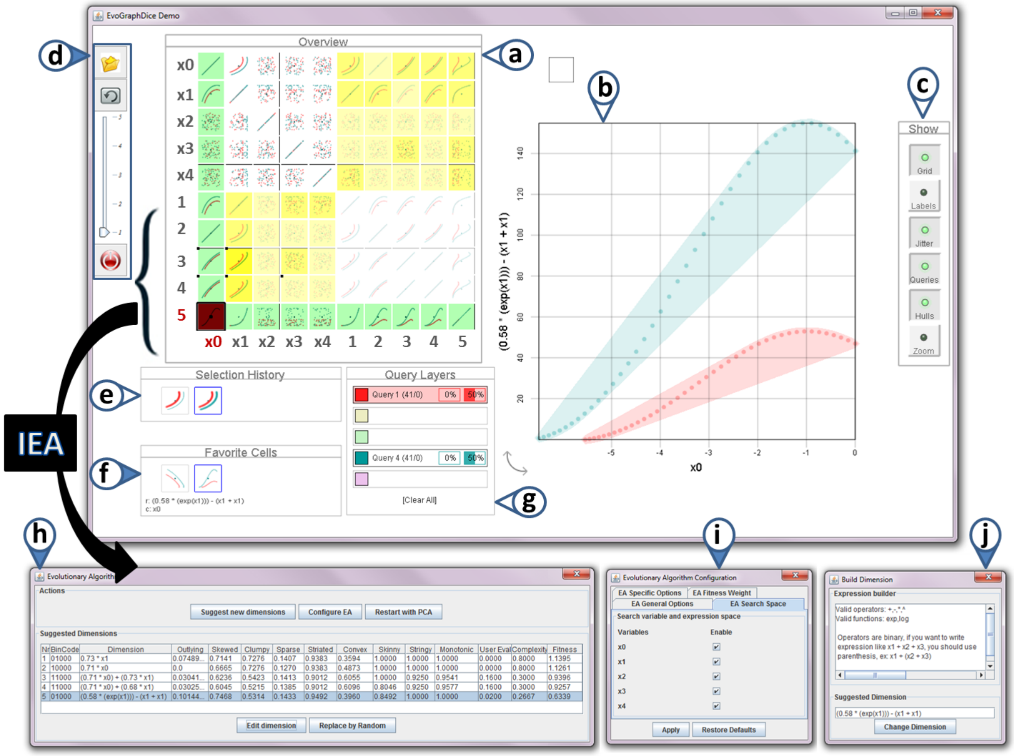


Figure 1: *EvoGraphDice prototype showing an exploration session of a synthetic dataset. New extensions to the GraphDice system are indicated by coloured label arrows. Widgets: (a) an overview scatterplot matrix showing the original data set of 5 dimensions (x0..x4) and the new dimensions (1..5) as suggested by the evolutionary algorithm. (b) main plot view. (c) tool bar for main plot view. (d) a tool bar with (top to bottom)"favorite" toggle button, "evolve" button , a slider to evaluate cells and a restart (PCA) button. (e) the selection history tool. (f) the favorite cells window. (g) the selection query window. (h) IEA main control window. (i) window to limit the search space. (j) dimension editor.*

lutions exactly like a natural population of individuals that live, fight, and reproduce, but the natural environment pressure is replaced by an "optimisation" pressure. In this way, individuals that reproduce are the best ones with respect to the problem to be solved. Reproduction consists of generating new solutions via variation schemes (the genetic operators), that, by analogy with nature, are called mutation if they involve one individual, or crossover if they involve two parent solutions. A *fitness function*, computed for each individual, is optimised by the EA. Evolutionary optimisation techniques are particularly efficient to address complex problems (irregular, discontinuous) where classical deterministic methods fail [Ban97, PLM08], but they can also deal with varying environments [JB05], or non computable quantities [Tak08]. More specifically, Interactive Evolutionary Algorithms (IEAs) are focussed on the optimisation of subjective quantities captured via a user interface.

Evolutionary Visual Exploration (EVE): we feel that Interactive Evolutionary Algorithms (IEA) are convenient for guiding the user in exploring complex datasets. This opinion is founded by the following characteristics of EAs: *(i) focus:* an IEA performs an optimisation, i.e. it drives the exploration towards "interesting" areas of the search space (areas of high fitness function and good user satisfaction), *(ii) diversity:* by nature, an IEA has a stochastic behaviour, and its population-based scheme allows to display a variety of solutions to the user at any time, *(iii) adaptation:* EAs are able to deal with time varying environments and are able to follow changes of user interest and focus [Lut06].

The contributions of this paper are: (1) a framework for Evolutionary Visual Exploration (EVE) that marries techniques from visual analysis and evolutionary computation to guide user exploration towards interesting views on the data; (2) a prototype tool (*EvoGraphDice*) [Evo] to demonstrate our framework; and (3) an observational study with five domain expert users to evaluate *EvoGraphDice*.

2. Related Work

Related work is organised as follows; (1) a brief overview of quality metrics used to describe specific properties of data projections; (2) description of quality metrics we use in this work as part of the automatic evaluation of scatterplots; and (3) a summary of related work to IEA.

Quality Metrics: faced with the overwhelming possibilities of exploration paths in multidimensional visualization, researchers in the field designed quality metrics that evaluate the various projections of the data, in the hope of focusing user search on the most promising views. In a recent survey, Bertini et al. [BTK11] used the data flow model to classify quality metrics into three types: metrics that draw information from the data space, from the image space or from both.

Amongst metrics calculated at the data space are clustering and outliers. The rank-by-feature framework [SS05], for instance, visualises an optimal set of features according to a user selected quality metric such as correlation or uniformity. They use axis-parallel projections to produce 1D or 2D views and color brightness to denote ranking scores. Amongst image based metrics are scagnostics [WW08] which describe measures of interest for pairs of dimensions based on their geometrical appearance on a scatterplot. The mixed metrics combine information from the data and image spaces at the same time. Peng et al. [PWR04], for example, combine data features such as correlation information with view features such as axes adjacency to measure clutter as a result of reordering visualization axes [BTK11].

When interaction with quality metrics is available, it is either to select a metric amongst others, or to set threshold values [BTK11]. Having to specify the type of ranking criteria requires users to be familiar with advanced statistical concepts. In our case, the quality metric is pre-defined as a vector of nine image-based measures (described in the next section) and the threshold values are adapted according to user feedback. Thus, our method leverages automatic tools to detect interesting features and human interpretation to derive meaning, validate the findings and guide the exploration without having to grasp advanced statistical concepts.

Scagnostics[†] are based on geometric graphs which are calculated from areas, perimeters and lengths of these graphs. They include nine measures to characterise scatterplots (Fig. 2) and are useful for quickly discovering regularities and anomalies in scatterplot matrices. The underlying algorithm detects different types of point distributions including multivariate normal, log normal, multinomial, sparse, dense, convex and clusters. It does so by binning, detecting outliers and computing measures based on the following three statistical properties: *shape* for convex, skinny and stringy distributions; *trend* for monotonic distributions; and

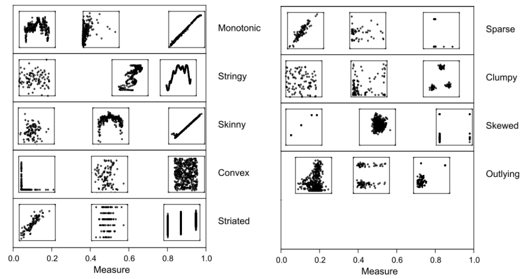


Figure 2: *Nine scagnostics measures from [WW08].*

density for skewed, clumpy, outlying, sparse and striated. These measures have proven statistical properties and are computable for moderately large data sets [WW08].

Visualization and IEA: visualization tools have been used in IEA both as representation and exploration tools to help users better evaluate the output of interactive evolutionary algorithms [HT00, LSA*06]. Despite efforts to design good user interfaces for IEA, human interaction with these systems usually raises several problems, mainly linked to the "user bottleneck" [PC97], human fatigue and slowness. Various solutions have been considered [PC97, Tak98, Ban97] such as reducing the population size (micro-EAs), constraining the search space to focus on *a priori* "interesting" areas, and deploying approximated user models (also called *surrogate functions*) to filter obvious bad solutions [LPLV05]. In the visualization community, work on parameter space exploration and optimisation relates to ours. Matkovic et al. [MGJH11], for instance, tried to interactively find an optimal combination of input parameters for a complex diesel engine injection system using visual analysis techniques. However, to our knowledge, we are the first to propose using IEA as optimisation tools to help navigate large search spaces.

3. EvoGraphDice

Since our main contribution in this work does not lie in a novel visualization system, but in enabling an IEA to guide user exploration, we used an existing visualization tool (GraphDice [EDF08, BCD*10]) to manage the various projections of the data. Views are organised in a scatterplot matrix (SPLOM) of 2D projections, Fig. 1(a). Users can do brushing and linking using a lasso tool. *EvoGraphDice* displays the dimensions proposed by the IEA as additional rows (and columns) in the SPLOM. The system initially displays dimensions returned by a PCA, after which the user can evolve new dimensions by pressing the "evolve" button, Fig. 1(d). The proposed views are displayed in yellow background; the darker the color the more interesting the view. The system provides an initial score (1 to 5) for each new view but the user can adapt this score using the slider in Fig. 1(d). User evaluated cells are flagged (small black square) to distinguish them from system evaluated

† Available as a free downloadable package in R from http://www.rforge.net/scagnostics/

cells. *EvoGraphDice* can be initialised at any time using the "restart" button which resets parameters of the IEA. Users can save views (Fig. 1(f)) and bring them back into the SPLOM if they have been replaced during the exploration.

The current population is also displayed as a table (Fig. 1(h)) where each row corresponds to a combined dimension described by a mathematical expression and various components of the fitness function such as the scagnostics measures. The user can edit an individual using the "dimension editor" in Fig. 1(j), and limit the dimension search space Fig. 1(i), which results in a system reset similar to precessing the "restart" button. Note that many EA parameters can be tuned, such as the fitness threshold and crossover/mutation/replacement rates (see [CBL12]).

Our prototype has been developed from a first version [CBL12] based on an IEA that only manipulated linear combinations of dimensions. Our new extensions are: (i) a Genetic Programming (GP) algorithm allowing the manipulation of non-linear combinations of dimensions as variable size mathematical formula, (ii) user assessment of proposed views is explicitly captured via a slider, (iii) a surrogate function based on scagnostics measurements is used to predict and simplify the interactions of the user with the IEA, (iv) color highlighting of cells is used to draw user attention to the most interesting views.

Search Space: The space searched by the evolutionary process is the set of all dimensions that can be built by combining the initial dimensions with operators and constants, encoded as trees according to the Genetic Programming (GP) framework [Koz92]. These combinations can be complex mathematical expressions containing quadratic, exponential or logarithmic terms (evolved expressions can be any combination using $+$, $-$, $*$, $/$, $(.)^{(.)}$, *exp* and *log* operators).

Genetic Engine: We have chosen to evolve a small set of combined dimensions, in order to let the user see all individuals of the population at a glance: if n is the number of initial dimensions, a population of another n combined dimensions is evolved. At each iteration, that is each time the user clicks on the "evolve" button, a new generation is produced by application of selection/crossover/mutation operators and then presented to the user whose judgment (evaluation) is explicitly collected via a slider.

Initialisation: A set of *a priori* interesting dimensions has been chosen as a starting point. A PCA analysis is performed [Smi02] on the original data and the corresponding n linear combinations form the initial population.

The fitness function, that is optimised by the genetic engine, is a sum of three terms:

1. **A surrogate function** f_{sc}, that plays the role of a predictor, and helps the system to better adapt to user needs. It is based on scagnostics measurements computed for every cell of each dimension y_i, the corresponding fitness term is a linear combination of the highest values of the scagnostics ($SC_k(y_i,x_j)$) of each scatterplot cell (y_i,x_j):
$$f_{sc}(y_i) = \sum_{k=1..9} w_k(\max_j SC_k(y_i,x_j)) \qquad (1)$$
The weights w_k that govern the relative importance of each scagnostic measurement are initialised to a uniform weight $(1/9)$. Then, as soon as enough interactions are recorded (n, the number of variables), w_k are updated via a simple multilinear regression on the m past interactions ($m \geq n$ corresponds to the length of the "memory" of the system).
2. **A Complexity term** that favours dimensions made of a small number of variables and simple mathematical expressions :
$$f_c(y_i) = \left(1 - \frac{nvars(y_i)}{n}\right) \times \frac{1}{depth(y_i)}, \qquad (2)$$
$nvars(y_i)$ is the number of original variables involved in the mathematical expression of y_i, and $depth(y_i)$ is the depth of the GP tree representing y_i.
3. **A user evaluation term,** $f_u(y_i)$, that is an average of the user evaluation for each cell corresponding to y_i (range of 1 to 5 from "bad" to "excellent").

Diversity management : The evolutionary mechanisms naturally tend to concentrate the population around good solutions. So for small populations sizes, there is a risk of premature convergence if no diversity preservation mechanism exists. In *EvoGraphDice*, each time a new dimension y'_i is generated, its Euclidean distance to the current population is computed. If y'_i is too close to one of the individuals of the current population, it is replaced by a random individual.

4. Case Studies with Expert Users

We conducted an observational study with five domain experts. During the study sessions, we encouraged participants to think-aloud and share their findings with the study facilitator. We wrote observations, conducted semi-structured interviews and questionnaires, video-recorded the sessions and logged user interactions. The following sections describe the study setup, observations and findings for each expert.

4.1. Method

Due to the open-ended style of exploration using *EvoGraphDice*, and the subjective nature of our fitness function, we chose a qualitative observational study methodology [Car08, MSM12, SMM12] that better suits our evaluation needs. We wanted to evaluate the usability and utility of our tool. In particular, we attempted to answer the following three questions: (i) is our tool understandable and can it be learnt; (ii) are experts able to confirm known insight in their data; and (iii) are experts able to evolve views that contain new insight or allow them to generate a new hypothesis, and if so how easy or difficult is it to reach those findings.

4.2. Participants and Apparatus

We evaluated our prototype with 5 domain expert users (2 female), ages $27-42$ (mean 34.2). Experts were academics and practitioners who had multidimensional datasets related to their domain of expertise (scientific simulation, medicine and geography) and were interested in further exploration. They consisted of one graduate student, four senior researchers and one medical surgeon. Participants had previously explored their datasets using graphical tools (e.g. Excel and JMP) or used statistical methods (PCA and regression analysis) but felt there is more to discover in their data than their current tools allowed them to. Experience with advanced multidimensional visualization tools varied from none, to experts who already used GraphDice or other SPLOM-based tools (two experts). None of our participants previously used dimension combination to analyse their data but three performed PCA-type analysis. The first three case studies ran at our research lab on an HP Z800 workstation PC with a $19''$ dual monitor (1280 x 1024 screen resolution). The last two case studies ran on a similar setup at the experts' institutions. Each session lasted on average 2.5 hours.

4.3. Tasks and Procedure

Participants were asked to carry out two main tasks: (**T1**) show in the tool what they already know about their data, hypothesis and questions they wanted answered; and (**T2**) explore their data in light of these hypotheses and research questions. The first task (and a training game) was designed to test if the tool is understandable, easy to learn, and can help experts rediscover known findings. The second open-ended one explores how domain experts use our tool to answer questions about their data and gain new insights.

Prior to the actual study, participants filled by e-mail a pre-questionnaire to elicit their background, knowledge about the dataset they want to explore, and experience with multidimensional data visualizations. In particular, they were asked to describe the dimensions of the data sets they provided, known relationships between variables, and hypotheses they wanted to investigate. The main study ran in two parts; first training then open exploration as follows:

Training: participants played a game designed to teach them how to operate the tool. A 5D dataset was synthesised with two enclosed curvilinear dependencies between two variables (x0 and x1) and random data for the rest of the dimensions. Participants were asked to evolve a scatterplot where it is possible to separate the two curves in Fig. 3 (left) with a straight line and were given around 20 minutes to complete the task (this task is equivalent to separating the two convex hulls in Fig. 3). Two participants successfully separated the two curves, while the remaining experts evolved views very close to a correct solution within the allocated time.

Open Exploration: the second part of the study ended after about one hour of exploration (a maximum limit of two hours was set based on a pilot to avoid user fatigue), and participants were encouraged to take breaks. A facilitator was present to answer experts' questions and discuss their findings. Throughout the study, a second screen with an open text editor and pen and paper were provided to the experts as means of writing down their exploration findings. At the end, participants filled in a short questionnaire rating aspects of the tool (5-point Likert scales), such as the ease of performing the two main tasks, and open ended questions regarding their exploration strategy and helpful features of the tool.

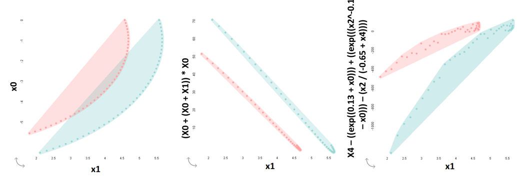


Figure 3: *Two different solutions (screenshots of plots) for the training game problem (left) that involve a simple dimension combination (middle) and a complex formula (right).*

4.4. Data Collection and Analysis

Participants were video-taped and log data of user interactions was gathered for further analysis (table1). Live and video observations, the results of the questionnaire, and the log analysis are described separately for each case study.

4.5. Expert 1: Electrical Consumption Profiles

Dataset: (9D) describes the electrical consumption of 900 anonymised businesses during non-peak (*npk*) and non-plan (*npn*) hours ('plan' refers to an agreed unit rate for a defined period of time) for winter (*W*) and summer seasons (*S*), their geographical altitude and the total consumption cost.

Goal: the expert wanted to investigate electricity consumption patterns of these businesses and their impact on the total cost of consumption. The expert noted in the pre-questionnaire that he would like to sum-up some dimensions in twos in order to focus on one aspect of consumption (e.g. non-plan and non-peak for summer), and therefore had a clear motivation for combining dimensions. This, he argued, may allow him to see interesting consumption profiles.

Observations: the expert hypothesised that altitude has an influence on electricity consumption during both summer and winter seasons. He also had some prior knowledge about existing outliers in the data. During the study he was able to quickly verify both of these hypothesis.

The most important new finding made by the expert, which was not part of the original search space, is a view showing a linear combination of the four parameters of interest to the expert (*npn* and *npk* for summer and winter) which brought to evidence in a quantitative manner that *npnW* consumption is the more correlated to the total consumption.

Expert	G	T1	T2	Q	Data	Size	D	LimitSearch	Evolve	Eval	OVisits	NVisits	Insight
1	-	4	4	3	business	9x900	1:10	3	3	16	40	105	2(1)
2	-	4	4	3	timeseries	7x78	1:33	4	3	8	114	115	4(3)
3	9	5	5	3	geometrical	12x67	0:49	4	21	90	99	344	2(1)
4	7	5	4	3	statistical	10x200	2:23	7	13	83	110	309	6(1)
5	-	3	2	4	geospatial	11x653	1:27	5	5	20	64	229	-

Table 1: *Log data showing: (G) the generation when a solution for the game was found, (T1&T2) experts' scores for ease of completing tasks T1&T2 on a 5-point Likert scale, 5 signifies "very easy", (Q) score for user agreement with EvoGraphDice cell evaluations on 5-point Likert scale, 5 indicates strong agreement, (Data&Size) type and size of dataset, (D) duration (hh:mm) of T2, (LimitSearch) breath of exploration indicated by the number of times the expert limited the search space, (Evolve) depth of exploration indicated by the maximum reached generation, (Eval) how many new cells were evaluated by the user, (OVisits&NVisits) number of times the expert visited the original cells and the new cells respectively, and (Insight) number of times the expert limited the search space and the generation (between parenthesis) where the insight was found.*

In the user's own words: "we always talk about this qualitatively. This is the first time I see concrete weights ... To understand what is a better fare, it is necessary to find a good approximation of the consumption profile", like the one found in Fig. 4. Thus this insight can lead to electricity plans that are better suited to clients' consumption needs.

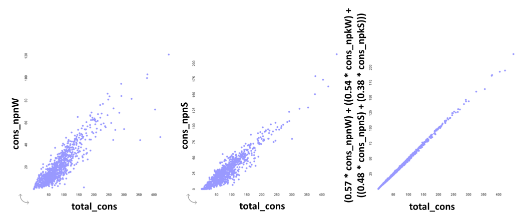


Figure 4: *Confirmed findings (left and centre) and new insight found by the expert (right): a linear combination of four parameters that approximates customer consumption.*

According to this participant, his exploration strategy was to look at propositions in detail along a row, e.g., to examine proposed dimensions plotted against the total consumption. Overall, the expert did not evolve many generations (depth of exploration was three generations at most), but used "limit the search space" facility three times, indicating that he was trying to formulate an interesting hypothesis more than he investigated one in depth. The solution he found was after limiting the search space for the second time.

The expert liked the ability to limit the search space and to enter formulae for the combined dimensions using the dimension editor, e.g. to invert a weight.

4.6. Expert 2: Biscuit Baking Process

Dataset: (7D) describes 78 data points recorded from several industrial biscuit training processes taken by experts in the industry. There are two input parameters relating to temperature settings and three output parameters relating to biscuits (weight loss, height and colour) and a timestep.

Goal: the expert wanted to visualise dependencies in the data between input and output parameters (intuitively such correlation should exist but its exact nature is not clear).

Observations: the expert was able to quickly verify known profiles in the data, for instance the influence of temperature on height and color of the biscuit.

The more general profile of the relationships between input and output parameters was not evident from the original dimensions, thus the expert looked at a wider space using combined dimensions. He observed that there might be some exponential factors that link outputs and inputs, in particular an exponential dimension of one of the input parameters (proposed by the GP) was linearly correlated between all output variables: "we would probably not have considered looking at exponential relationships" indicating a surprising finding and, thus, the ability of the tool to encourage lateral thinking. Further investigation showed that the exponential of temperatures has a specific meaning in thermally activated processes (explained by the Arrhenius law [Wik]).

To look for new relationships, the expert's strategy was to evolve a few generations and choose a visualization that showed linear or quadratic relationships. Like the first expert, he edited promising views using the "dimension editor" to see if this had better or worse impact on the relationship between variables. Most importantly he tried to reduce formula complexity to make better sense of the relationship.

The expert found the evolution functionality and the preview matrix useful to reach interesting cells. However, he wanted to examine scores per dimension as well as per views, e.g. to see if a dimension is always highly ranked; and for the tool to highlight new dimensions that have not been visited before.

4.7. Expert 3: Anatomical Planning for Surgery

Dataset: (12D) of which half describe anatomical and geometrical values related to a 3D planning of a surgical operation (total hip arthroplasty) for 67 patients, and the other half represent values for the same parameters after surgery.

Goal: the expert wanted to investigate whether there is a correlation between the planned values and the final values for each of the investigated parameters, and, if it exists, how strong is this correlation. Since there are many parameters to examine with potentially many interactions between them, the expert focussed on examining the offsets for the cup anteversion parameter (AntvCupSupine) which corresponds to the orientation angle of the cup of the hip prosthesis.

Observations: the expert already knew that there is a relationship between the planned and real values for the AntvCupSupine parameter. This was easily verified in the original dimension space. To explore this further, the expert examined the view showing the before and after values, then made three lasso selections corresponding to over-fit, best-fit and under-fit values (Fig. 5 left), then examined brushed cells in the original search space.

In terms of new insight, the expert found a new cell where the two problematic groups (in red and blue) were separated from the well-restored group (in green) with the exception of one data point. The proposed dimension had a simple formula that involved two original dimensions. Views showing such separation may correspond to special geometrical settings or anatomical features for the observed patients. The expert noted that he needs to examine these patients more carefully with special attention to the selected parameters which can then lead to better pre and post-surgery results.

The expert followed the training game example as his exploration strategy, which may explain the big depth of exploration (21 generations): he made lasso selections of groups of data points, and evolved views that he scored highly depending on whether the overlap between the clusters is minimised. He examined the proposed dimensions in relation to the AntvCupSupine parameter and made use of the "favorite" facility to compare interesting cells that were replaced in the next generations. Notably, he evaluated more than 26% of new visited cells. This seemed to be an important part of his exploration strategy.

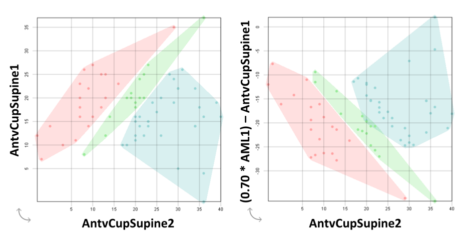


Figure 5: *Selections of over-fit (red), best-fit (green) and under-fit (blue) parameter values (left), and (right) a finding by the expert showing a separation between the two groups of interest in relation to a new parameter.*

The expert commented that he liked the direct visual interaction with the data but he did not like the uncertainty in whether a solution existed and whether the tool will find it. For example, he was interested to see if the degree of separation between his data groups became smaller between evolutions. He suggested adding more adapted tools for selecting data clusters and including statistical information.

4.8. Expert 4: Pareto Front Exploration

Dataset: (10D) describes the outpt of a genetic algorithm that was used to calibrate a city growth and emergence model. The data represents a set of parameter values (7 dimensions) and their objective fitness scores (3 dimensions). The explored dataset only includes the first best 200 parameter values that the algorithm found according to the three objectives of the calibration model (i.e. the Pareto front of the global parameter space exploration).

Goal: the expert wanted to explore the dataset from the two different perspectives (parameter and objective space) as well as the interaction between the two spaces, e.g. does a special profile in the parameter space correspond to a special profile in the objective space?

Observations: prior to the study, the expert had an idea about some characteristics of the data, e.g. there are two large clusters that can be differentiated by the value of one parameter (pAdoption). This type of calibration was also known to produce a characterisable response in the objective space. This hypothesis was easily verified using *EvoGraphDice* via brushing and linking between cells in the parameter space and the objective space.

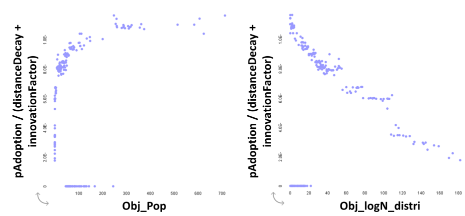


Figure 6: *An interesting combined dimension from the parameter space and its impact on two objective dimensions.*

In terms of new insight, the expert was able to find an interesting combined dimension that gave a good correlation for two parameters of the objective space (Fig. 6). The expert commented that this combination may be an important finding because it involves parameters that affect only one part of the simulation model. This indicates that those parameters, at least for these two output indicators, work together; and that this linear combination could be one way to reduce the complexity of the model.

As for strategies, the expert mentioned primarily limiting the search space (7 times), evolving (13 generations) and examining cells that had monotonic or striated distributions. She also made good use of visual queries and cell evaluation; 27% of new visited cells were evaluated by the participant.

The expert liked the ability to limit the search space and the freedom the tool offers to explore the data by "evolution". She added that the tool was helpful in reaching insight because of its ability to visualise and suggest combinations of dimensions that actually had a visual pattern: "in Excel it is difficult to find a formula that would give a nice pattern".

4.9. Expert 5: Urban Organisation and Perception

Dataset: (11D) describes geo-spatial information about 653 inhabitants and their urban environments including their profiles (e.g. gender, age), perception of their neighbourhood and objective variables describing their street such as the distance to the nearest metro station and type of district.

Goal: as with other experts, this participant was interested in finding relationships between variables; specifically she wanted to identify groups of inhabitants having similar profiles and to find the most discriminating variables for these individuals in order to make sense of the formed groups.

Observations: the expert had already found interesting correlations between the different original variables using her own statistical tools. She was able to confirm these findings. For instance, that an individual's perception of the size of their neighbourhood was dependent on the distance to the next metro station, but this was only true if context (type of district) was taken into account. However, the expert was aware of correlations requiring an interaction between two variables against a third but found it difficult to see them using *EvoGraphDice*. The expert noted two major difficulties that may have hindered the exploration and thus lack of early insight or hypothesis generation: (a) difficulty in determining the criteria for scoring patterns without knowing what a good pattern is in advance; and (b) the nature of data about human behaviour and perception has high variability, thus examining averages, for instance, is more appropriate.

Despite the aforementioned difficulties, the expert found two interesting views where clusters and outlier groups seem to correspond to a known profile (Fig. 7). However, she was not able to fully interpret the proposed combined dimensions as the choice of variables made sense, but the overall interpretation of the pattern was not clear to the participant.

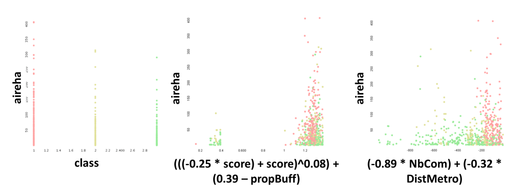


Figure 7: *Two interesting combined dimensions (centre and right) found by the system and their impact on one objective dimension (aireha). Brushing and linking to an original view (left) shows interesting profiles.*

This expert's exploration strategy was to limit the search space to $3-4$ variables and examine their interaction with one original dimension (e.g. perception of space). She also made selections and examined the brushed views in the original space. Since the expert did not evolve many generation (5 max) and only evaluated a few cells (20 overall)–due to the aforementioned difficulties– the system did not learn well the type of distributions the expert was looking for.

The expert tended to agree with the system's proposed scores (Table 1 Q), which she found interesting because of the choice of variables and the simplicity of the proposed formulae. As interpretation of results was difficult using the current point-based presentation, the expert noted that showing aggregated values and variance would help her better understand the views.

5. Summary of Results

Almost all participants were able to easily confirm prior knowledge about their datasets (2 x 'very easy', 2 x 'easy', 1 x 'neutral'). One expert found this task challenging because of the lack of data aggregation that her type of analysis requires. Overall, participants confirmed known correlation, clusters or outliers in their data. In the remainder of this section, we summarise our study findings concerning new found insight, successful tasks and exploration strategies.

5.1. Insight Generation and Tasks

If we include hypothesis formation as part of insight generation, similar to work by Saraiya et al. [SND05], *EvoGraphDice* helped our participants generate new insight in the form of distinct observations about the data (4 experts), new hypothesis (1 experts) and better formulation of research questions (4 experts). Distinct observations found by the experts were either clustering, linear or non-linear relationships, and similarly to generated hypotheses, they always linked a dimension in the original data set and a new proposed dimension. The subjective evaluation of ease of task T2 (table 1) shows most experts found it easy to find new insight: 1 x 'very easy', 3 x 'easy' and 1 x 'not easy'. Not surprisingly, those who reached a concrete new finding scored the tool highly in comparison to those who did not.

The found solutions were regarded by the experts as interesting because they had one or more of the following properties: (i) a visual pattern such as those modeled by the scagnostics measures; (ii) a simple formula involving few dimensions; (iii) a selective choice of dimensions (corresponding to an unformulated hypothesis or an inherent aspect of their data model); and (iv) a domain value. Regarding the latter point, not all participants were able to state the immediate domain value but in general, our participants stated that *EvoGraphDice* helped them:

- interact visually with data (experts 3)
- try out alternative scenarios by editing dimensions (experts 1,2)
- think laterally (expert 2)
- quantify a qualitative hypothesis (expert 1)
- formulate a new hypothesis or refine an existing one (1-4)

5.2. Exploration Strategies

Overall, participants followed the same exploration pattern consisting of first examining the original dimensions then inspecting and evaluating the first generation of the proposed dimensions (returned by the PCA) followed by one or more iterations of the following steps: (i) limit the search space; (ii) select and rank cells; (iii) evolve; and (iv) interpret and verify. However, the frequency of using some tools (e.g. "evolve" vs. "limit the search space") varied depending on whether the expert had an *a priori* focused hypothesis (i.e. a research question involving typically $3-4$ dimensions). We observed that the looser the initial hypothesis, the more often they tried to change the search space; and the more focused the hypothesis the more generations they inspected. Indeed, these two strategies of *exploration* and *exploitation* are supported by EAs [Ban97] where on the one hand the user wants to visit new regions of the search space and on the other hand they want to explore solutions (combined dimensions) close to one region of the search space.

6. Discussion

Most of our experts were able to formulate interesting hypothesis or reach new insight requiring looking at data in terms of a combination of dimensions. Our approach consists of proposing new views based on automatically calculated metrics and user feedback. On the one hand, our method is complimentary to PCA, clustering and regression analysis that automatically find data patterns, and optimise a fit. On the other hand, we allow users to interactively select examples of visual pattern types they are interested in, and that may not be easy to express mathematically. Users can then verify the new relationships they find in *EvoGraphDice* by using the dedicated automatic data analysis tools.

In comparison to automatic analysis such as in statistics and data mining, our approach offers: (i) **Intuitiveness:** a visual approach to interact with data requiring no prior statistical knowledge; (ii) **Interactivity:** rather than fitting the data to pre-defined shapes in a static manner, using an IEA the user can dynamically steer the exploration process towards a pattern of interest. These patterns can involve dimension concatenations that are not obvious at the outset of the exploration; (iii) **Flexibility:** ability to edit and try out alternative dimension combination scenarios, or limit the search space. (iv) and **Adaptability:** the system can adjust to user change of interest over time.

There are limitations to using our tool, such as the types of datasets to explore and issues related to the interpretation of combined dimensions which we discuss below. Moreover, the issue of the convergence of the genetic algorithm is an interesting one given that the IEA deals with optimisation. However, this is not easy to study in our case as there is no unique solution to converge to, rather the optimisation is dynamically adapted to follow user interest over time.

First, we are constrained by the SPLOM representation of *EvoGraphDice* which does not provide a natural way to interact with some dataset types such as timeseries. Data with high variability provides additional challenges that we do not currently address, such as detecting and evolving aggregated patterns. In addition, we tested our prototype with user-provided datasets that are small to medium sized, having dimensions between $7-12$. Although our algorithm can deal with a large number of data points, it may not handle well larger number of dimensions as complex combined dimensions may be difficult to avoid. In this case, a dimension reduction technique can be applied to the dataset before feeding the results to *EvoGraphDice*.

Second, not all variables can be combined, therefore the user should as soon as possible limit the search space to "combinable" dimensions. This in a sense requires the user to have some domain knowledge and to make an initial hypothesis about the data. The proposed dimensions can involve complex or unforeseen combinations yielding a visual pattern but one that can be difficult to interpret. To help address this issue, we used "complexity" of a combined dimension as a component of the IEA fitness function. Nonetheless, our method can still yield complex dimensions that are difficult to interpret. We noticed that our participants only looked at the combined dimensions in relation to an original dimension, most likely to ground the observation. This problem of interpretation, however, is common to all tools that offer dimension combination.

7. Conclusion and Future Work

We presented a prototype tool (*EvoGraphDice*) for supporting Evolutionary Visual Exploration (EVE) that combines visual analysis with interactive evolutionary computation to help steer the exploration towards interesting views on the data. Our method complements PCA, clustering and regression types of analysis, offering additional features such as interactivity and adaptability. We conducted an observational study with domain experts and found that our tool allowed users to evolve characteristics that are not visible in the original dimensions space. Our experts were able to try out different scenarios, think laterally, quantify qualitative hypotheses and formulate new ones.

Future work for our tool includes longitudinal user studies to explore in detail the long term evolution of user focus, as well as addressing issues such as improving the IEA to detect more complex visual patterns (beyond those currently detected by Scagnostics), handling data with high-dimensionality and bridging *EvoGraphDice* and existing statistical packages to combine powerful statistical analysis with flexible and intuitive visual exploration.

Our work demonstrated that tightly combining visualization and optimisation techniques can yield exciting results in data analysis and opens new venues for research, but also highlights challenges such as monitoring algorithm convergence, history visualization of diverging exploration paths, and appropriate methodologies for evaluation.

Acknowledgements:
We thank all our expert users for their participation in this study, and J-D Fekete for his useful feedback and comments.

References

[Asi85] ASIMOV D.: The grand tour: a tool for viewing multi-dimensional data. *SIAM J. Sci. Stat. Comput. 6*, 1 (Jan. 1985), 128–143. 1

[Ban97] BANZHAF W.: *Handbook of Evolutionary Computation*. Oxford University Press, 1997, ch. Interactive Evolution. 2, 3, 9

[BCD*10] BEZERIANOS A., CHEVALIER F., DRAGICEVIC P., ELMQVIST N., FEKETE J.-D.: GraphDice: A System for Exploring Multivariate Social Networks. *Computer Graphics Forum (Proc. EuroVis 2010) 29*, 3 (2010), 863–872. 3

[BTK11] BERTINI E., TATU A., KEIM D.: Quality metrics in high-dimensional data visualization: An overview and systematization. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (Dec. 2011), 2203–2212. 1, 3

[Car08] CARPENDALE S.: Evaluating information visualizations. In *Information Visualization*, Kerren A., Stasko J. T., Fekete J.-D., North C., (Eds.). Springer-Verlag, Berlin, Heidelberg, 2008, pp. 19–45. 4

[CBL12] CANCINO W., BOUKHELIFA N., LUTTON E.: Evographdice: Interactive evolution for visual analytics. In *IEEE Congress on Evolutionary Computation, June 10-15* (2012). June 10-15, Brisbane, Australia. 4

[EDF08] ELMQVIST N., DRAGICEVIC P., FEKETE J.-D.: Rolling the Dice: Multidimensional Visual Exploration using Scatterplot Matrix Navigation. *IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 2008) 14*, 6 (2008), 1141–1148. 3

[Evo] EVOGRAPHDICE:. http://www.aviz.fr/Research/EVE. Accessed: 10/03/2013. 2

[Fri87] FRIEDMAN J. H.: Exploratory projection pursuit. *J. Am. Statistical Assoc 82*, 397 (1987), 249–266. 1

[Gol89] GOLDBERG D. E.: *Genetic Algorithms in Search, Optimization and Machine Learning*, 1st ed. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1989. 1

[HT00] HAYASHIDA N., TAKAGI H.: Visualized IEC: interactive evolutionary computation with multidimensional data visualization. In *IECON 2000. 26th Annual Conference of the IEEE* (2000), vol. 4, pp. 2738–2743. 3

[Hub85] HUBER P. J.: Projection Pursuit. *The Annals of Statistics 13*, 2 (1985), 435–475. 1

[JB05] JIN Y., BRANKE J.: Evolutionary optimization in uncertain environments-a survey. *IEEE Trans. Evolutionary Computation 9*, 3 (2005), 303–317. 2

[Koz92] KOZA J. R.: *Genetic Programming* . MIT Press, 1992. 4

[LPLV05] LUTTON E., PILZ M., LÉVY VÉHEL J.: The fitness map scheme. application to interactive multifractal image denoising. In *CEC2005* (Edinburgh, UK, September, 2-5 2005), IEEE Congress on Evolutionary Computation. 3

[LSA*06] LLORÀ X., SASTRY K., ALÍAS F., GOLDBERG D. E., WELGE M.: Analyzing active interactive genetic algorithms using visual analytics. In *Proceedings of the 8th annual conference on Genetic and evolutionary computation* (New York, NY, USA, 2006), GECCO '06, ACM, pp. 1417–1418. 3

[Lut06] LUTTON E.: Evolution of fractal shapes for artists and designers. *IJAIT, International Journal of Artificial Intelligence Tools 15*, 4 (2006), 651–672. Special Issue on AI in Music and Art. 2

[MGJH11] MATKOVIC K., GRACANIN D., JELOVIC M., HAUSER H.: Interactive visual analysis supporting design, tuning, and optimization of diesel engine injection. *Proceedings of IEEE Visualization 2011 (Discovery Exhibition)* (2011). 3

[MSM12] MEYER M., SEDLMAIR M., MUNZNER T.: The Four-Level Nested Model Revisited: Blocks and Guidelines. In *Proceedings of the VisWeek Workshop Beyond Time and Errors: Novel Evaluation Methods for Information Visualization (BELIV)* (2012), ACM Press. 4

[PC97] POLI R., CAGNONI S.: Genetic programming with user-driven selection: Experiments on the evolution of algorithms for image enhancement. In *Genetic Programming 1997: Proceedings of the Second Annual Conference* (1997), Morgan Kaufmann, pp. 269–277. 3

[PLM08] POLI R., LANGDON W. B., MCPHEE N. F.: *A field guide to genetic programming*. Published via http://lulu.com and freely available at http://www.gp-field-guide.org.uk, 2008. (With contributions by J. R. Koza). 2

[PWR04] PENG W., WARD M. O., RUNDENSTEINER E. A.: Clutter reduction in multi-dimensional data visualization using dimension reordering. In *Proceedings of the IEEE Symposium on Information Visualization* (Washington, DC, USA, 2004), INFOVIS '04, IEEE Computer Society, pp. 89–96. 3

[Smi02] SMITH I.: A tutorial on principal component analysis, 2002. 4

[SMM12] SEDLMAIR M., MEYER M., MUNZNER T.: Design Study Methodology: Reflections from the Trenches and the Stacks. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis) 18*, 12 (2012), 2431–2440. 4

[SND05] SARAIYA P., NORTH C., DUCA K.: An insight-based methodology for evaluating bioinformatics visualizations. *IEEE Transactions on Visualization and Computer Graphics 11*, 4 (July 2005), 443–456. 8

[SS05] SEO J., SHNEIDERMAN B.: Rank-by-feature framework for interactive exploration of multidimensional data. *Information Visualization 4*, 2 (2005), 99–113. 3

[Tak98] TAKAGI H.: Interactive Evolutionary Computation: System Optimization Based on Human Subjective Evaluation. *INES'98* (1998). 3

[Tak08] TAKAGI H.: New topics from recent interactive evolutionary computation researches. In *Knowledge-Based Intelligent Information and Engineering Systems* (2008), p. 14. 2

[Wik] WIKIPEDIA: Arrhenius equation. http://en.wikipedia.org/wikiArrhenius_equation. Accessed: 10/03/2013. 6

[WW08] WILKINSON L., WILLS G.: Scagnostics Distributions. *Journal of Computational and Graphical Statistics 17*, 2 (2008), 473–491. 3

A Comparison of Visualizations for Identifying Correlation over Space and Time

Vanessa Peña-Araya, Emmanuel Pietriga, Anastasia Bezerianos

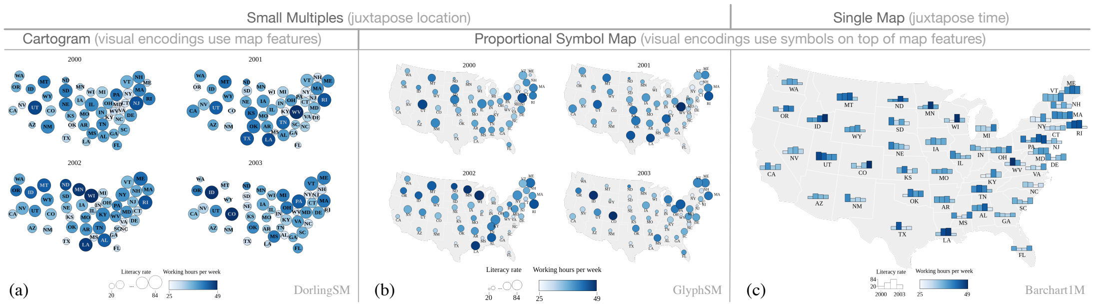


Fig. 1. The three visualizations compared in our study. (a) Dorling cartograms as small multiples, (b) proportional symbols (circles) on maps as small multiples, and (c) proportional symbols (bar charts) on a single map. In this example, each map shows the values of two artificially-created variables over four years. In each case, both variables have an overall positive correlation (Pearson correlation coefficient ≥ 0.75) and no monotonic evolution.

Abstract— Observing the relationship between two or more variables over space and time is essential in many domains. For instance, looking, for different countries, at the evolution of both the life expectancy at birth and the fertility rate will give an overview of their demographics. The choice of visual representation for such multivariate data is key to enabling analysts to extract patterns and trends. Prior work has compared geo-temporal visualization techniques for a single thematic variable that evolves over space and time, or for two variables at a specific point in time. But how effective visualization techniques are at communicating correlation between two variables that evolve over space and time remains to be investigated. We report on a study comparing three techniques that are representative of different strategies to visualize geo-temporal multivariate data: either juxtaposing all locations for a given time step, or juxtaposing all time steps for a given location; and encoding thematic attributes either using symbols overlaid on top of map features, or using visual channels of the map features themselves. Participants performed a series of tasks that required them to identify if two variables were correlated over time and if there was a pattern in their evolution. Tasks varied in granularity for both dimensions: time (all time steps, a subrange of steps, one step only) and space (all locations, locations in a subregion, one location only). Our results show that a visualization's effectiveness depends strongly on the task to be carried out. Based on these findings we present a set of design guidelines about geo-temporal visualization techniques for communicating correlation.

Index Terms—geo-temporal data, bivariate maps, correlation, controlled study, bar chart, Dorling cartogram, small multiples

1 Introduction

Understanding phenomena often requires looking at multiple variables, their inter-relationships, and how these evolve over time. Take Hans Rosling's visualization of the demographics of countries in his seminal 2006 TED talk [55]. Looking at the life expectancy and the fertility rate together is key to understanding the phenomenon at hand. Watching their co-evolution provides many of the insights unveiled by the speaker.

In many cases, the data will also feature a spatial dimension. Rosling refers to individual countries, but also different groups of countries multiple times. The spatial dimension plays an important role in his story, even if it is only indirectly represented in the scatterplot. Again, understanding the interplay between the considered variables, and the spatial arrangement of the entities they describe, can yield key insights.

- *Vanessa Pena-Araya is with Univ. Paris-Sud, CNRS, INRIA, Université Paris-Saclay. E-mail: vanessa.pena-araya@inria.fr.*
- *Emmanuel Pietriga is with Univ. Paris-Sud, CNRS, INRIA, Université Paris-Saclay. E-mail: emmanuel.pietriga@inria.fr.*
- *Anastasia Bezerianos is with Univ. Paris-Sud, CNRS, INRIA, Université Paris-Saclay. E-mail: anastasia.bezerianos@lri.fr.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

This famous example illustrates the potential of multivariate geo-temporal data visualization as a storytelling device. The speaker communicates insights about two variables that are related thematically, and that describe a phenomenon that is situated both spatially and temporally [2]. Beyond data storytelling, geo-temporal visualization can also support the analysis of such phenomena. The context, however, is different. While animation can illustrate temporal evolution when telling a story, it will often not be as effective for analysis purposes [65]. Moreover, depending on the application domain considered, information about group membership (*e.g.*, a country belonging to a particular continent) might not be sufficient to understand what role the spatial dimension plays in the phenomenon. Thus more detailed information about the topological relationship between entities might be necessary.

The problem of designing an effective visual representation in this context is challenging, as multiple data of different nature must be combined, each having specific characteristics: the thematic variables that describe the first-class entities in the dataset (life expectancy, fertility rate), the spatial properties of those entities (countries, continents), and the evolution of the thematic variables over time (years). Design choices will influence how well the representation can enable analysts to detect correlations between variables over space and time. It is thus important to identify guidelines to inform such designs.

Prior studies have compared geo-temporal visualization techniques for a single variable that evolves over space and time [21, 39, 40, 58].

Others have looked at two variables on a map (bivariate maps), but at a specific point in time [15, 18, 45]; or at how to visualize the correlation between two variables [31, 52, 53, 69], including visualizations that can be used to depict temporal evolution [26], but not in a geospatial context. To our knowledge, how effective visualization techniques are at communicating correlation between two thematic variables, that evolve over both space and time, remains to be studied.

We identify the different strategies used to combine thematic, spatial and temporal data into a visualization. The first design choice to be made concerns the combination of thematic variables in the representation: is the representation **juxtaposing all locations for a given time step**; or **juxtaposing all time steps for a given location**. The second choice concerns the visual encoding of thematic variables: either **overlaying symbols on top of map features**; or **using visual channels of the map features themselves**.

We discuss design variations for each strategy and identify three candidate techniques (see Fig. 1). Our study is designed to evaluate participants' ability to identify whether two variables are correlated over time or not, and if they are, if there is a pattern to their evolution. As we expect the techniques to fare differently depending on the number of time steps and the number of geographical entities to consider, we test them on tasks that vary both in temporal and in geographical granularity. Our results confirm this intuition, leading to a set of design guidelines about visualization choices for effectively communicating correlations in thematic geo-temporal data.

2 Related work

We first review some of the available visualizations categorized by how they combine space and time, and then how thematic variables are encoded to create multivariate maps. We finally discuss research related to perception studies of visualizations of correlated geo-temporal data.

2.1 Visualizing Combined Dimensions of Space and Time

Maps are the most direct visual representation of geo-temporal data. When combining both dimensions of space and time, thematic variables can be displayed by either juxtaposing locations (*e.g.*, small multiples of compact map representations); or juxtaposing time (*e.g.*, glyphs that represent multiple time steps overlaid on locations on a single map).

Juxtaposing location. From this category, small multiples are the most popular technique. For example, Johnson *et al.* [29] use small multiples to observe the correlation of Internet adoption with GDP and with population over the years. Animation can also be considered as a technique that juxtaposes location on maps that are presented in a sequence. Animation has been used to smooth the transition between views [9], or combined with symbols to depict change [32].

Juxtaposing time. The most common approach is to use glyphs in 2D (*e.g.*, [3, 17, 33, 47, 61]) or 3D (*e.g.*, [64]) on top of a single map. Additionally, the 3rd dimension has been used to juxtapose time over a map. For example, Space Time-Cubes [35] arrange time steps on the z-axis, effectively piling up the maps that correspond to each one of them. They have been used in several applications, *e.g.*, [19, 43, 50].

2.2 Visually Encoding Thematic Variables

Visually encoding data on a map can be done using two main strategies: mapping thematic attributes to visual properties of the map features; or overlaying symbols (*e.g.*, basic shapes such as circles, or glyphs such as pie charts and bar charts) on top of a base map, which remains untouched. As stated by Elmer [15], the number of possibilities to create bivariate or multivariate maps can range from dozens to hundreds (the declarative model of Jo *et al.* [28] for multiclass density maps shows numerous examples). Thus in this section we focus on those representations that are most commonly used or studied.

Encodings that use visual channels of the map features. Choropleth maps are among the most popular in this category [22, 41, 62]. They visually encode thematic attribute values using the map features' fill color. A bivariate type of choropleth, called value-by-alpha maps, allows for two variables to be displayed at the same time by combining color hue and transparency for each map feature [18].

	Juxtapose location	Juxtapose time	No time
Visual encodings use symbols on top of map features	[6] [16] [58]*	[33]† [39, 40]	[15]* [18]* [30]* [63]† [71] [38] [70] [4]
Visual encodings use map features	[21] [44] [45]*	[39, 40] [45]*	[15]* [18]* [30]* [63]† [23]* [24] [42] [62]

Table 1. Categorization of studies comparing geo-spatial visualizations. The first two columns represent the juxtaposition strategy. The third groups studies which compare visualizations that do not include time. The two rows represent the categories of visual encodings (symbols or map features). (*) indicates studies that consider more than one quantitative variable, and (†) studies that consider one quantitative and one qualitative variable. Note that some references are included in more than one cell as they make comparisons across categories.

Cartograms, use size as a visual encoding channel, and deform geographical shapes proportionally to the variable of interest [46]. There are four major types of cartograms: contiguous, non-contiguous, Dorling and rectangular. Contiguous cartograms distort regions to make their size reflect the thematic variable's value, preserving topology, and in particular adjacency, at the cost of statistical accuracy. Non-contiguous cartograms rescale each region of the map independently. They yield better statistical accuracy but fail to preserve topology (geographical regions are no longer contiguous). Dorling cartograms [12] yield more abstract representations of the geographical entities, replacing each region with a circle (Fig. 1-a). The circle's area can be mapped to a thematic variable. The position of circles is computed so as to preserve the overall topology, putting each circle as close to its original location as possible, adjusting their actual position to avoid circles overlapping one another. Finally, rectangular cartograms are similar to Dorling cartograms, but use rectangles to represent each region, yielding even more abstract representations of the geographical entities. Bivariate cartograms [66] use color or shade to encode a second variable in addition to that mapped to size. A recent variation on bivariate cartograms was presented by Nusrat *et al.* [45], in which two variables are visually encoded with size.

Encodings that use visual channels of symbols on top of map features. Overlaying thematic glyphs on top of a base map (*"symbols on maps"* [25]) gives more flexibility compared to mapping data to the attributes of the map features themselves. A wide variety of glyphs can be used to encode multivariate data. They are typically placed on top of geographical regions, on an independent layer. Proportional circles are the most frequently-used shape, but other basic shapes like squares, triangles or any other symbol can also be used [66]. Beyond simple shapes, more elaborate glyphs have been proposed; from generic glyph designs such as star glyphs or Chernoff faces [7] to domain-specific ones such as those used in meteorology [68].

2.3 Perception Studies on Correlated Geo-Temporal Data

We now summarize the studies we consider most relevant to geospatial visualization. From the extensive literature, we selected a subset using keyword searches involving *maps*, *geographical*, *geo-temporal*, *empirical study*, *evaluation*. We filtered out papers that were more than 20 years old, ones that consider numerical metrics but not visual perception (*e.g.*, [1, 41]), or that evaluated a new proposed technique in isolation (*e.g.*, [14, 37]). The final set of articles can be seen in Table 1.

We observe that most work on evaluating map-based visualizations does not focus on temporal evolution. From the results of those that do, we conclude that choosing the best-suited technique will depend on the task. For example, for analyzing statistical data over time and space, the results of Boyandin *et al.* [6] indicate that users get more insights with small multiples than with animation. This is confirmed by Robertson *et al.* [54] for the analysis of trends using non geo-spatial visualizations. For identifying moving patterns, Griffin *et al.* [21] show that animation leads to better results than small multiples. Other studies that consider temporal change focus on comparing only two points in time (*e.g.*, [44, 45]). They do not provide insights about the compared techniques' performance for identifying trends over space and time.

Regarding visual encoding, we observe that most studies do not focus on more than one quantitative variable at the same time. Particularly regarding correlation, two of them study user performance for tasks that require analyzing the relationship between two variables. The first, from Gao *et al.* [18], compares value-by-alpha maps with non-contiguous cartograms and proportional symbol maps. The latter displayed better overall performance. The second is from Elmer [15], who evaluated eight different visual encodings for bivariate maps. He focused on studying the effectiveness of different combinations of visual variables for the analysis of patterns. His results indicate that the eight combinations were consistent in accuracy, showing the utility of bivariate maps. Time was not considered in these studies.

Other research studies the perception of spatial autocorrelation [4,34] (how much a phenomena is dependent on spatial location). Yet other studies investigate the perception of correlation in visualizations that do not involve maps [26, 31, 52, 53, 69]. While such studies relate to our work, none of them considers all dimensions (correlation of *two* variables, over both space *and* time) simultaneously.

3 Study Rationale and Hypothesis

The literature describes many visualization techniques capable of encoding two thematic variables in a geo-temporal context. As it would be impractical to test them all, we discard general strategies that are ill-suited to the context of visual analysis, and identify representative techniques based on the strategies briefly introduced earlier. We then motivate our tasks, formulate our hypotheses, and explain how we have generated the synthetic datasets used in the study.

3.1 Selection of Visualization Techniques

Our first decision is to discard techniques that use animation to convey the temporal evolution of thematic variables. There has been much discussion about the role of animations [8] and their effectiveness [65], with sometimes-contradictory findings. But there seems to be relatively broad consensus that they are ineffective for detailed analyses of multiple variables over sequences of many time steps: showing only a single step at a time, they require users to remember previously-seen steps, thereby increasing cognitive load [27].

We also discard techniques that use 3D representations. These can provide more opportunities for mapping data attributes to visual variables (see, *e.g.*, [64]), which can be useful when visualizing multivariate data. But they typically force users to interact more with the representation, and require more elaborate means of navigation because of the higher number of degrees of freedom, among other pitfalls [60].

To make our study tractable, we make one final choice: to focus on visualizations based on how they represent the information, independently of any interaction technique. This means that we consider only static visualizations, in which elements can neither be filtered nor highlighted. As we discuss later in Sec. 7.1, follow-up studies should investigate how adding interaction impacts performance, but as this is the first empirical study to investigate the perception of correlation over space and time, there are already many factors to include before considering interaction techniques.

Based on these choices, we identify strategies used to combine thematic, spatial and temporal data into one visual representation. 1) We first categorize visualizations according to how they organize thematic variables. They can juxtapose values for all locations at a given time step, yielding **small-multiples** maps. Or they can juxtapose values for all time steps at a given location, yielding a **single map**. 2) We then categorize visualizations according to how thematic variables are visually encoded [15]. They can be mapped to the visual properties of symbols overlaid on top of the corresponding map features, eventually forming a **proportional symbol map** [18]. Or they can be mapped to the visual properties of the map features themselves. Both choropleth maps and cartograms fall in this category, but we only consider **cartograms** here. Indeed, encoding two thematic variables on choropleth maps is mostly limited to fill color hue, saturation and brightness, but these often interfere in terms of visual perception. Variations on the original design exist, such as, *e.g.*, Banded Choropleth Maps [14], but have not proven effective so far.

Combinations of these different strategies each yield multiple design variations. To avoid having to handle an unmanageable number of conditions, we choose at most one design per combination of strategies, and limit ourselves to designs that are actually used in practice. Those choices are rationalized below, taking into account the fact that our two thematic variables are quantitative in nature.

Proportional Symbol Map + Small Multiples: these techniques juxtapose values for all locations at a given time step. They consist of multiple identical base maps, one for each time step, with symbols superimposed on top of map features. The symbols' visual channels encode the thematic variables, showing individual values for the corresponding time step. We select circles, as they are the most frequently used shape [66], mapping the thematic variables to their radius and fill color brightness, respectively. This technique, which we refer to as `GlyphSM` in the study, is illustrated in Fig. 1-b.

Proportional Symbol Map + Single Map: these techniques juxtapose values for all time steps at a given location. They consist of a single base map. Because all values for all time steps are juxtaposed, we can create miniature bar charts [28], encoding one of the thematic variables using bar length instead of circle radius. Length is considered a more effective encoding channel than area, and this also makes for a more compact glyph than juxtaposed circles would. The second variable is mapped to each bar's fill color brightness. This technique, which we refer to as `Barchart1M`, is illustrated in Fig. 1-c.

Cartogram + Small Multiples: these techniques juxtapose values for all locations at a given time step and encode thematic attributes directly on the map features, without using symbols. They consist of multiple cartograms, one for each time step in the dataset. Among all variations on cartograms (discussed in Sec. 2.2), prior studies have shown that contiguous cartograms and Dorling cartograms perform best overall [44]. We chose Dorling cartograms over contiguous cartograms as results of previous studies indicate they yield higher statistical accuracy and are better suited to *summarize* tasks, therefore better aligned with the analysis of correlations. This technique, which we refer to as `DorlingSM`, is illustrated in Fig. 1-a.

Cartogram + Single Map: while instances of this combination do exist, all the ones we identified are somewhat contrived. Indeed, it is difficult to have a single small glyph meet all requirements: (i) show two thematic variables; (ii) show individual values for each of them, (iii) for each time step; and (iv) preserve the global topology of map features. One possibility would be to take the above Dorling cartogram, slice the circles radially into as many time steps (transforming them into pie charts), and map the thematic variables to each slice's radius and fill color, effectively creating a rose chart. Such a design, however, makes it difficult to compare values across entities. Other possibilities exist, involving, *e.g.*, augmented donut charts or treemaps, but none of these is in reasonably widespread use and none stands out as a promising technique. We thus did not include this combination in the study.

3.2 Task Motivation

Our goal was to compare the effectiveness of visualization techniques, when it comes to identifying the correlation between two variables and its evolution over time. We had no hypothesis about which part is more difficult: detecting different types of correlations (positive / negative / non-existent), or characterizing their evolution (following a trend or not). We thus treat them as a single integrated task, that requires viewers to identify both potential correlations and their trends. We varied the combinations of these factors in our tasks to cover their range, but without exhaustively testing all combinations (Sec. 3.4) and without making any assumption about their difficulty. Such integrated tasks fall under *"characterize the relationship among multiple map features"* in Roth's task taxonomy [56].

To construct our tasks, we used the geo-temporal framework proposed by Peuquet [51], that describes the linked triad of "what", "where" and "when". Each task corresponds to a question of the type *when* + *where* $\rightarrow$ *what*, where *what* is the participant's characterization of the correlation and its evolution.

We varied the *when* and *where* in a way similar to other research (*e.g.*, [20, 59]), using three granularity levels. In particular, the classifi-

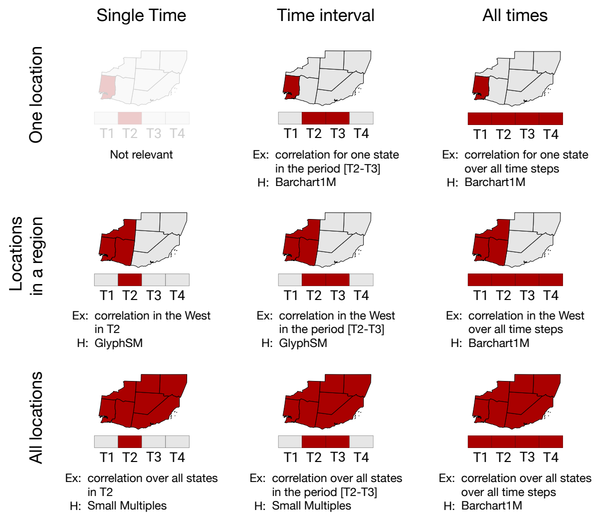


Fig. 2. Summary of tasks based on spatio-temporal granularity. In each cell, the image illustrates the task, together with an example (Ex), and our hypothesis (H) about which visualization will perform best overall.

cation of granularity levels for time (*when*) is divided in (i) one time, (ii) a time interval, and (iii) all times. Space (*where*) is categorized as (i) one location, (ii) locations in a region, and (iii) all locations. Crossing the spatial and temporal dimensions results in a matrix of nine possible tasks illustrated in Fig. 2, together with a concrete example. Correlation at one location in one point in time (top-left cell) is not meaningful and was discarded as a task. We thus ended up with eight possible spatio-temporal tasks.

We hypothesized that the best-performing visualization would depend on the task considered. Specific hypotheses are detailed in Sec. 3.3, and the techniques hypothesized to perform best for each task are also indicated in Fig. 2.

3.3 Hypotheses

The following hypotheses capture our expectations and were formulated before data was collected:

H1: We expect small multiples (`GlyphSM`, `DorlingSM`) to result in better performance (less time and fewer errors) than single maps (`Barchart1M`) for tasks that require analysis at *one point in time* only. The search for the desired point in time is done only once across small multiples, and then the focus is on the spatial information that is grouped closely together. Whereas for a single map the specific point in time needs to be identified repeatedly across map features (bar charts).

H2: For *time intervals* in a single location, we expect better performance (time or errors) for a single map, as all information is colocated (one bar chart) (**H2.1**). When it comes to locations in a region, or to all locations, small multiples (`GlyphSM`, `DorlingSM`) will fare better than single maps (`Barchart1M`) (**H2.2**). We expect that repeatedly identifying the right time interval across multiple locations in a single map will make this visualization slower and lead to more errors.

H3: We expect single maps, that juxtapose time (`Barchart1M`), to result in better performance (time or errors) than small multiples (`GlyphSM`, `DorlingSM`) for tasks that require analysis over all time steps. Indeed, small multiples require users to continuously change their focus between many maps to see trends for locations and make comparisons. This is not the case for single maps as they allow getting an overview of the behavior at each location quickly and identify trends.

H4: We expect that among small multiple techniques, `GlyphSM`, which overlays symbols on a base map, will feature better performance across all tasks. Cartograms (`DorlingSM`) adjust the layout of features in each map independently, thus making it hard to identify and match them across small multiples.

3.4 Dataset and Task Construction

For the setup of our experiment we use the map of the United States (*i.e.*, map features are US states) over nine years of temporal evolution (*i.e.*, a point in time is a year).

The geography of the US states provides good diversity in terms of size of individual features (*e.g.*, Texas compared to South Carolina) and density of those features (*e.g.*, west coast compared to east coast). In trials involving a single location, we varied the size of target features (smaller & larger states) and density of the surrounding geographic area. We grouped locations in contiguous regions using the four cardinal points: north, south, west and east. These regions were selected as they represent common geographic division of countries or other administrative levels. Regions were determined by drawing an imaginary line that divided the country into two equally-sized areas, vertically for east and west, horizontally for north and south. This resulted in areas of varying density across trials. To avoid participants fixating their gaze over discontinuous areas, especially for tasks involving a subset of locations in a region (Fig. 2, second row), we removed Alaska and Hawaii from the map. This resulted in a total of 48 locations, a fair amount of locations to analyze.

Regarding time granularity, we define all time spans to be nine years long (a number that utilizes the space of a small multiple setup). Time intervals were made of four consecutive steps, selected in the middle of the range so as not to favor single maps – identifying the first or last part of the small bar charts is much easier. Four years represent almost half of the total time steps, which allows us to balance the amount of patterns (correlations to identify) and noise (additional data-points to make the task realistic).

The two variables were presented to participants as *literacy rate* and *working hours per week*. Nevertheless, to control the displayed correlation and trends within the different spatio-temporal constraints, we used artificially-created datasets. We initially created variables that followed normal distributions, as other perception studies about the visualization of correlation do [26,53]. With this type of distributions, it is common that points do not follow strict patterns of both increasing at the same time (in case of positive correlation), or one increasing as the other decreases (in case of negative correlation). This is not a problem with scatterplots, as the overall distribution of many points helps convey the overall relationship. However, in our case, the number of points in time was small, minimum 4 for time intervals and maximum 9 for all time steps. Thus, even if one point did not follow the pattern, it would suggest that there was no correlation. We instead generated pairs of points using a random linear regression model with added Gaussian-centered noise,[1] as the difference between values could be evaluated more clearly. The obtained points were checked to ensure that they follow the pattern for the desired time range. To make the generated distributions closer to actual literacy rate and working hours per week, we scaled our generated data between values extracted from Rosling's GapMinder example. For instance, for the variable assigned to literacy rate, we scaled between a minimum within $[20,30]$ and a maximum within $[75,85]$. For the variable assigned to working hours per week, the minimum varied within $[25,35]$ and the maximum within $[40,50]$.

For each task, we created a dataset that followed particular spatio-temporal patterns. The possible correlation patterns were: positive correlation ($r \geq 0.75$) with and without monotonic evolution; negative correlation ($r \leq -0.75$) with and without monotonic evolution; and no correlation ($|r| \leq 0.2$). These patterns were enforced for the space and time granularities considered in each task (*e.g.*, a time range or all times).[2] We added distractors for the locations and time points that were not the focus of the task by including 1/3 of data points that did not follow the assigned pattern.

To increase reliability, our design included three repetitions per task, that were aggregated in our analysis. To avoid learning for each repetition, we varied the selected location, region, point in time and time interval. We generated one dataset per task repetition that, for

[1] Data was created with Scikit-learn [48], using `make_regression`.

[2] We note that for tasks that require analysis in one point in time, it was not relevant to create two variables with monotonic evolution.

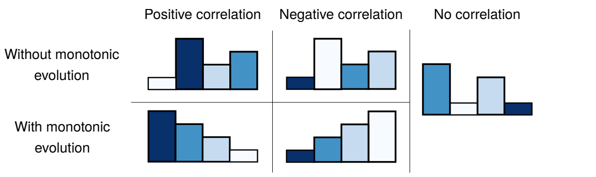


Fig. 3. Schematic illustration of all possible answers for tasks in Fig. 2. The three trial repetitions included combinations, such that each correlation type (positive, negative, no-correlation) appeared once. When temporal evolution was applicable, one of the positive/negative correlations was coupled with monotonic evolution, while the other was not.

the spatial and temporal constraints required by the question (time and space granularity), followed a different correlation pattern. For the three repetitions, there was always: one trial with no correlation, one with correlation (positive or negative) but no monotonic evolution, and one with correlation (positive or negative) and monotonic evolution. Fig. 3 illustrates the different configurations.

To avoid participants remembering answers across visualizations, from each generated dataset we derived two additional variations by shuffling data over states, over years, or both. Thus, for each task repetition displayed in each visualization, the participants would observe a dataset with the same structure but with different layouts. In total, this resulted in 80 datasets: 8 tasks $\times$ 3 repetitions $\times$ 3 datasets (1 original + 2 shuffled variations) = 72 for main trials + 8 for training.

4 Study Design

The study was designed to evaluate, for each of the tasks, the three visualizations introduced earlier. Supplemental material containing dataset generation code, experiment data, analysis scripts and detailed results are available at `http://ilda.saclay.inria.fr/spacetimecorr`.

4.1 Experimental Design

We used a within-subjects design where all participants were exposed to all three visualization techniques. For each technique, a participant had to perform 8 training trials, and 8 measured tasks $\times$ 3 repetitions = 24 main trials. Repetitions considered one of each possible correlation types: Positive, Negative or No-Correlation. For tasks that involved analysis over time, answers also included monotonic choices (Fig. 3).

Technique presentation order and dataset variations were counterbalanced across participants using Latin squares. Tasks were grouped by time granularity (one point in time, time interval, all times) and their order of presentation was counterbalanced as well. For each time granularity, the order of geographical granularity was randomized. Within each group of space and time granularity, the three task repetitions were also randomized. In total, the experiment consisted of 18 participants $\times$ 3 visualizations $\times$ 8 tasks $\times$ 3 repetitions = 1296 trials.

4.2 Apparatus and Participants

We used a 27" Apple Thunderbolt Display set to its default resolution (2560$\times$1440 pixels). The web user interface was implemented in Django and visualizations were generated with D3 [5] and Vega [57]. We made sure that all visualizations were of similar size by keeping their width consistent (adjusting height to keep the original aspect ratio). All visualizations fit comfortably on the screen and did not require scrolling. More specifically, the dimensions were 1350$\times$996 pixels for `GlyphSM` and `DorlingSM`, and 1350$\times$849 pixels for `Barchart1M`.

We recruited 18 participants before starting the experiment, a number that allowed us to counterbalance technique presentation order. We continuously recruited participants until we arrived at this pre-defined number. Our participant exclusion criteria included: not completing all conditions, or failing any of the 3 training trials. Given the complexity of the task, we assumed task learning would transfer across techniques. Thus, an excluded participant would have to be replaced with another participant with an equivalent configuration of technique, dataset and task presentation ordering. We had to replace a single participant who declared during the second session that she had misunderstood how to perform the tasks in the first session.

Fig. 4. Web interface used to conduct the experiment. Visualization = `GlyphSM`; task performed on a Time Interval, for All Locations.

From the final 18 participants (10 female and 8 male), none reported any color deficiency. All had normal or corrected-to-normal vision. Age ranged from 23 to 40 ($M = 27.6$, $SD = 4.9$) and most of them were students (13/18) from either a PhD or a Masters' program. Their backgrounds were mainly HCI, Computer Science and Visualization. They were all volunteers, and did not receive any monetary compensation.

4.3 Procedure

First pilots of our study showed that conducting the tasks was mentally demanding. We thus divided the study in three sessions, one per visualization, performed on three different days (that could be consecutive and at most 9 days apart). Each session consisted of three parts: introduction, training, and main trials. In the first session, participants signed a consent form, were told that they could withdraw at any time, and filled out a demographic questionnaire.

1) Introduction and training. The experimenter explained the visualization to be used in the session, along with examples of how correlation and monotonic evolution looked on it. Further training was conducted, that consisted in answering eight trials, one per task (described next). After finishing each trial, the system would indicate if the answer was correct or not. If participants made no error and declared that they had no further question, they would start the main trials. Otherwise, the experimenter would add further training trials.

2) Completion of main trials. Fig. 4 shows a trial screenshot. On the left are the overall progress, the question asked and possible answers. On the right is the generated visualization for that condition. Before each trial a map was shown, highlighting the location(s) that the trial would be about. Our aim was to reduce potential bias due to prior knowledge of the United States' geography, and to ensure there was no ambiguity about geographical features to consider such as, *e.g.*, which states constitute a region.

Participants completed 24 main trials per session (visualization). In this phase, they did not get any feedback about the correctness of their answers. They were instead asked to report the level of confidence in the answer they had just given (low, medium, high).

Once trials were completed for a visualization, participants filled out a post-hoc questionnaire about the strategies used to complete the eight tasks, and how easy it was to complete each one of those tasks. After finishing the third session, participants filled out a final questionnaire, in which they were asked to rank the visualizations. A representative image of each visualization was displayed in the form to help participants remember them. The entire experiment (3 sessions) took approximately one hour and a half.

4.4 Measures

For each task, we defined three metrics, two objective, one subjective:

- `Task completion time:` measured from the moment participants saw the trial screen until they submitted an answer. We computed the average over the 3 repetitions.
- `Error rate:` computed as the number of incorrect answers per task multiplied by the total number of repetitions.

- `Self-reported confidence:` measured on a 3-point Likert scale (high, medium, low).

For each technique, we recorded:

- `Strategies to complete the trials:` described as free text.
- `Self-reported difficulty to complete each type of task:` measured on a 5-point Likert scale from very easy (5) to very difficult (1).

5 RESULTS

We analyze, report, and interpret all our inferential statistics using graphically-reported point estimates and interval estimates [11, 13].

We report sample means for **Completion Time** and **Error Rate** and 95% confidence intervals (CIs), indicating the range of plausible values for the population mean. For our inferential analysis we use means of differences and their 95% confidence intervals (CIs).[3] We use BCa bootstrapping to construct all confidence intervals (10,000 iterations). Since in our *per-task* analysis we test specific predictions rather than a universal null hypothesis, no correction for multiple comparisons was performed [10, 49]. A p-value approach of our technique can be obtained following the recommendations from Krzywinski and Altman [36]. Finally, we also report percentages for self-reported **Confidence** results.

We analyzed a total of 1296 trials (18 participants × 72 trials). All reported analyses were planned before the experiment started.

We first provide an overview across tasks.[4] Since our hypotheses are task dependent, we then perform a detailed per-task analysis.

5.1 Overall results across tasks

Completion Time: Fig. 5 shows completion times of all tasks collectively. Mean times per technique are on the left, mean differences on the right. Mean times are shorter for `GlyphSM` (23.7sec) followed by `DorlingSM` (26sec) and `Barchart1M` (30.7sec). There is strong evidence that `Barchart1M` is slower than `GlyphSM` (by 7.0sec on average) and evidence that it is also slower than `DorlingSM`, although the difference is smaller (4.5sec on average).

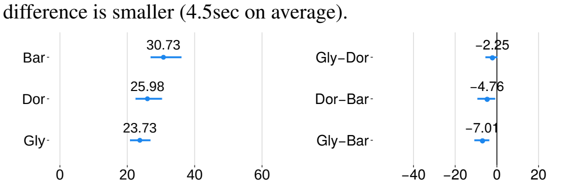


Fig. 5. Left: Mean Completion Time in seconds for each visualization, for all tasks. Right: Pairwise comparisons for each visualization. Error bars represent 95% Bootstrap confidence intervals.

Error Rate: Fig. 6 shows error rates for all tasks collectively, with mean error rates per technique on the left and mean differences on the right. Mean error rates are lower for `GlyphSM` (7.4%) followed by `DorlingSM` (8.1%) and `Barchart1M` (8.6%). There is no evidence that error rates were different across techniques. Thus the main differentiation we can make across techniques comes from completion time.

Confidence: Fig. 7 shows the self-reported confidence for each visualization, for all tasks. Confidence is high for all three visualizations in more than half the trials, although more so for `GlyphSM` (64% of trials) than for `DorlingSM` (57%) and `Barchart1M` (53%).

5.2 Results per task

Next we report results per task, grouped by *temporal granularity* for legibility purposes (analyses were performed per task). The values

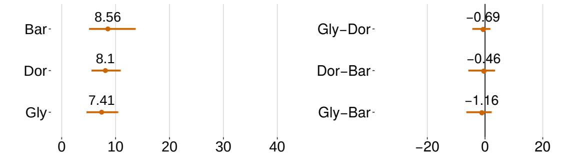


Fig. 6. Left: Mean Error Rate in % for each visualization, for all tasks. Right: Pairwise comparisons for each visualization. Error bars represent 95% Bootstrap confidence intervals.

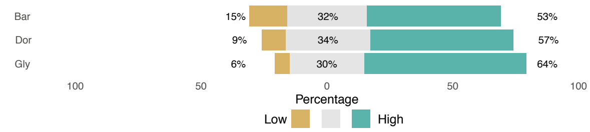


Fig. 7. Self-reported confidence across all tasks per visualization.

and CIs for means and differences of means for both Completion Time and Error Rate can be seen separately for each task in Fig. 8, with the direction of our hypothesis indicated by a gray background. Self-reported Confidence per task can be seen collectively in Fig. 9.

5.2.1 SINGLETIME correlation tasks

In tasks involving a single time step we expect small multiples techniques to fare better (**H1**). Completion times and error rates (means and CIs) for these tasks are found in the leftmost column of Fig. 8.

Completion Time: is faster with small multiples (`GlyphSM`, `DorlingSM`) and slower for `Barchart1M` for both geographic granularity tasks. Looking at mean differences, there is strong evidence that `Barchart1M` is slower than both small multiples techniques, by more than 27sec for REGION, and by more than 32sec for ALLLOCATIONS tasks. Results are inconclusive for the difference between `GlyphSM` and `DorlingSM` in both tasks.

Error Rate: Similar to the results for completion time, for both REGION and ALLLOCATIONS tasks, `GlyphSM` had the best performance, followed by `DorlingSM` and `Barchart1M` with the highest error rate. Looking at mean differences, there is strong evidence that `Barchart1M` is more error prone than `GlyphSM` for both types of geographic granularities. There is weak evidence that `Barchart1M` is also more error-prone than `DorlingSM` for REGION (but no evidence of a difference for ALLLOCATIONS). Finally, `DorlingSM` appears more error-prone than `GlyphSM` for both tasks (strong evidence of this difference for REGION, and weak for ALLLOCATIONS).

Confidence: (self-reported by participants) corroborates these findings. For both tasks that considered SINGLE TIME, confidence is high for small multiples techniques (`GlyphSM` and `DorlingSM`) but low for `Barchart1M` (see top of Fig. 9).

Summary for SINGLETIME: Overall, the tendencies for the two tasks that focus on correlations for a SINGLE TIME are similar, irrespective of whether we consider a geographical region or all locations. We have evidence that using the small multiples visualizations (`GlyphSM`, `DorlingSM`) takes less time (less than 20sec) and causes less errors than `Barchart1M`, supporting **H1**. There is also evidence of differences between `GlyphSM` and `DorlingSM` when it comes to errors, with `DorlingSM` being more error prone, supporting **H4**.

5.2.2 TIME INTERVAL correlation tasks

In time interval tasks, we expect different performance across geographic granularities (**H2**), with a single map (`Barchart1M`) faring better for tasks involving one location (**H2.1**), and small multiples faring better for tasks involving a region or all locations (**H2.2**). Completion times and error rates (means and CIs) are found in the middle column of Fig. 8.

Completion Time: When considering ONELOCATION, we observe that completion time is indeed on average lower for `Barchart1M` (22.2sec), followed by `GlyphSM` (25.8sec) and then `DorlingSM` (29.3sec). Looking at the mean differences, there is evidence that `Barchart1M` is faster than `DorlingSM` (by 7sec on average). It may

[3] A CI of *differences* that does not cross 0 provides evidence of differences - the further away from 0 and the smaller the CI the stronger the evidence.

[4] We counterbalanced visualization order across participants to mitigate learning (Sec. 4.1). An unplanned analysis indicates that although participants improved over sessions (performed best in the 3rd visualization presented than in the 1st), there was indeed no evidence of asymmetric learning across `Barchart1M` and `GlyphSM`, thus counterbalancing worked for them (there is some Time improvement for `DorlingSM`). Analyses/charts are available as supplementary material.

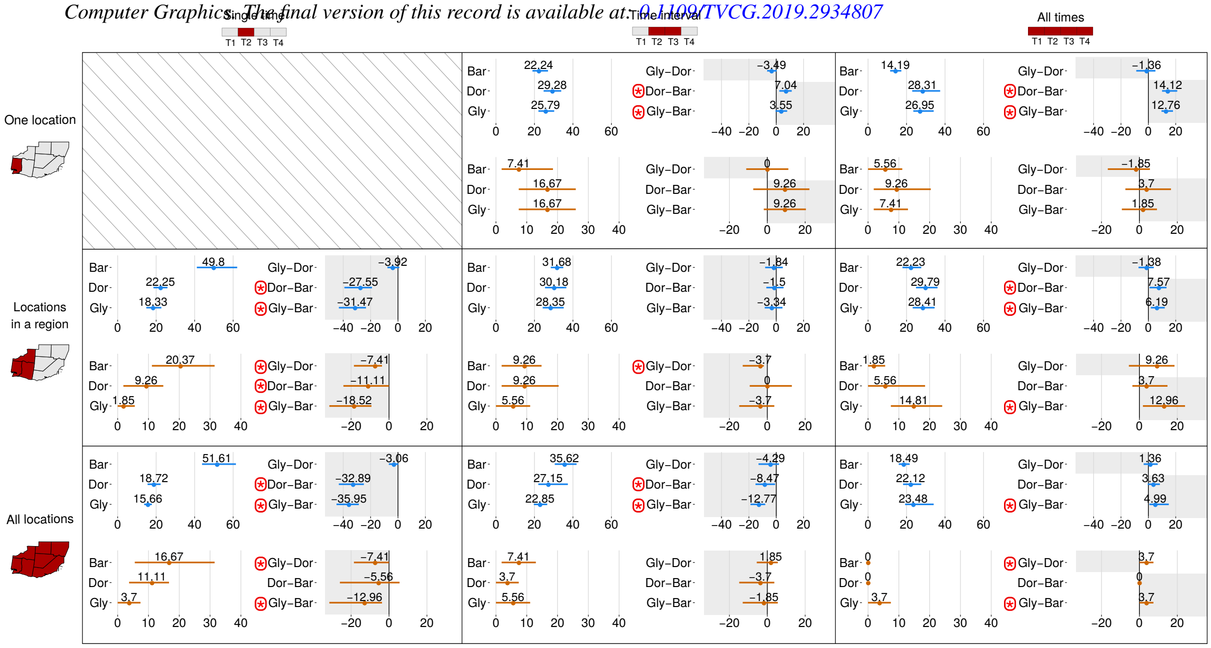


Fig. 8. Results for Completion Time (sec) and Error Rate (in %) for each task in Fig. 2. In each cell (task), Mean values per visualization are seen on the left and means of pairwise differences on the right. Error bars represent 95% Bootstrap confidence intervals. Gray rectangles indicate the direction of our hypotheses. Evidence of differences are marked with a ⊛ (the further away from 0 and the tighter the CI, the stronger the evidence).

be the case that `Barchart1M` is also faster than `GlyphSM` and that `GlyphSM` is faster than `DorlingSM`, but evidence is not conclusive.

The completion time for REGION is close for all three techniques (`GlyphSM` 28.4sec, `DorlingSM` 30.2sec, and `Barchart1M` 31.7sec) and we do not have evidence of differences looking at the mean differences. The same pattern is found in ALLLOCATIONS as `GlyphSM` (22.8sec) is faster than the other techniques, followed by `DorlingSM` (27.1sec) and `Barchart1M` (35.6sec). Looking at the mean differences, we have evidence that `Barchart1M` is slower than both `GlyphSM` and `DorlingSM` (by 12.7sec and 8sec on average). We do not have evidence of a difference between `GlyphSM` and `DorlingSM`.

Error Rate: For these tasks, we observe that the lowest error rate depends on the geographical granularity considered. `Barchart1M` is better for ONELOCATION (7.4%), `GlyphSM` for REGION (5.6%) and `DorlingSM` for ALLLOCATIONS (3.7%). Looking at mean differences for ALLLOCATIONS there is indeed evidence that `DorlingSM` is more error prone than `GlyphSM` (by 3.7% on average) for REGION, but no evidence of other differences.

Confidence: The second row of Fig. 9 shows the self-reported confidence for TIME INTERVAL. We observe that confidence for ONELOCATION is high in more than half of the trials for `Barchart1M` and `GlyphSM` (over 60%), but lower for `DorlingSM` (45%). For tasks in REGION and ALLLOCATIONS, we observe that it is higher for both `GlyphSM` and `DorlingSM` (over 60%) and lower for `Barchart1M` (54% and 50% respectively).

Summary for TIMEINTERVAL: The tendencies for the three tasks that focus on correlations for a time interval change significantly depending on the spatial granularity. For a single location, `Barchart1M` is faster than the small multiple techniques (`GlyphSM`,`DorlingSM`), supporting **H2.1**. This behavior is reversed when considering all locations on the map. `Barchart1M` becomes the slowest visualization, supporting the part of **H2.2** related to all locations. In both tasks, we found no evidence of difference in error rates. The situation is less clear when multiple locations in a region have to be considered. We found no evidence of differences for any of the measures, contrary to the prediction of **H2.2** related to geographical regions. We observe no difference between `GlyphSM` and `DorlingSM`. **H4** is thus not supported.

5.2.3 ALL TIME

In tasks involving all time steps we expect a single map (*i.e.*, `Barchart1M`) to fare better (**H3**). Completion times and error rates (means and CIs) for these tasks are in the rightmost column of Fig. 8.

Completion Time: is lower with `Barchart1M` than with both small-multiples visualizations. Looking at the mean differences, there is strong evidence that `Barchart1M` is faster than `GlyphSM` and `DorlingSM` for both ONELOCATION and REGION tasks. For ALLLOCATIONS task, there is also strong evidence that `Barchart1M` is faster than `GlyphSM` (by 4.9sec on average) but evidence is not conclusive regarding `Barchart1M` being faster than `DorlingSM`. There is no evidence of a difference between `GlyphSM` and `DorlingSM` for any geographical granularity.

Error Rate: is lowest in `Barchart1M` for ONELOCATION and REGION tasks. For ALLLOCATIONS, the error rate is 0% for both `Barchart1M` and `DorlingSM` (and thus, no CI is computed). There is evidence that `Barchart1M` is less prone to errors than `GlyphSM` for REGION, but this evidence is weak for ALLLOCATIONS (and we see no evidence of a difference for ONELOCATION). There is also weak evidence that `DorlingSM` is also less error prone than `GlyphSM` (by 3.7%) for ALLLOCATIONS.

Confidence: is high in over 60% of trials for most visualizations and geographic granularities, with high-confidence trials for `DorlingSM` being a bit lower (around 50% of trials) for the ONELOCATION and REGION.

Summary for ALLTIME: The tendencies for the three tasks that focus on correlations over all time steps are fairly similar, with `Barchart1M` being generally faster than small multiples (`GlyphSM`, `DorlingSM`), thus supporting **H3**. Again, we do not find evidence of a difference between `GlyphSM` and `DorlingSM`. **H4** is not supported.

6 PER-TASK DISCUSSION AND DESIGN RECOMMENDATIONS

We observed that, overall, small multiples were faster across tasks, but their error rates were not different from those of a map with bar charts. Nevertheless, as hypothesized, looking at the individual tasks we see that the performance changes depending on the task at hand. Next,

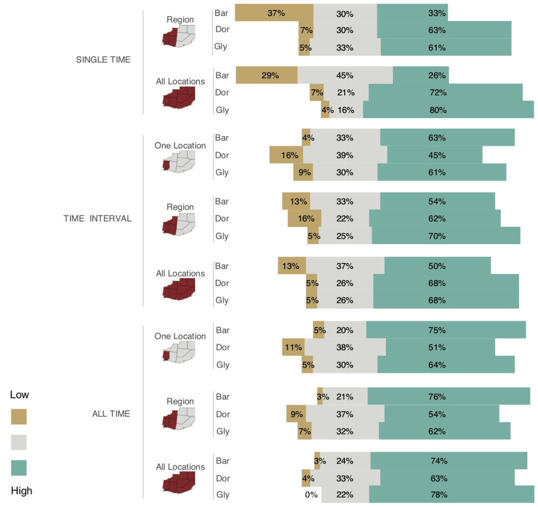


Fig. 9. Reported self-confidence per task (in %).

we summarize and discuss our findings, and distill them into design recommendations (summarized in Fig. 10).

SINGLETIME: The correlation of thematic variables on geographic maps has been studied before for a single point in time [15, 18]. We add to these findings, by identifying that the tendencies for correlation tasks on a single point in time are similar, irrespective of whether we consider a geographical region or all locations. Using the small multiple visualizations (`GlyphSM`, `DorlingSM`), participants were almost twice as fast as when using a single map with bar charts (`Barchart1M`), as they only needed to focus on a single cell of the small multiples, since that cell juxtaposes all spatial information for that point in time. The tasks are slower with a single map with bar charts (`Barchart1M`), since participants needed to visually search for the specific time step across multiple bar charts and synthesize their findings. Error rates for these tasks follow similar trends. Our findings thus confirm **H1**.

When it comes to small multiples, there is a tendency for the proportional symbol map (`GlyphSM`) to be less error prone than the Dorling cartogram (`DorlingSM`), supporting **H4**. This is likely the case because the position of symbols shifts between multiples in the cartogram case, making it hard to re-identify them. This tendency was also observed when comparing proportional symbol maps with non-contiguous cartograms in the work of Gao *et al.* [18]. However, in their case, it was for the overall performance over multiple tasks, not just for correlation identification, and the differences observed were not significant.

These tendencies were consistent with the self-perceived difficulty of conducting these tasks in the exit questionnaire. It was stated often that it is hard to make analyses for one time step with bar charts.

R1: *For identifying correlations at a specific point in time, small-multiples visualizations are better.*

TIMEINTERVAL: When participants have to identify correlation tendencies and evolution over a time interval, the situation is less clear. The tendencies change significantly depending on the geographic granularity (consistent with **H2**). When considering a single location, a single map with bar charts (`Barchart1M`) is faster than the small-multiples techniques (`GlyphSM`, `DorlingSM`), as all temporal information is grouped closely together and participants just needed to identify the temporal interval on a single bar chart. Whereas for small multiples, after identifying the relevant time cells, participants needed to then identify, in each cell, the specific location and collate their findings. This is consistent with **H2.1**.

The findings are reversed when considering all locations, consistent with **H2.2**. Here, a map with bar charts is slower, because it is the visualization where information is scattered and needs to be collated from across different areas. Participants first had to go through all (or almost all) bar charts to identify the specific interval, and collate the information to identify tendencies. Whereas for small multiples, they only needed to focus on a few time steps and look for overall patterns.

One of the most interesting findings from this study is the inconclusive evidence for tasks where a geographic region has to be considered across a time interval (this part of **H2.2** is not confirmed). The lack of observed differences may be due to low statistical power. But we believe it is more likely due to this task being more balanced in the amount of information that needs to be collated across different areas for the different techniques. Here, for a single map with bar charts, participants still had to identify the specific bars across multiple bar charts – but not all of them. When using small multiples, they could focus on a few time steps, but still had to identify the desired geographic area in each one of them. There is likely a tradeoff when it comes to tasks that involve spatial regions and time intervals. When considering subsets of time, it looks like the less spatial locations have to be considered, the better a single map is. Inversely, the more spatial locations, the better small multiples become. More generally, it is likely that a single map with bar charts likely works best for simple geography and complex temporal patterns, and small multiples when geography is complex but the temporal variability is simple. Future work needs to determine exactly when to transition between visualizations. We are not aware of any previous work that has considered correlation tasks that require gathering information across subsets of space and time.

R2: *For identifying correlations and temporal evolution over a subset of time steps and a subset of locations, there is no clear winner. If there are only a few locations, consider using a single map with bar charts. If there are many locations, prefer small multiples.*

ALL TIME: The tendencies for the three tasks that focus on correlations for all time steps are again consistent, with `Barchart1M` being faster, in accordance with **H3**. Even though participants had to collate both spatial and temporal information, a single map with bar charts was faster. This representation makes it easy to see trends over time (correlation and monotonic evolution) that are juxtaposed in the individual bar charts. Collating this information seems to be fast irrespective of how many geographic regions are taken into account. Small multiples seem slower, likely because determining temporal trends necessitates comparing several locations across cells before identifying a trend.

The self-perceived difficulty to conduct the task for all time steps was also consistent with objective measures. A single map with bar charts was perceived, overall, as easier to use than both small-multiples visualizations, and several participants commented that it was easy to observe evolution over time on the single map with bar charts.

R3: *For identifying correlations and temporal evolution over all time steps, irrespective of the number of locations, a single map with bar charts is better.*

Small multiples: We found evidence that the two small-multiple techniques (`GlyphSM`, `DorlingSM`) were different mainly when considering a single point in time (partially confirming **H4**), with `DorlingSM` being slower and more error prone. Participants' comments indicate that they had difficulty matching a location, or sets of locations, across small multiples with `DorlingSM`, since positions of locations shifted. Nevertheless, this cost is not seen in tasks considering more than one time step. This may be due to low statistical power, or because this cost is small when it comes to more challenging tasks (time intervals or all time steps) that require collating information across small multiples.

R4: *For small multiples, there is some evidence that proportional symbol maps are better than Dorling, especially for a single time point.*

7 General Discussion

Our findings generally followed our original hypothesis. We thus believe that our reasoning, that difficulty in each technique depends on whether the information to be collated is juxtaposed or distant, is sound; and that our results reflect true tendencies.

The one exception relates to correlation tasks on subranges in time and space. We had originally thought that small-multiple variations would prevail in this situation, but we were unable to detect a trend. We believe that our setup of this task may reflect a similar difficulty in collating temporal information (for small multiples) and spatial

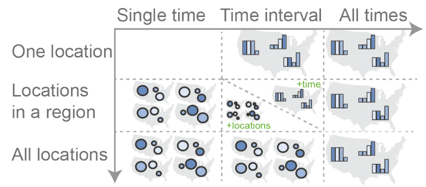


Fig. 10. Summary of our recommendations for the different tasks. For tasks on subranges of space and time (middle cell) there is no clear winner, but the table structure suggests small multiples work best for increased spatial complexity, and a single map with bar charts for increased temporal complexity (although the transition point is not known).

information (for the single map with bar charts). Looking at Fig. 10, we observe that this middle cell is a transition cell between tasks better suited to small multiples and tasks better suited to single maps. For example, if we look at the middle column (Time interval), it seems that one (and likely a few) geographic locations are best seen on single maps, but as more locations are added small multiples become more competitive. Or if we look at the middle row (Locations in a region) it seems that one (and likely a few) time steps may be better viewed with small multiples, but as time increases a single map with bar charts becomes better. It is interesting to consider what are the tipping points of these shifts (number of time steps, number of geographic locations), in order to determine when to transition between visualizations.

For all visualizations, collating information across different areas (bars from different bar charts for the single map, and locations across cells for the small multiples) is challenging. In our study, we focus on static visualizations, but the addition of highlighting would likely reduce the differences we found, by making it easier to collate information (*e.g.*, highlight Washington in all small multiples, or 2011 in all bar charts). Nevertheless, we believe the high-level effects would still hold (to a lesser extent) as they are due to the fact that information is dispersed across the visualization. If filtering is considered, we believe behavior will likely revert to the results at the corners of Fig. 10. For example, filtering on time interval 2009-2011 would either remove or fade other years out, making this a task similar to ones over all time steps. Similarly, if the East US is the focus, the system would remove or fade other locations out, making this a task similar to those involving all locations. More importantly, the actions performed to select or filter time steps or geographical locations could themselves be used as an indication of what is the user's focus, and used to transition to the best visualization for the task.

7.1 Limitations and Future Work

Interaction was deliberately not considered in this first study, as we primarily aimed at evaluating the specific influence of space and time at different granularities on users' ability to identify correlations with different visualizations. Thus, we wanted to avoid adding further factors to an already complex experimental design. Our discussion section above provides initial thoughts about how interaction could affect our results, but further work is needed to verify them, and to consider the use of interaction as a means to transition between visualizations.

The number of steps used to detect correlation in our tasks is limited (nine time steps per location), which required us to use datasets with a strong relationship between two variables. Data extracted from measures of real world phenomena is unlikely to present such strong patterns, making it harder to detect potential correlations. This could alter our results, although we believe the general trends would persist.

Another limitation is that we only used a single map of the US, which necessarily captures only a subset of geographical configurations. It is possible that countries with more diverse shapes (*e.g.* Chile, Italy or United Kingdom) would lead to different results, as the identification of individual locations or regions might be different. We attempted to mitigate any bias in identifying the locations of interest by displaying, prior to each trial, the geographic region of interest. Nevertheless, further experimentation is needed. Moreover, diversity of irregularity of locations can impact spatial autocorrelation in geospatial visualizations that use irregular geometries to represent thematic variables, such as choropleths [4, 42, 67]. While it is possible that effects might differ somewhat in other types of maps, we feel that the general trends should hold: our techniques use regular shapes to represent thematic variables, and thus the size and number of items compared likely weigh more in the complexity of the task (*e.g.*, occlusion or clutter of elements might impact the interpretation of patterns). To this end, in our trials we varied the size of locations and their density. Finally, although the analysis of data using a map of a known country could have led to bias given preconceptions about the geographical distribution, we believe this to be unlikely given the extensive training, and the number of map features and time steps involved.[5] In summary, while we believe that overall trends would persist across different maps, future work needs to consider more diverse geographic maps.

For the small multiples tested, we expected that Dorling cartograms (*i.e.*, visualizations that use visual channels of the maps features themselves to encode thematic variables) would perform worst than proportional symbol maps, as was the case in previous work [18, 30]. In our context this was observed mainly when considering tasks at a single point in time. It is very likely this effect will be more pronounced in other spatial tasks that involve more continuous geographic changes and correlations that vary spatially (*e.g.*, identifying transmission patterns).

We recruited users who were already knowledgeable about visualization, and gave them additional training. Opportunities for such training may not be available to the general public. While we believe general trends will still apply, it is possible that non-trained users would have lower accuracy rates or would not dedicate as much time as our participants to perform the tasks. Additionally, they might be more familiar with one of the three tested techniques, which would bias results in its favor. Future work should investigate the learning curve of each technique and analyze how well they fare when used by novices with a more diverse background and lack of training. A next step in that direction would be to conduct a crowdsourced study.

Finally, we decided to combine two different association tasks in one (*i.e.*, the type of correlation and its evolution), as we felt they were tightly coupled when performed in the context of geo-temporal analysis. Due to this combination, our analysis does not provide finer details on the difficulty of each subtask. Future work could study each one separately to gain more insights about how correlation and trends are detected individually. For example, we expect that complex temporal tasks, such as detecting and characterizing monotonic evolution, is easier on single maps with bar charts (as each one directly encodes this evolution); whereas complex geographical tasks, such as detecting transmission patterns, may be easier with small multiples.

8 Conclusions

We presented a study on identifying correlation in spatio-temporal visualizations. We considered eight tasks that associate two variables over different granularity levels for both time and space. The compared visualizations combine different strategies to represent thematic variables: juxtaposing either time or space (a single map with bar charts *vs.* small-multiple maps); encoding variables either using symbols overlaid on top of map features, or using visual channels of the map features themselves (proportional symbol maps *vs.* cartograms). We provide a set of design recommendations depending on the task at hand. In our context, the technique using the map features' own visual channels to encode thematic variables (cartograms) performed worst only when a single point in time was considered. Our results further indicate that for tasks that consider the evolution over all time steps, a visualization that represents data on a single map (juxtaposing time) is more effective and easier to interpret than small multiples. Small multiples (juxtaposing space) are better suited for tasks that require the comparison of variables for one point in time over several geographical locations. When dealing with time intervals and spatial regions, our results suggest that there is a continuum of performance between visual representations (juxtapose time *vs.* space), raising questions for future research.

[5] We did not find any warning signs of such pitfalls (*e.g.*, participants taking very little time to finish the tasks and making numerous errors).

References

[1] M. J. Alam, S. G. Kobourov, and S. Veeramoni. Quantitative measures for cartogram generation techniques. *Comput. Graph. Forum*, 34(3):351–360, June 2015. doi: 10.1111/cgf.12647

[2] L. An, M.-H. Tsou, S. E. Crook, Y. Chun, B. Spitzberg, J. M. Gawron, and D. K. Gupta. Space–time analysis: Concepts, quantitative methods, and future directions. *Annals of the Association of American Geographers*, 105(5):891–914, 2015. doi: 10.1080/00045608.2015.1064510

[3] N. Andrienko and G. Andrienko. Interactive visual tools to explore spatio-temporal variation. In *Proc. of the working conference on Advanced visual interfaces (AVI)*, pp. 417–420. ACM, 2004. doi: 10.1145/989863.989940

[4] R. Beecham, J. Dykes, W. Meulemans, A. Slingsby, C. Turkay, and J. Wood. Map LineUps: Effects of spatial structure on graphical inference. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):391–400, Jan. 2017. doi: 10.1109/TVCG.2016.2598862

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

[6] I. Boyandin, E. Bertini, and D. Lalanne. A qualitative study on the exploration of temporal changes in flow maps with animation and Small-Multiples. *Comput. Graph. Forum*, 31(3pt2):1005–1014, June 2012. doi: 10.1111/j.1467-8659.2012.03093.x

[7] H. Chernoff. The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association*, 68(342):361–368, 1973. doi: 10.1080/01621459.1973.10482434

[8] F. Chevalier, N. H. Riche, C. Plaisant, A. Chalbi, and C. Hurter. Animations 25 years later: New roles and opportunities. In *Proc. of the International Working Conference on Advanced Visual Interfaces*, AVI '16, pp. 280–287. ACM, 2016. doi: 10.1145/2909132.2909255

[9] P. Craig, N. R. Seïler, and A. D. O. Cervantes. Animated geo-temporal clusters for exploratory search in event data document collections. In *18th International Conference on Information Visualisation (IV)*, pp. 157–163. IEEE, 2014. doi: 10.1109/IV.2014.69

[10] G. Cumming. The new statistics: Why and how. *Psychological Science*, 25(1):7–29, 2014. doi: 10.1177/0956797613504966

[11] G. Cumming and S. Finch. Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist*, pp. 170–180, 2005. doi: 10.1037/0003-066X.60.2.170

[12] D. Dorling. *Area Cartograms: Their Use and Creation*, chap. 3.7, pp. 252–260. John Wiley & Sons, Ltd, 2011. doi: 10.1002/9780470979587.ch33

[13] P. Dragicevic. Fair statistical communication in HCI. In J. Robertson and M. Kaptein, eds., *Modern Statistical Methods for HCI*, chap. 13, pp. 291–330. Springer International Publishing, 2016. doi: 10.1007/978-3-319-26633-6_13

[14] Y. Du, L. Ren, Y. Zhou, J. Li, F. Tian, and G. Dai. Banded choropleth map. *Pers. Ubiquit. Comput.*, 22(3):503–510, June 2018. doi: 10.1007/s00779-018-1120-y

[15] M. E. Elmer. Symbol considerations for bivariate thematic maps. *In Proceedings of the 26th International Cartography Conference (ICC)*, 2013.

[16] C. Fish, K. P. Goldsberry, and S. Battersby. Change blindness in animated choropleth maps: An empirical study. *Cartogr. Geogr. Inf. Sci.*, 38(4):350–362, Jan. 2011. doi: 10.1559/15230406384350

[17] G. Fuchs and H. Schumann. Visualizing abstract data on maps. In *Proceedings of the 8th International Conference on Information Visualisation (IV)*, pp. 139–144. IEEE, 2004. doi: 10.1109/IV.2004.1320136

[18] P. Gao, Z. Li, and Z. Qin. Usability of value-by-alpha maps compared to area cartograms and proportional symbol maps. *Journal of Spatial Science*, pp. 1–17, Mar. 2018. doi: 10.1080/14498596.2018.1440649

[19] P. Gatalsky, N. V. Andrienko, and G. L. Andrienko. Interactive analysis of event data using space-time cube. *Proceedings of the 8th International Conference on Information Visualisation (IV)*, pp. 145–152, 2004. doi: 10.1109/IV.2004.1320137

[20] S. Goodwin, J. Dykes, A. Slingsby, and C. Turkay. Visualizing multiple variables across scale and geography. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):599–608, Jan. 2016. doi: 10.1109/TVCG.2015.2467199

[21] A. L. Griffin, A. M. MacEachren, F. Hardisty, E. Steiner, and B. Li. A comparison of animated maps with static small-multiple maps for visually identifying space-time clusters. *Annals Assoc. of American Geographers*, 96(4):740–753, 2006. doi: 10.1111/j.1467-8306.2006.00514.x

[22] D. Guo, J. Chen, A. M. MacEachren, and K. Liao. A visualization system for space-time and multivariate patterns (vis-stamp). *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1461–1474, 2006. doi: 10.1109/TVCG.2006.84

[23] H. Hagh-Shenas, S. Kim, V. Interrante, and C. Healey. Weaving versus blending: a quantitative assessment of the information carrying capacities of two alternative methods for conveying multivariate data with color. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1270–1277, Nov. 2007. doi: 10.1145/1179622.1179848

[24] R. Han, Z. Li, P. Ti, and Z. Xu. Experimental evaluation of the usability of cartogram for representation of GlobeLand30 data. *ISPRS International Journal of Geo-Information*, 6(6):180, 2017. doi: 10.3390/ijgi6060180

[25] R. L. Harris. *Information Graphics: A Comprehensive Illustrated Reference*. Oxford University Press, 2000.

[26] L. Harrison, F. Yang, S. Franconeri, and R. Chang. Ranking Visualizations of Correlation Using Weber's Law. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1943–1952, Dec. 2014. doi: 10.1109/TVCG.2014.2346979

[27] M. Harrower. The cognitive limits of animated maps. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 42(4):349–357, 2007. doi: 10.3138/carto.42.4.349

[28] J. Jo, F. Vernier, P. Dragicevic, and J. Fekete. A declarative rendering model for multiclass density maps. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):470–480, Jan 2019. doi: 10.1109/TVCG.2018.2865141

[29] T. Johnson, C. Acedo, S. Kobourov, and S. Nusrat. Analyzing the evolution of the internet. In *17th IEEE Eurographics Conference on Visualization (EuroVis)*, vol. 17, 2015. doi: 10.2312/eurovisshort.20151123

[30] S. Kaspar, S. I. Fabrikant, and P. Freckmann. Empirical study of cartograms. In *25th International Cartographic Conference (ICC)*, vol. 3, p. 5, 2011.

[31] M. Kay and J. Heer. Beyond weber's law: A second look at ranking visualizations of correlation. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):469–478, Jan. 2016. doi: 10.1109/TVCG.2015.2467671

[32] S. Kim, S. Jeong, I. Woo, Y. Jang, R. Maciejewski, and D. Ebert. Data flow analysis and visualization for spatiotemporal statistical data without trajectory information. *IEEE Transactions on Visualization and Computer Graphics*, PP(99):1–1, 2017. doi: 10.1109/TVCG.2017.2666146

[33] S. Kim, R. Maciejewski, A. Malik, Y. Jang, D. S. Ebert, and T. Isenberg. Bristle maps: a multivariate abstraction technique for geovisualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(9):1438–1454, Sept. 2013. doi: 10.1109/TVCG.2013.66

[34] A. Klippel, F. Hardisty, and R. Li. Interpreting spatial patterns: An inquiry into formal and cognitive aspects of Tobler's first law of geography. *Annals of the Association of American Geographers*, 101(5):1011–1031, 2011. doi: 10.1080/00045608.2011.577364

[35] M.-J. Kraak. The space-time cube revisited from a geovisualization perspective. In *Proc. 21st International Cartographic Conference (ICC)*, pp. 1988–1996, 2003.

[36] M. Krzywinski and N. Altman. Points of significance: Error bars. *Nature Methods*, 10(10):921–922, Oct. 2013. doi: 10.1038/nmeth.2659

[37] J. Li, Z.-P. Meng, M.-L. Huang, and K. Zhang. An interactive radial visualization of geoscience observation data. In *Proceedings of the 8th International Symposium on Visual Information Communication and Interaction*, VINCI '15, pp. 93–102. ACM, 2015. doi: 10.1145/2801040.2801061

[38] Y.-N. Li, D.-J. Li, and K. Zhang. Metaphoric transfer effect in information visualization using glyphs. In *Proceedings of the 8th International Symposium on Visual Information Communication and Interaction*, VINCI '15, pp. 121–130. ACM, 2015. doi: 10.1145/2801040.2801062

[39] M. A. Livingston and J. W. Decker. Evaluation of trend localization with multi-variate visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2053–2062, Dec. 2011.

[40] M. A. Livingston and J. W. Decker. Evaluation of multivariate visualizations: a case study of refinements and user experience. In *Visualization and Data Analysis 2012*, vol. 8294, p. 82940G. SPIE - International Society for Optics and Photonics, Jan. 2012. doi: 10.1117/12.912192

[41] L. McNabb, R. S. Laramee, and R. Fry. Dynamic choropleth maps – using amalgamation to increase area perceivability. In *Proceedings of the 22nd International Conference Information Visualisation (IV)*, pp. 284–293, July 2018. doi: 10.1109/IV.2018.00056

[42] L. McNabb, R. S. Laramee, and M. Wilson. When size matters: towards

evaluating perceivability of choropleths. In *Proceedings of the Conference on Computer Graphics & Visual Computing*, pp. 163–171. Eurographics Association, 2018. doi: 10.2312/cgvc.20181221

[43] T. Nakaya and K. Yano. Visualising crime clusters in a space-time cube: An exploratory data-analysis approach using space-time kernel density estimation and scan statistics. *Transactions in GIS*, 14(3):223–239, 2010. doi: 10.1111/j.1467-9671.2010.01194.x

[44] S. Nusrat, M. J. Alam, and S. Kobourov. Evaluating cartogram effectiveness. *IEEE Transactions on Visualization and Computer Graphics*, 24(2):1077–1090, Feb. 2018. doi: 10.1109/TVCG.2016.2642109

[45] S. Nusrat, M. J. Alam, C. Scheidegger, and S. Kobourov. Cartogram visualization for bivariate Geo-Statistical data. *IEEE Transactions on Visualization and Computer Graphics*, 24(10):2675–2688, Oct. 2018. doi: 10.1109/TVCG.2017.2765330

[46] S. Nusrat and S. Kobourov. The state of the art in cartograms. *Comput. Graph. Forum*, 35(3):619–642, June 2016. doi: 10.1111/cgf.12932

[47] J. H. Park, S. Nadeem, and A. Kaufman. GeoBrick: exploration of spatiotemporal data. *The Visual Computer*, 35(2):191–204, Feb 2019. doi: 10.1007/s00371-017-1461-y

[48] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. Scikit-learn: Machine learning in python. *Journal of machine learning research*, 12(Oct):2825–2830, 2011.

[49] T. V. Perneger. What's wrong with bonferroni adjustments. *BMJ*, 316(7139):1236–1238, 1998. doi: 10.1136/bmj.316.7139.1236

[50] S. Peters and L. Meng. Visual analysis for nowcasting of multidimensional lightning data. *ISPRS International Journal of Geo-Information*, 2(3):817–836, 2013. doi: 10.3390/ijgi2030817

[51] D. J. Peuquet. It's about time: A conceptual framework for the representation of temporal dynamics in geographic information systems. *Ann. Assoc. Am. Geogr.*, 84(3):441–461, Sept. 1994. doi: 10.1111/j.1467-8306.1994.tb01869.x

[52] R. A. Rensink. The nature of correlation perception in scatterplots. *Psychon. Bull. Rev.*, 24(3):776–797, June 2017. doi: 10.3758/s13423-016-1174-7

[53] R. A. Rensink and G. Baldridge. The perception of correlation in scatterplots. *Comput. Graph. Forum*, 29(3):1203–1210, Aug. 2010. doi: 10.1111/j.1467-8659.2009.01694.x

[54] G. Robertson, R. Fernandez, D. Fisher, B. Lee, and J. Stasko. Effectiveness of animation in trend visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1325–1332, Nov 2008. doi: 10.1109/TVCG.2008.125

[55] H. Rosling. The best stats you've ever seen. `https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen`, 2006. TED, Last accessed: 2019-03-27.

[56] R. E. Roth. An empirically-derived taxonomy of interaction primitives for interactive cartography and geovisualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2356–2365, Dec. 2013. doi: 10.1109/TVCG.2013.130

[57] A. Satyanarayan, R. Russell, J. Hoffswell, and J. Heer. Reactive vega: A streaming dataflow architecture for declarative interactive visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):659–668, Jan 2016. doi: 10.1109/TVCG.2015.2467091

[58] R. Scheepens, H. v. d. Wetering, and J. J. v. Wijk. Non-overlapping aggregated multivariate glyphs for moving objects. In *IEEE Pacific Visualization Symposium*, pp. 17–24, Mar. 2014. doi: 10.1109/PacificVis.2014.13

[59] J. Schiewe. Task-Oriented visualization approaches for landscape and urban change analysis. *ISPRS International Journal of Geo-Information*, 7(8):288, July 2018. doi: 10.3390/ijgi7080288

[60] B. Shneiderman. Why not make interfaces better than 3d reality? *IEEE Comput. Graph. Appl.*, 23(6):12–15, Nov. 2003. doi: 10.1109/MCG.2003.1242376

[61] A. Slingsby. Tilemaps for summarising multivariate geographical variation. In *IEEE VIS Poster*, 2018.

[62] J. Stewart and P. J. Kennelly. Illuminated choropleth maps. *Annals of the Association of american Geographers*, 100(3):513–534, 2010. doi: 10.1080/00045608.2010.485449

[63] H. Sun and Z. Li. Effectiveness of cartogram for the representation of spatial data. *The Cartographic Journal*, 47(1):12–21, 2010. doi: 10.1179/000870409X12525737905169

[64] C. Tominski, P. Schulze-Wollgast, and H. Schumann. 3D information visualization for time dependent data on maps. In *Proceedings of the 9th International Conference on Information Visualisation (IV)*, pp. 175–181. IEEE, 2005. doi: 10.1109/IV.2005.3

[65] B. Tversky, J. B. Morrison, and M. Betrancourt. Animation: Can it facilitate? *Int. J. Hum.-Comput. Stud.*, 57(4):247–262, Oct. 2002. doi: 10.1006/ijhc.2002.1017

[66] J. A. Tyner. *Principles of map design*. Guilford Publications, 2014.

[67] M. O. Ward, G. Grinstein, and D. Keim. *Interactive data visualization: foundations, techniques, and applications*. AK Peters/CRC Press, 2015.

[68] C. M. Wittenbrink, A. T. Pang, and S. K. Lodha. Glyphs for visualizing uncertainty in vector fields. *IEEE Transactions on Visualization and Computer Graphics*, 2(3):266–279, Sept. 1996. doi: 10.1109/2945.537309

[69] F. Yang, L. T. Harrison, R. A. Rensink, S. L. Franconeri, and R. Chang. Correlation judgment and visualization features: A comparative study. *IEEE Transactions on Visualization and Computer Graphics*, 25(3):1474–1488, Mar. 2019. doi: 10.1109/TVCG.2018.2810918

[70] B. Yost and C. North. Single complex glyphs versus multiple simple glyphs. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pp. 1889–1892. ACM, 2005. doi: 10.1145/1056808.1057048

[71] J. Zhang, A. Malik, B. Ahlbrand, N. Elmqvist, R. Maciejewski, and D. S. Ebert. TopoGroups: Context-Preserving visual illustration of Multi-Scale spatial aggregates. In *Proceedings of the Conference on Human Factors in Computing Systems*, CHI '17, pp. 2940–2951. ACM, 2017. doi: 10.1145/3025453.3025801

French Abstract - Résumé

French Abstract - Résumé

Les données numériques que notre société génère augmentent chaque année. Les initiatives de la science et du gouvernement favorisent le partage de ces données, qui atteignent facilement des pétaoctets chaque année. Par exemple, le Large Hadron Collider a généré plus de 200 pétaoctets de données. En outre, nous avons de plus en plus accès à des données collectées à titre personnel, par exemple grâce aux API des téléphones portables et des montre connectées. Ces technologies nous donnent aujourd'hui accès à une quantité de données sans précédent.

Néanmoins, la compréhension des données reste difficile, car leur volume dépasse de loin ce qu'une personne peut raisonnablement consommer et comprendre. Les mesures statistiques sont pas suffisantes, car par exemple des données très différents peuvent avoir les mêmes statistiques, comme le démontrent le Quartet d'Anscome et les travaux plus récents de Matejka et Fitzmaurice. Et les approches d'apprentissage automatique peuvent apprendre de fausses corrélations et nécessitent toujours une apprentissage (supervisée ou non); elles ne sont donc pas fiables lorsque les motifs d'intérêt ne sont pas fréquents dans l'ensemble de données. Ainsi, les approches de "sensemaking" basées sur l'inspection visuelle et l'interaction restent une alternative importante.

La visualisation de l'information combine l'interaction homme-machine, la conception visuelle et la théorie de la perception, afin de proposer des représentations visuelles des données qui amplifient la cognition et aident à la compréhension des données. Venant d'un formation sur l'interaction homme-machine (IHM), cette définition de la visualisation interactive a résonné en moi, car elle invoque la définition plus générale d'une interface homme-machine. Une interface est traditionnellement considérée comme le moyen pour les humains et les machines de communiquer. Lorsque j'ai commencé mes recherches sur la visualisation pendant mon post-doc en 2008, j'ai établi un parallèle avec cette définition. Je vois la visualisation d'informations comme un moyen par lequel les êtres humains peuvent communiquer avec leurs données et les machines qui les stockent et les traitent. Donc si on considère la visualisation comme un canal de communication, avec les humains d'un côté et les machines de l'autre, plus la bande passante est élevée, plus la visualisation est efficace. On est aujourd'hui confronté aux limites de ce canal étant donné la quantité de données à notre disposition. Présenter visuellement de grandes quantités d'informations reste un défi, les travaux précédents montrant qu'un million d'éléments affichés sur un écran traditionnel est proche de la limite pratique.

Dans mon manuscrit je présente comment les recherches dans le domain de la visualisation proposent des nouvelles façons d'amplifier cette bande de communication entre les machines et les êtres humains, des deux côtés du canal de communication (humain et machine). Par exemples des travaux regardent le côté des êtres humaines, étudient les limites des êtres humains (sujet central des études sur la perception visuelle); ou bien examinent l'impact de différentes plateformes de visualisation comme les tables ou les écrans muraux, les téléphones portables et les montres intelligentes, ou même les visualisations tangibles. Il est également possible d'augmenter la puissance du côté humaine en permettant à plusieurs utilisateurs de collaborer et de partager leur expertise. De l'autre côté du canal, les nouvelles technologies d'affichage peuvent afficher une plus grande quantité de données. Par exemple, les grands écrans ou les écrans multiples ont un nombre de pixels plus important que les écrans traditionnels. Et les appareils mobiles ou de réalité augmentée peuvent. La combinaison de la visualisation avec des méthodes automatisées peut

également aider à orienter les algorithmes d'exploration de données vers des modèles visuels intéressants, en faisant un meilleur usage des ressources informatiques. Beaucoup des travaux se situent entre les deux côtés, et considèrent à la fois des contraintes humaines et des contraintes de calcul ou de rendu. De nouvelles représentations visuelles et interactions appropriés peuvent distiller les aspects importants des données, rendant moins des pixels mais plus d'informations saillantes pour le spectateur, amplifiant à la fois la largeur de bande machine et humaine.

Parmi ces stratégies, mes propres recherches augmentent ce bande de communication de deux façons :

▷ S'éloigner des ordinateurs de bureau traditionnels pour s'orienter vers des grandes dispositifs collaboratifs (comme les murs d'images) qui peuvent à la fois restituer de plus grandes quantités de données et accueillir plusieurs utilisateurs ;

▷ Trouver des représentations visuelles appropriées et leurs limites, afin de montrer les informations importantes qui peuvent être comprises et exploitées.

Méthodologie de recherche et inspiration J'ai commencé ma carrière de chercheur en interaction homme-machine dans une institution et un laboratoire où l'IHM était un domaine bien établi. Cela m'a donné l'occasion d'accéder à plusieurs cours qui couvraient les méthodes de conception, d'évaluation et d'analyse de l'IHM. Les origines interdisciplinaires de l'IHM font que le domaine comprend une pléthore de méthodologies qui viennent de la psychologie expérimentale et les facteurs humains (recherche plus quantitative, avec des évaluations contrôlées aux utilisateurs), de la sociologie et l'anthropologie (souvent de la recherche qualitative, comme des observations, des entretiens, des enquêtes contextuelles, essayant généralement de comprendre les besoins des utilisateurs), de la conception et le génie logiciel (se concentrant sur le processus de conception, comme la co-conception et le prototypage itératif), etc.

Ainsi, comme la plupart des chercheurs formés en IHM, je suis consciente et ouverte à de plusieurs méthodologies pour répondre aux questions de recherche et analyser mes résultats. J'ai apporté ces méthodes à ma recherche en visualisation, en choisissant mon approche méthodologique en fonction de l'objectif du travail (mentionné dans chaque section et résumé à la fin de chaque chapitre dans le manuscrit). Mes traveaux utilisent donc plusieurs types des méthodes comme des expériences qualitatives contrôlées lorsqu'on essaie d'isoler différents facteurs, de la conception participative et aux observations qualitatives lors de l'introduction d'un nouveau système d'analyse visuelle, etc. Lorsque j'ai commencé de travailler dans le domain de la visualisation en 2008 (sur l'amélioration de la lisibilité des réseaux sociaux), le domaine avait juste commencé à réfléchir profondément aux méthodes d'évaluation (par exemple le workshop BELIV venait d'apparaître dans la conférence AVI et CHI). Cette réflexion active sur les méthodes d'évaluation des systèmes de visualisation, et ma conviction que je pouvais y contribuer, ont été l'une des raisons principales pour lesquelles j'ai me suis orientée vers la recherche en visualisation.

La plupart des recherches en visualisation mentionnées dans mon manuscrit (par moi ou par autres chercheurs) varient en termes de méthodologie de recherche, c'est-à-dire les étapes qu'on suit et les preuves qu'on fournit. Celles-ci sont généralement dictées par les objectifs de chaque recherche. Néanmoins, je dirais que nos objectifs de recherche, tels qu'ils sont énoncés dans nos résultats de recherche respectifs (publications, rapports, systèmes) ne sont pas toujours à l'origine de nos travaux. Plus généralement, il n'est pas toujours évident d'où viennent les idées / inspiration de recherche. Dans chaque chapitres de ce manuscrit, je décris le but de ma recherche, mon méthodologie, mais j'ajoute aussi une réflexion sur l'inspiration de ce travail. Mon but est de commencer un dialogue sur la manière de mieux capturer, documenter, et communiquer l'inspiration, evolution et itération des nos recherches.

4.6 Surfaces collaboratifs

Grâce à leur grande taille et de leur haute résolution, les grands dispositifs collaboratifs tels que les murs et les tables interactifs, peuvent repousser les limites de rendu des écrans de bureau. Ils peuvent aussi naturellement permettre à plusieurs utilisateurs de collaborer et d'explorer des données simultanément. La migration des visualisations interactives vers ces surfaces collaboratifs soulève des questions difficiles. Par exemple, quelles sont les visualisations appropriées pour de tels environnements, comment aider les utilisateurs à explorer et à interagir avec ces données, et plus généralement comment aider la collaboration autour des ces surfaces, en particulier les murs d'images.

Interaction. Une partie de mon travail explore les défis de l'interaction et de la visualisation, en examinant d'abord comment interagir dans un environnement où les souris et les claviers ne sont pas nécessairement appropriés car les utilisateurs souvent se déplace devant le mur. Mon travail sur Smarties [CBF14], réalisé en collaboration avec O. Chapuis et S. Franzeskakis, porte sur le prototypage d'un support d'interaction. Il s'agit d'un toolkit permettant de développer facilement une interface mobile pour des applications murales, permettant aux concepteurs de personnaliser la surface tactile et les widgets des appareils mobiles synchronisés. En tant que boîte à outils, l'utilité de Smarties a été démontrée à l'aide d'exemples de cas d'utilisation et d'applications. Mon travail sur SketchSliders [TBJ15], en collaboration avec T. Tsandilas et T. Jacob, permet aux utilisateurs d'esquisser leurs propres interfaces pour analyser leur données. Cette approche offre une flexibilité aux utilisateurs finaux, qui peuvent esquisser à la volée les composants interactifs dont ils ont besoin pendant leur exploration, en créant des interfaces personnalisées et adaptées. La méthodologie derrière la conception des SketchSliders était différente que celle de Smarties. Nous avons commencé par comprendre dans quelles situations les utilisateurs seraient intéressés par l'esquisse de leur interface, en utilisant une configuration de type "Wizard of Oz", où les participants esquissaient l'interaction qu'ils souhaitaient et où un expérimentateur l'appliquait aux données. Cela nous a permis de recueillir un ensemble concret de dessins, qui ont ensuite été ajoutés à notre prototype. Étant donné la nature créative des croquis, les expériences contrôlées ne sont pas appropriées pour la validation de nos prototypes. Nous avons plutôt mené des séances d'analyse ouvertes avec des experts en visualisation, afin d'observer comment ils utilisent les SketchSliders dans la pratique. Le travail de SketchSliders a reçu une mention honorable à CHI 2015.

Perception. Une autre question qui se pose quand on utilise les murs d'images est la façon dont la grande surface des affichages muraux peut affecter comment nous voyons et comprenons les visualisations. Dans notre travail sur les études de la perception visuelle avec P. Isenberg, nous examinons [BI12] trois variables visuelles (longueur, surface, angle) qui sont considérées comme des éléments constitutifs de visualisations complexes. Et nous montrons que notre perception de ces variables change en fonction de notre emplacement devant le mur d'écran. Cela soulève des questions sur la façon dont nous devons coder et visualiser les visualisations sur ses dispositifs, et souligne l'importance du mouvement physique, car il peut aider à corriger cette distorsion. La méthodologie suivie dans notre travail est similaire à d'autres expériences de perception, qui préconise une expérience contrôlée, avec des conditions bien équilibrées et des difficultés variables. Dans Hybrid Image Visualizations [IDW+13], avec des collègues de l'équipe AVIZ nous montrons comment tirer parti des différences de perception en fonction de la distance. Nous montrons comment combiner deux visualisations qui sont filtrées à l'aide de filtres des différentes fréquences, de sorte que l'une devienne visible lorsqu'elle est vue de loin et l'autre lorsqu'elle est vue de près. Cette combinaison augmente la quantité d'informations qui peut être rendues sur le mur d'images. Notre principale contribution ici a été l'explication de la théorie derrière l'approche, et les outils pour créer de telles visualisations. Comme pour Smarties, notre méthodologie a consisté à fournir de nombreux exemples divers, à explorer les limites de l'approche, plutôt que de mener une étude sur les utilisateurs.

Collaboration. Dans le cadre de la thèse de A. Prouzeau (co-encadré avec O. Chapuis), nous examinons s'il existe des différences quantitatives entre la collaboration en utilisant un mur d'images et des ordinateurs de bureau coordonnés. Étant donné la nature quantitative de notre question de recherche, notre méthodologie consiste en une expérience contrôlée [PBC17b]. Pour une simple tâche de coordination, nous avons montré que le mur d'images était plus lent, mais qu'il permettait d'obtenir des résultats de qualité plus constante. En tant que chercheurs nous cherchons toujours pour des preuves mesurables de le bénéfice de l'utilisation de ces surfaces collaboratifs.

Nous étudions ensuite l'impact de la technique de sélection sur la coordination, toujours pendant la collaboration autour des murs d'images. Nous regardons en particulier les graphes, une structure très pertinente pour des domaines comme la biologie, le trafic, etc. Nous avons constaté qu'avec une technique de sélection basique, qui a une faible empreinte visuelle, les participants avaient la tendance à diviser une tâche non divisible, ce qui se traduit à une précision base [PBC17a]. Alors qu'avec une technique de sélection basée sur la propagation, qui a une plus grande empreinte visuelle, les participants étaient plus précis. Dans ce travail, nous utilisons une méthodologie mixte, en commençant par une expérience contrôlée pour comparer les deux techniques, dans le cadre d'une tâche spécifique (trouver le chemin le plus court). Comme la propagation est une technique nouvelle dans des contextes collaboratifs, nous voulions aussi voir comment les participants s'en approprieraient dans d'autres scénarios d'analyse. Nous avons donc ensuite mené une étude plus ouverte où ils ont effectué d'autres tâches de topologie sans entraînement.

Enfin, nous considérons la collaboration dans un contexte d'utilisation spécifique, les centres de commande et de contrôle. L'accent de ce travail est mis sur un groupe d'utilisateurs spécifique et leurs besoins, donc nous avons a suivi une méthodologie de conception centrée-utilisateur. Au lieu de commencer avec une solution spécifique en tête, nous avons observé et interviewé les utilisateurs. Cela nous a amené à une question secondaire (concevoir et tester différentes visualisations pour la prévision du trafic [PBC16a]) et un prototype avec des techniques de awareness pour aider les opérateurs à se coordonner et à passer du travail individuel au travail collaboratif [PBC18].

Dans l'ensemble, les travaux présentés dans ce chapitre ont renforcé que les murs d'images peuvent augmenter la bande de communication entre les humains et leurs données. En ce qui concerne la puissance de calcul humaine, ils peuvent rassembler des collègues ayant des compétences diverses et mener une coordination élevée (par rapport à d'autres dispositifs de collaboration tels que des ordinateurs de bureau connectés). Du côté machine, leur haute densité de pixels peut aussi accueillir plus d'informations que les écrans traditionnels. En outre, ils permettent également de visualiser les informations avec des granularités différentes en fonction de la distance de vue, et peuvent même combiner deux visualisations différentes qui sont chacune vues à des distances spécifiques.

4.7 Représentations appropriées

La conception de visualisations interactives n'est pas facile. En tant que leur concepteurs, nous devons nous assurer que les représentations ou systèmes que nous proposons fonctionnent bien avec des données réelles, et peuvent aider avec les tâches des utilisateurs. Et tandis que l'utilité des systèmes conçus autour des besoins réels des utilisateurs est claire pour les utilisateurs eux-mêmes, elle ne nous aide pas nécessairement à comprendre les phénomènes autour (les mécanismes de perception visuelle et de compréhension des informations présentées). La recherche en visualisation aborde à répondre au question "quelles sont les visualisations interactives appropriées", mais par différentes perspectives / pointes de vue. Par exemple, on peut d'abord considérer l'utilisateur final et ses besoins. Ou on peut commencer par les propriétés des données et les tâches qu'on vait effectuer. On peut également être motivé par les algorithmes de requêtage ou d'autres technologies disponibles. Ou on peut rechercher une compréhension plus profonde des propriétés fondamentales et de l'impact des représentations visuelles, afin de fournir des guides de conception.

Ce chapitre présente un panorama mon travail sur la création de représentations visuelles appropriées, qui proviennent de différentes perspectives.

Utilisateur. Je décrit d'abord les travaux qui ont tenté de répondre aux besoins d'utilisateurs spécifiques, en commençant par les analystes de Business Intelligence [EB11, EB12, EAB13], travaux effectués pendant le doctorat de M. Elias que j'ai co-encadré avec M.-A. Aufaure. Le chapitre décrit aussi notre travail avec les généalogistes [BDF+10] (en collaboration avec P. Dragicevic, J.-D. Fekete, J. Bae, B. Watson), et les neuroscientifiques [GTPB19] qui fait partie du travail de A. Gogolou que j'ai co-encadré avec T. Palpanas et T. Tsandilas. Notre travail avec des analystes de Business Intelligence sur la narration (storytelling) [EAB13] a reçu le prix Brian Shackel dans INTERACT 2013. Ici, nous avons suivi une méthodologie de conception centrée-utilisateur: en commençant par comprendre le contexte et les défis auxquels les experts du domaine sont confrontés (avec des interviews et des workshops), nous avons ensuit poursuit des sessions de conception pour des solutions possibles, et finalement on a eu des commentaires et des retours de nos experts sur les conceptions finales. Les besoins de ces utilisateurs ont motivé des solutions innovantes, telles que des annotations contextuelles pour les tableaux de bord, qui attachent des annotations aux points de données ou aux requêtes (indépendamment du graphique et de l'agrégation) [EB12].

Question. La deuxième partie du chapitre décrit mes recherches passées qui partent de questions fondamentales et théoriques, y compris notre recherche sur l'existence de biais cognitifs lors de la prise de décisions à l'aide de visualisations. Je décris ces travaux autour de questions fondamentales sur la façon dont les êtres humaines prennent des décisions à l'aide de visualisations, qui font parti du doctorat d'E. Dimara que j'ai co-supervisé avec P. Dragicevic [DBD17a, DBBF19, DBD17b, DBD18], y compris une taxonomie des biais cognitifs [DFP+20] (avec E. Dimara, P. Dragicevic, S. Franconerri, C. Plaisant, G. Bailly). Je résume également les recherches sur la perception pour des représentations visuelles spécifiques, telles que les glyphes [FIB+14, FIBK17] (avec J. Fuchs, P. Isenberg, D. Keim et E. Bertini) et les linecharts [IBDF11] (avec P. Isenberg, P. Dragicevic, et J.-D. Fekete). Notre travail sur le biais cognitif "attraction effet" a reçu une mention honorable dans IEEE VIS 2018 [DBD17a].

Outil. La troisième partie du chapitre résume comment nous avons adapté un outil de visualisation existant, ScatterDice [EDF08], à utiliser dans différents contextes. D'abord on a introduit la variation GraphDice pour visualiser les graphes multidimensionnelles (au lieu des scatterplots) et on a examiné le potentiel de l'outil pour des données et des tâches spécifiques à l'analyse des réseaux sociaux [BCD+10]. De même, EvoGraphDice a commencé avec l'idée de combiner l'outil GraphDice avec l'apprentissage automatique (des algorithmes évolutifs interactifs) pour faciliter l'exploration de données de très haute dimension, offrant des vues intéressantes. Ce travail a également commencé par un outil, mais a mené une grande séquence de travaux sur l'application de l'utilisation d'un outil qui combine le calcul humain et algorithmique à différents contextes et domaines [BTBL13], ainsi que des questions sur l'évaluation de ces types des systèmes [BBL18].

Données/Tâche. La dernière partie examine des exemples de mon travail sur des données et tâches spécifiques. En particulier, avec mes collègues, nous avons étudié des données géo-temporelles et des nouveaux tâches. Dans une première étude, nous avons examiné comment visualiser la corrélation entre deux variables thématiques (par exemple, l'espérance de vie et le taux de fertilité) à travers l'espace et le temps [PPB20]. Dans la seconde, nous avons examiné les caractéristiques uniques des mouvements de propagation (par exemple, un virus) et comparé différentes visualisations dans la façon dont elles les transmettent [PBP20]. Tous les deux ont fourni des guides pour les différentes tâches et visualizations possibles.

4.8 Perspectives

Au-delà des environnements collaboratifs traditionnels (murs, tables) qui font parti de mon travail, nous voyons un nombre croissant de casques de réalité augmentée (RA). Elles pourraient accroître les capacités des dispositifs collaboratifs, en affichant des informations dans l'environnement. Les plateformes de collaboration traditionnelles (murs, tables) et les casques RA, varient considérablement en termes de coût, de taille, de résolution et de modes d'interaction. Au future je vais étudier quelle technologie est la plus appropriée pour des tâches données et les besoins des utilisateurs. Plus particulièrement, je étudierai quand les grandes dispositifs telles que les murs sont nécessaires pour l'analyse collaborative, quand des casques de réalité augmenté sont suffisantes, et si un mélange de ces technologies est intéressant dans des situations collaboratives.

Je voudrais également explorer la relation entre l'interaction et la perception visuelle. Lorsque les utilisateurs manipulent activement les informations, cela peut affecter leur compréhension. Avec mes collègues, nous avons déjà vu par exemple que le choix des techniques d'interaction peut affecter la coordination et la qualité des tâches simples sur des graphes. Des travaux récents autour de la notion d'interaction dans la visualisation, soulignent que même si les avantages de l'interaction sont généralement reconnus, elle est rarement au centre de nos études de recherche en visualisation. Plus généralement, l'espace de la conception d'interfaces pour les outils d'analyse visuelle reste assez inexploré. Récemment de nouveaux paradigmes d'interaction sont étudiés dans le domain de la visualisation. Parmi eux est la manipulation directe pour modifier les encodages visuels dans une visualisation (une approche adoptée de manière non systématique par plusieurs systèmes). Notre travail sur les SketchSliders est complémentaire, au lieu de personnaliser les encodages visuels, nous considérons la personnalisation et création des outils dont les analystes auront besoin dans leur exploration. Cette approche mérite d'être étudiée plus. Nos participants ont souvent eu le sentiment que les décisions qu'ils ont prises concernant la création de leurs outils les ont aidés à structurer leur analyse. En plus l'ajustement et la copie de ces outils leur ont permis de garder une trace des explorations alternatives. Ils ont commenté que leurs choix d'outils et les annotations capturent des aspects subtils de l'exploration qui est important de saisir et de stocker. L'étude des implications de la création d'outils à la volée pour l'analyse visuelle est une direction sur laquelle j'aimerais continuer à travailler.

Lorsqu'il s'agit de questions plus fondamentales, il reste encore beaucoup de questions sur la façon dont nous prenons des décisions à l'aide de systèmes de visualisation. Notre article sur la taxonomie des biais cognitifs a identifié de nombreux problèmes ouverts à étudier, comme : la façon dont notre mémoire des visualisations peut être biaisée, comment l'utilisation de l'automatisation (ex l'apprentissage automatique) en conjonction avec la visualisation peut affecter nos décisions, si on peut utiliser les visualisations pour atténuer ces biais (comme nous l'avons fait avec l'attraction effect), etc. Des travaux récents dans la communauté de la visualisation ont commencé à considérer les biais possibles lors de la compréhension et de l'analyse des données. Plus important encore, cela soulève des questions fondamentales sur la manière dont nous menons nos propres recherches. Par exemple, on sait qu'il existe un biais de confirmation dans la recherche. Mais est-il possible que d'autres biais affectent nos résultats et nos conclusions? Par exemple, lorsque nous présentons aux utilisateurs trois systèmes ou techniques à classer, introduisons-nous des effets d'attraction ou des biais de compromis, amenant les participants à favoriser certains systèmes et techniques? J'aimerais poursuivre les travaux sur ce sujet et, au-delà des biais spécifiques, tenter de mener une méta-analyse des travaux dans notre domaine qui ont pu être affectés par de tels biais.

Dans me perspectives est d'explorer deux questions de haut niveau sur la façon dont nous menons des recherches en visualisation interactive. Le premier est, comme mentionné ci-dessus, l'étude des biais possibles qui nous affectent en tant que chercheurs en interaction et visualisation. Le deuxième est d'essayer d'analyser l'inspiration derrière notre travail en tant que communauté et d'explorer s'il existe des moyens de le capturer et de le partager avec les autres. Bien que plusieurs articles de recherche reconnaissent clairement leur inspiration, cela n'est pas toujours systématiquement fait lors de la rédaction d'un article de recherche, car l'espace peut être limité, ou parce qu'en tant que chercheurs, nous devons présenter une histoire cohérente de notre travail. De plus, nos articles capturent rarement l'itération de nos idées et techniques. Les domaines intéressés par la conception de solutions aux problèmes pernicieux (wicked problems) [KR70] ont développé des méthodologies pour saisir leur processus de rationalisation de la conception [Lee97] (c.à.d. leurs raisons, justifications, alternatives envisagées et compromis évalués). Dans une moindre mesure, IHM a une tradition de suivre le processus de conception itérative et les articles du domaine parlent souvent des itérations précédentes et des résultats pilotes. Par contre, ces informations sont perdues dans une grande partie des articles en visualisation. Mais ces détailles pourraient être importantes pour d'autres chercheurs dans le domaine (renseigner ou réutiliser). La crise de la réplication de la recherche a suscité des efforts dans notre communauté de recherche (entre autres domaines) pour adopter des pratiques scientifiques ouvertes qui peuvent faciliter la réplication et la reproduction des études (pré-enregistrement des études, partager du code, des scripts de données et d'analyse et des résultats). Il est peut-être le moment de déterminer si nous devons également partager plus systématiquement nos inspirations et nos étapes et résultats de conception. Au cours des années suivantes, je prévois étudier comment différents membres de notre communauté de recherche voient et suivent les informations liées à l'inspiration et à l'itération de leur travail, s'ils peuvent voir la valeur d'en partager et communiquer aux autres, ainsi que chercher les méthodologies provenant d'autres domaines que nous pourrions adopter(comme la logique de conception) et comment les adapter en face des pratiques et des contraintes de notre domaine.

Last updated: July 7, 2020

Updated since original submission: Acknowledgements and minor fixes/typos.