



HAL
open science

Post-WIMP Interaction for Information Visualization

Bongshin Lee, Arjun Srinivasan, Petra Isenberg, John Stasko

► **To cite this version:**

Bongshin Lee, Arjun Srinivasan, Petra Isenberg, John Stasko. Post-WIMP Interaction for Information Visualization. Know Publishers, 14 (1), pp.95, 2021, Foundations and Trends® in Human-Computer Interaction, 978-1-68083-827-5. 10.1561/1100000081 . hal-03261048

HAL Id: hal-03261048

<https://inria.hal.science/hal-03261048v1>

Submitted on 15 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Post-WIMP Interaction for Information Visualization

—*Author Version*—

Suggested Citation: Bongshin Lee, Arjun Srinivasan, Petra Isenberg and John Stasko (2021), "Post-WIMP Interaction for Information Visualization", : Vol. 14, No. 1, pp 1–95. DOI: 10.1561/11000000081.

Bongshin Lee

Microsoft Research
bongshin@microsoft.com

Arjun Srinivasan

Georgia Institute of Technology
arjun010@gatech.edu

Petra Isenberg

Université Paris-Saclay, CNRS, Inria, LISN
petra.isenberg@inria.fr

John Stasko

Georgia Institute of Technology
stasko@cc.gatech.edu

This article may be used only for the purpose of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval.

now

the essence of knowledge

Boston — Delft

Contents

1	Introduction	3
2	History	8
2.1	Pen and/or Touch Interaction	9
2.2	Natural Language Interaction	22
2.3	Mid-air Gestures, Proxemics, Gaze, and Immersive Interfaces	30
2.4	Workshops	33
3	Reflections and Viewpoints	36
3.1	Towards Synergistic Multimodal Visualization Interfaces . .	36
3.2	Towards the Ultimate Synergy: Post-WIMP + WIMP . . .	41
3.3	Towards Consistent Post-WIMP Interactions for InfoVis . .	43
3.4	Designing Speech-driven Interactions with Visualizations .	47
3.5	Evaluating InfoVis Systems with Post-WIMP Interactions .	49
3.6	Towards Technology-Inspired Research	54
4	Looking Ahead	56
4.1	Creative Adaptations from Broader HCI Research	56
4.2	Toolkits and Frameworks for Prototyping Post-WIMP Visu- alization Interactions	59
4.3	Context-Aware Visualization Interfaces	61
4.4	Post-WIMP Interfaces for Data Manipulation	63

4.5	Interacting with Data Physicalizations	64
4.6	Reaching Broader Audiences	66
5	Conclusions	69
5.1	Limitations	69
5.2	Concluding Remarks	70
	Acknowledgements	72
	References	73

Post-WIMP Interaction for Information Visualization

Bongshin Lee¹, Arjun Srinivasan², Petra Isenberg³ and John Stasko⁴

¹*Microsoft Research; bongshin@microsoft.com*

²*Georgia Institute of Technology; arjun010@gatech.edu*

³*Université Paris-Saclay, CNRS, Inria, LISN; petra.isenberg@inria.fr*

⁴*Georgia Institute of Technology; stasko@cc.gatech.edu*

ABSTRACT

Well-designed interactions for visualization systems hold great promise both for empowering people to effectively complete their tasks and for providing more natural and engaging user experiences. Leveraging advancements in hardware and software technology, the visualization research community has made considerable progress on providing novel input and interaction experiences. These technologies continue to advance at a fast pace, and thus it seems timely to look back at what has been achieved so far and contemplate what might be possible in the future. In this monograph, we first present a condensed summary of research efforts investigating post-WIMP interaction techniques in visualization systems. We also include research from the broader HCI community and several product releases from industry that we believe to be relevant and have influenced visualization interfaces. Furthermore, we reflect on our own projects that investigated post-WIMP InfoVis interaction and systems. We discuss the main challenges we faced and lessons we learned, and we reflect on how our perspectives and viewpoints on post-WIMP for InfoVis have evolved over the

course of these projects. Finally, we identify several open research directions that will help realize the full potential of post-WIMP interaction for and with InfoVis, expanding the boundaries of InfoVis and reaching a broader audience.

1

Introduction

Interaction, a means for people to express their goals and intentions to systems, plays a critical role in information visualization (InfoVis). The importance of interaction grows as the size and complexity of data increases. Well-designed interactions empower people to effectively complete tasks with visualized data, which simply may not be possible with static visualizations. Furthermore, such interactions can provide a more fluid and engaging experience.

Over the past few decades, both industry and the broad human-computer interaction (HCI) research community have made significant advancements in hardware and software technologies that can be leveraged to support novel interaction techniques. We—InfoVis and HCI researchers—have been increasingly empowered to take a step closer to the post-WIMP user interfaces van Dam envisioned (van Dam, 1997), where the “interface disappears,” enabling people to focus on their tasks rather than worrying about user interfaces consisting of icons, buttons, and menus. Touch-enabled devices are now affordable and prevalent, and they are also becoming increasingly compatible with complementary devices such as digital pens or styluses. Today, it is common to interact with touch input, and pinch & swipe gestures have become

de facto standards. Pen interaction allows people to draw, manipulate, and annotate digital artifacts, such as documents and charts in a more flexible and fluid way. Natural language understanding and generation has made astonishing progress, enabling people to talk to systems and helping them understand the visualized data.

Such advancements are particularly timely, encouraging, and important, presenting novel research opportunities across a wide range of application domains. The advancements also expand both the group of people who could use and benefit from InfoVis and the usage contexts and environments in which InfoVis will be used (Lee *et al.*, 2020). Unlike interaction with InfoVis systems so far, going forward, we cannot and should not assume that one person will be performing data analysis while sitting in front of a desktop computer, with access to a mouse and keyboard Isenberg *et al.*, 2011. Multiple people with different roles and backgrounds may investigate data together on devices such as a wall-sized, multitouch-enabled display. Designers may create novel and engaging visual representations using a tiltable digital canvas equipped with a stylus. Lay people may consume visual data stories on their tablets while lying on a couch or their own data with their smartphones on a bus. Athletes may track how their speed and heart rate change while they are running.

The InfoVis research community has recognized the potential and importance of more natural, flexible, and fluid interactions, going beyond a traditional desktop environment (Lee *et al.*, 2012; Roberts *et al.*, 2014). The community has made slow but steady progress adapting and leveraging advancements in input and interaction technologies in novel ways. We assert that the InfoVis context serves as a great platform to demonstrate the benefits and potential of novel interaction technologies. It is time to take stock of what has been investigated so far and to look into the future.

In this monograph, we discuss recent explorations into post-WIMP interaction for InfoVis. Rather than conducting a comprehensive survey, we instead provide a more personal retrospective of our own projects and experiences. While our reflection is focused on our own research, we refer to other relevant research and commercial systems that have enabled and influenced our research and thinking both directly and

indirectly. We begin this retrospective with a brief overview of research endeavors that have employed and adapted post-WIMP interaction techniques in InfoVis systems. We include and highlight major milestones we believe to be worth mentioning, for example, being the first attempt/application or having a major impact, sparking a thread of research projects. To provide a better context, we also include not only research and activities from the broader HCI community but also product releases from industry (e.g., the Perceptive Pixel, iPhone/iPad, and various Surface devices) that we found relevant and influential to novel interactions within InfoVis. We first collected a purposive sample of 110 publications each author found worth mentioning and categorized them by interaction modality, target device, publication year, and venues. From this sample we selected a representative subset to cover in this monograph. (Throughout the writing process, we incorporated additional papers including the ones suggested by reviewers.) We chose not to cover an in-depth analysis of a wide variety of visualization-related research areas, for example, looking into general HCI research on post-WIMP interaction or including post-WIMP interaction for Scientific Visualization, as we wanted our retrospective to stay close to our personal experiences in the InfoVis area. We acknowledge that what we cover here is by no means exhaustive or comprehensive and that a systematic review on post-WIMP interaction for InfoVis or visualization more broadly would be a good complement to this monograph.

In the scope of this monograph, we generally follow van Dam's notion of post-WIMP (van Dam, 1997), focusing on interaction techniques that do not require a mouse and keyboard. One exception to this is the inclusion of natural language interaction that uses a keyboard because it obviates the need to fully rely on menus and toolbars, and segues into speech-based interaction. Conversely, we exclude interactions purely based on mouse-driven direct manipulation: while they are widely adopted and highly valuable, they inherently rely on a mouse and widgets. Similarly, we do not discuss augmented reality (AR) & virtual reality (VR) applications if their interaction relies solely on a WIMP metaphor (e.g., interacting with menus and buttons in AR/VR through a controller).

We also discuss key insights and knowledge collectively gained and

synthesized from our own projects that employed and adapted post-WIMP interaction techniques in InfoVis systems. We aim to reflect on the challenges we faced, lessons we learned along with the misconceptions we initially had, and how our perspectives and viewpoints of post-WIMP for InfoVis have evolved over the course of different projects. The key reflections and viewpoints that are explored in-depth are:

1. As we explore the benefits and strengths of multimodal interactions, we must strive to maximize the synergies between different modalities while identifying the most suitable operation-to-modality mappings and logical (i.e., semantically meaningful) combinations of modalities.
2. Post-WIMP and WIMP interfaces have their own strengths and weaknesses, and they can complement each other. We thus can and should create synergy between post-WIMP and WIMP to improve and enhance overall user experiences.
3. While it is important to design natural and novel interactions, to increase the resulting systems' practical viability (beyond research settings), we need to consider consistency in interactions across different visualizations, especially within one system. Going beyond a single system, it may not be possible (or even desirable) to design a universal interaction set. However, it may be beneficial to consider an interaction model that can characterize and describe interactions that can be commonly used across different visualizations and systems.
4. Although the two styles of natural language input—typing and speech—share several characteristics, speech has unique challenges from a system design and development aspect. We discuss three key challenges—triggering of speech input, lack of assistive features like autocomplete, and transcription errors, with potential workarounds to guide future work exploring InfoVis systems with speech-based interactions.
5. As acknowledged by several researchers for other contexts (Olsen Jr, 2007; Stasko, 2014; Thudt *et al.*, 2017; Ren *et al.*, 2018), the

traditional comparative studies emphasizing performance metrics (i.e., time and accuracy) are often not appropriate for evaluating post-WIMP visualization systems. We share their viewpoints and suggestions regarding the use of more appropriate evaluation methods and metrics by properly adapting existing methods or devising new ones.

6. Problem-oriented research is a well-established approach with great values (Shneiderman, 2016). However, to spark creative innovations that post-WIMP offers, we need to embrace technology inspired research as well. This will help us take a step closer to our overarching goal, enabling people to stay in the flow of their visual exploration instead of worrying about how to manipulate user interfaces.

Beyond these discussions, to pave the way for future research, we also highlight six open research directions: further leveraging the extensive HCI research on interaction & interfaces, the development of post-WIMP interaction prototyping toolkits for InfoVis, context-aware visualization interfaces, post-WIMP interfaces for data manipulation, interacting with data physicalization, and developing mobile and accessible visualization systems that cater to broader audiences.

Although this monograph describes a personal perspective, we hope it will help readers gain an initial understanding of how post-WIMP interaction techniques have evolved in the context of InfoVis, where we currently stand, and more importantly where we can go in the future. Ultimately, we hope to inspire researchers and practitioners in both the InfoVis and the broader HCI communities. InfoVis-oriented readers will get an overview of the applications and unrealized potential of post-WIMP interactions, while HCI researchers and practitioners may see untapped opportunities where their novel and innovative technologies and interaction techniques might flourish.

2

History

“A post-WIMP interface to me is one containing at least one interaction technique not dependent on classical 2D widgets such as menus and icons.” — van Dam (1997)

This section takes a look at the history of post-WIMP interfaces for InfoVis. It is not meant to be exhaustive, but is instead a curated review of key events, technologies, introductions, and papers that we feel have most contributed to this growing area. To begin, we describe some of the most important contextual contributions from human-computer interaction at large. Next, we focus on interfaces for visualization involving pen and/or touch interaction, followed by those adding natural language interaction to visualization. We move on to discuss alternative interaction methods such as the use of proxemics and AR/VR technologies. Finally, we conclude with a review of key meetings and workshops in the development of this research topic.

From a graphical user interfaces perspective, there are a handful of notable milestones before van Dam argued for the new thinking about fourth-generation UIs, which he called post-WIMP user interfaces in 1997. WIMP (Windows, Icons, Menus, Pointer) interaction was developed at Xerox PARC in 1973 and later popularized with Apple’s

introduction of the Macintosh in 1984 (van Dam, 1997). In the meantime, Bolt (1980) proposed an inspiring multimodal interaction technique (“put-that-there”) that supports “naturalness” in issuing a command by augmenting voice with simultaneous pointing. This was followed by a seminal paper by Shneiderman (1983), which introduced the concept of “direct manipulation.” Cohen *et al.* (1989) showed the potential synergy from the integration of natural language and direct manipulation with two applications built for manufacturing environments. Another paper worth mentioning is an inspiring paper by Weiser (1999), which introduced the ubiquitous computing paradigm. Although this paper does not focus on interaction, it advocates new input technologies, such as live boards and stylus, while describing a new way of thinking about computers that vanish into the background.

Beaudouin-Lafon (2000) introduced a new interaction model called Instrumental Interaction that extends and generalizes the principles of direct manipulation, encompassing a wide range of interaction techniques. Instrumental Interaction describes graphical user interfaces in terms of domain objects and interaction instruments: similar to how we use tools and instruments to interact with physical objects in the real world, interaction instruments mediate the interaction between users and domain objects. We note that Beaudouin-Lafon’s notion of post-WIMP is not fully aligned with van Dam’s even though the Instrumental Interaction model is presented as a way to support the leap from WIMP to post-WIMP interfaces.

2.1 Pen and/or Touch Interaction

Pen and touch input modalities have been extensively investigated. In this section, we describe several milestones and visualization-based research that are pertaining to pen and/or touch interaction. We present them mostly in a chronological order: while not clearly separated, we notice several clusters around device form factors—tabletops, tablets (touch only first followed by pen- & touch-enabled), wall-sized displays, and stationary tiltable displays.

2.1.1 Advent of DiamondTouch, iPhone, & PixelSense in 2000s

Dietz and Leigh (2001) at MERL (Mitsubishi Electric Research Laboratories) developed the DiamondTouch table technology and presented a paper at the ACM Symposium on User Interface Software and Technology (UIST). By proving the capability of detecting simultaneous touches from multiple people, DiamondTouch tables (Figure 2.1) facilitated academic research, being cited more than 1800 times as of May 2021. Leveraging a DiamondTouch table as an input device, the DTLens system (Forlines and Shen, 2005) provided a set of interactions for multiple simultaneous zoom-in-context lenses to enable group exploration of spatial data. Furthermore, DiamondTouch inspired the IEEE International Workshop on Horizontal Interactive Human-Computer Systems Tabletop held for the first time in 2006. This workshop became an annual academic conference called ITS (the ACM International Conference



Figure 2.1: DiamondTouch table detects simultaneous touches from multiple people. This photo is licensed under the Creative Commons Attribution 3.0 License. <https://en.wikipedia.org/wiki/File:MERL-LOBBY.JPG>

on Interactive Tabletops and Surfaces) in 2009, which later became ISS (the ACM International Conference on Interactive Surfaces and Spaces) in 2016. ITS/ISS has published several visualization-focused systems and research projects that have investigated post-WIMP interactions.

The year 2007 was notable for touch technologies: Apple released the first iPhone while Perceptive Pixel released the Multi-Touch Collaboration Wall, and Microsoft announced the first version of PixelSense, a hardware and software platform that allowed to build and deploy multi-touch applications. The research behind the Multi-Touch Collaboration Wall was initially published as a technical note in UIST 2005 (Han, 2005). The technical note presented a simple, inexpensive, and scalable technique for robust multi-touch sensing based on frustrated total internal reflection: the technique was readily available for others to use. Like DiamondTouch, this technique sparked much multitouch research, being cited more than 1600 times as of May 2021. The original Microsoft Surface, a 30-inch touchscreen tabletop, was unveiled in 2007.

A few research projects in InfoVis used the original Microsoft Surface. Cambiera (Isenberg and Fisher, 2009), for example, was a system that allowed two people to analyze document collections together. The pair, sitting on the long side of the Microsoft Surface, used touch to search for, organize, and read/analyze text documents. Cambiera was one of the first data analysis environments for tabletops and researchers showed that participant pairs could effectively solve one of the IEEE VAST challenges using it (Isenberg *et al.*, 2010; Isenberg *et al.*, 2012).

Once direct-touch had become a popular interaction technique to apply to visualization, researchers attempted to understand it more deeply. North *et al.* (2009), for example, characterized gestures participants used for selecting multiple related items from a larger group. Dwyer *et al.* (2009) examined how people manipulate graph-layout using multi-touch while implementing some of the above grouping mechanisms. The authors found that participants made good use of multiple fingers for manipulating graph layouts. With a “scrapbook” application, Hinckley *et al.* (2010) identified nine key design considerations for pen+touch interaction. Although this research was conducted in the context of HCI, the key considerations are still relevant and valuable when designing pen and touch interactions for InfoVis systems.

The Microsoft Surface was replaced by a new model, the Samsung SUR40 for Microsoft Surface in 2011, and renamed as “PixelSense” after the debut of the Surface tablet in 2012. Wigdor and Wixon (2011)—who worked on the development of the Surface product at Microsoft—published a book titled, “Brave NUI World,” introducing natural user interfaces (NUIs). This book discusses the challenges and mistakes in creating touch- and gesture-based user interfaces along with practical design guidance to address and avoid them.

Frisch *et al.* (2009) investigated how people edit node-link diagrams when pen and multi-touch input is available on interactive tabletops. Through the collection and analysis of user-preferred gestures, they provided insights into the suitability of gestures and bimanual interaction on tabletops, along with a rich set of user-defined gestures for node-link diagram manipulations. Volda *et al.* (2009) discussed a set of design considerations for designing practical interaction techniques for InfoVis on tabletops. They also presented two interaction techniques, i-Loupe and iPodLoupe, that reflect different design choices. Schmidt *et al.* (2010) designed a set of multi-touch interaction techniques (TouchPlucking, TouchPinning, TouchStrumming, TouchBundling, and PushLens) to support edge selection and manipulation for node-link diagrams. Some of these techniques are inspired by the existing interactions (EdgePlucking (Wong and Carpendale, 2007) and EdgeLens (Wong *et al.*, 2003)) designed with a mouse input for the desktop. Compared to single-point interfaces, flexible combinations of these techniques (either sequentially and simultaneously) leveraging multi-touch, bi-manual input broadened the interaction possibilities.

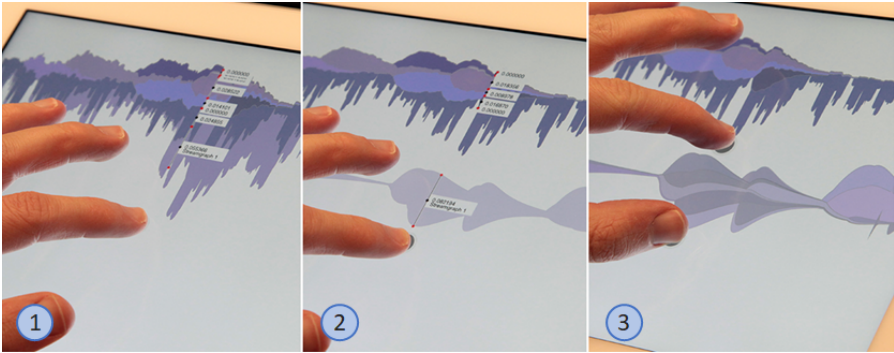
2.1.2 Introduction of iPad in 2010

Beginning with the first iPad on April 3, 2010, Apple released a series of iPad tablets with varying sizes and computing power. This enabled a rich set of tablet-based InfoVis research projects. For example, TouchWave (Baur *et al.*, 2012) (Figure 2.2a) showcased how multi-touch gestures could be used to interact with hierarchical stacked graphs on a tablet, enabling a breadth of operations including retrieving values, extracting layers, sorting, and re-scaling layers, among others.

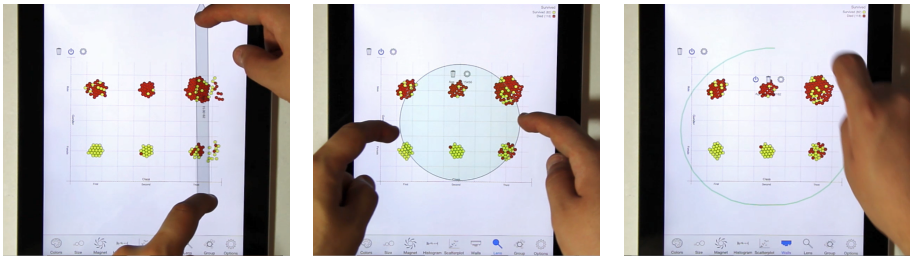
Kinetica (Rzeszotarski and Kittur, 2014) (Figure 2.2b) illustrated how multitouch gestures coupled with physics-based affordances applied to a scatterplot can enable a fluid and enjoyable data exploration experience.

Sadana and Stasko presented a series of tablet-based visualization systems initially focusing on interactions with one visualization (i.e., scatterplots) (Sadana and Stasko, 2014), and later on interactions in the context of multiple coordinated views (Sadana and Stasko, 2016a) (Figure 2.2c) as well as advanced selection techniques such as generalized selection (Sadana and Stasko, 2016b). In doing so, they showed how tablet-based visualization systems can be designed to effectively support analytical capabilities offered in desktop systems, highlighting the corresponding design considerations and challenges. To provide generalized selection capabilities, the system allowed people to use the thumb of their non-dominant hand holding the tablet as a kind of “clutch.” This activity simulated and substituted for keyboard modifiers such as “control” and “shift” to activate different selection modes.

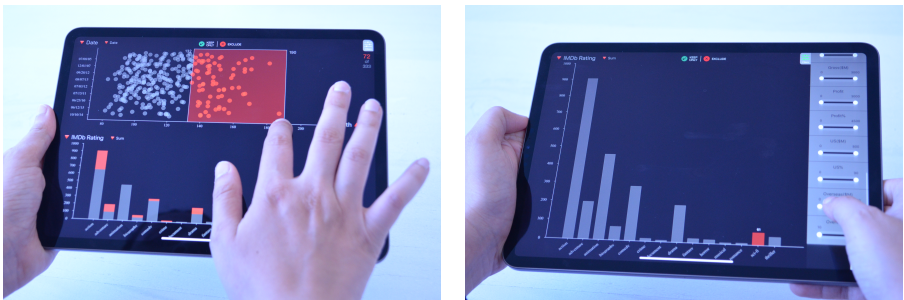
In addition to Apple, other manufacturers also built and released touch-enabled devices, which were also used for touch-based InfoVis research. For example, Drucker *et al.* (2013) designed and compared two versions of tablet-based interfaces for bar charts—WIMP and FLUID—that could be easily operated with touch operations. The WIMP interface employed a control panel which was placed on the side of the screen while the FLUID interface, based on their earlier design work, ensured that all touch interactions occur on the bar chart, striving to minimize the buttons and controls. Their study participants performed better with the FLUID interface. SmartCues (Subramonyam and Adar, 2018) enabled quick access to details on demand by letting people select items and query their details through multitouch interactions. The requested details are provided with labels, highlights, reference lines, and shaded reference bands as overlays. It is implemented as a reusable & extensible web-based library: the library user (i.e., developers, designers) has full control over how to use the default interaction set. We note that SmartCues is not specifically designed for tablet devices: the SmartCues study was conducted using (23-inch) HP Sprout and Microsoft Surface Book. However, we expect its multitouch interactions to work for a tablet form factor as the figures included in the paper imply.



(a) TouchWave (Baur *et al.*, 2012) proposed the touch interaction designed for hierarchical stacked graphs, including (1) showing values with a vertical ruler on demand, (2) extracting a layer by dragging, and (3) showing sublayers by pinching after reaching a scaling threshold.



(b) Kinetica (Rzeszotarski and Kittur, 2014) provided physics-based interactions to allow people to explore multivariate data on tablet devices.



(c) Later version of Tangere (Sadana and Stasko, 2016a) expanded its design to support multiple coordinated views on iPad.

Figure 2.2: Three example projects developed on iPad to examine the touch interaction for InfoVis.

2.1.3 Arrival of the Surface Family in 2010s

Microsoft acquired Perceptive Pixel (along with its team members) in 2012 and released the Surface Hub in 2015, which was equipped with multi-touch and multi-pen capabilities. This enabled a series of research projects leveraging pen and touch on large displays for both data exploration and communication. SketchStory (Lee *et al.*, 2013) demonstrated how freeform sketching with a pen coupled with simple touch interactions can be leveraged to create engaging data-driven presentations. PanoramicData (Zraggen *et al.*, 2014) and SketchInsight (Lee *et al.*, 2015) (Figure 2.3b), conversely, showed how a combination of pen and touch can enable more natural data exploration on an infinite canvas. We note that prior research was conducted to inform the design of SketchInsight before more capable devices became available. SketchVis (Browne *et al.*, 2011) explored the possibility of sketch-based interaction for supporting interactive data exploration by bringing data to digital whiteboards. To investigate the use of both pen and touch for data exploration on whiteboards, Walny *et al.* (2012) conducted a Wizard-of-Oz study, exploring the interaction patterns that people employ during visual data analysis, in particular the interplay between pen and touch interactions.

For their initial investigation, some projects also considered only touch interaction with the Perceptive Pixel or Surface Hub. For instance, Tangraphe (Thompson *et al.*, 2018) supported a series of multitouch gestures to support fundamental operations with node-link diagrams—namely, selection, adjacency-based exploration, layout modification, and navigation. VisWall (Agarwal *et al.*, 2019) coupled direct combination (Holland and Oppenheim, 1999) and derivable visualizations (Dunne *et al.*, 2012; Zraggen *et al.*, 2014) to enable rapid construction of multivariate visualizations using attributes of previously created visualizations (Figure 2.4). By blending visualization recommendations and natural interactions, VisWall enabled users (particularly novices) to focus on the underlying data attributes rather than specifying and reconfiguring visualizations.

Microsoft released the first Surface Pro in 2013 and Surface Book in 2015, 2-in-1 detachable tablets, which take pen input in addition to



(a) SketchStory recognizes a few sketch gestures for chart invocation, and automatically completes charts on the fly using presenter-provided icons and binding them to the underlying data.



(b) SketchInsight allows people to visually explore the data by drawing simple charts and directly manipulating them with pen and touch interaction.

Figure 2.3: SketchStory (Lee *et al.*, 2013) and SketchInsight (Lee *et al.*, 2015) investigated more engaging and fluid ways to present and explore data leveraging pen and touch interaction.

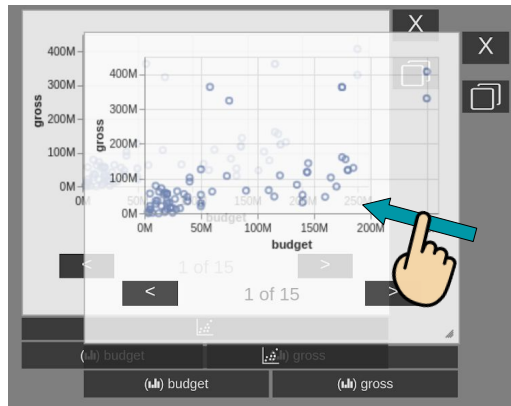


Figure 2.4: Example illustrating visualization construction through direct combination in VisWall (Agarwal *et al.*, 2019). In this case, dragging tiles for two quantitative attributes close to each other “merges” them into a scatterplot.

multitouch. In addition, Apple released iPad Pro and Apple Pencil in 2015: in addition to iPad Pro, the Pencil worked with later generations of the iPad family. These enabled several projects using tablet devices in the context of data communication/presentation as well as data exploration. WordlePlus (Jo *et al.*, 2015) illustrated how pen and touch together can give people higher control over wordle creation by allowing them to manipulate word size, placement, and groupings while incorporating additional visuals and animations. Pen complemented touch in two ways: (1) it allowed people to add and delete a word through writing and striking out as they normally do with a physical pen and (2) it addressed the fat finger problem by letting people select a small word by drawing a circle around it (once selected, the word can be moved with touch). More recently, DataSelfie (Kim *et al.*, 2019b) enabled people to compose custom graphical elements that best match their personal data with freeform sketching. It facilitated the construction of personalized visual vocabularies, i.e., mappings between the data and hand-drawn visuals (or predefined icons and emojis).

TouchPivot (Jo *et al.*, 2017) (Figure 2.5) combines pen and touch interactions with WIMP-style interface elements and visualization recommendation to help novices conduct data exploration on tablets. Several of TouchPivot’s interactions are designed to be performed on



Figure 2.5: TouchPivot (Jo *et al.*, 2017) assists visual data exploration on tablet devices by leveraging pen and touch interactions. Keeping novices in mind, TouchPivot combines familiar WIMP components and novel post-WIMP interactions.

the widgets. For example, TouchPivot employs the fan menu (at the bottom left corner) to enable rapid exploration. People can access data columns by dragging their thumb along the arc of the fan, and pivot the data by the focused column to preview the distribution of values in the panel at the bottom right. Once they confirm the pivot operation by lifting their thumb from the fan menu, the system updates the data table and moves the chart to the main chart view at the top right. Similarly, while exploring thumb + pen interaction on tablets, Pfeuffer *et al.* (2017a) applied their thumb + pen techniques for manipulating and analyzing data in spreadsheets. The pen interaction performed by people’s dominant hand is augmented by thumb interaction with marking menus using the non-dominant hand, while holding the device.

DimpVis (Kondo and Collins, 2014) (Figure 2.6) is a touch-based direct manipulation technique for time navigation, supporting the in-



Figure 2.6: With DimpVis (Kondo and Collins, 2014), users can select a data item and follow visual hint paths that indicate how the selected item has changed over time using touch. (The visual hint path is shown as a gray line while the pink and blue dots indicate a finger point and currently selected item, respectively.)

tegrated interaction concept (Baur *et al.*, 2012). It provides visual hint paths to guide a navigation and indicate how a selected data item has changed over time in time-varying bar charts, scatterplots, heatmaps, and pie charts. *SketchSliders* (Tsandilas *et al.*, 2015) is a sketch-based interface that allows people to sketch range sliders in arbitrary shapes on mobile devices (tablets and smartphones) to explore multi-dimensional datasets on a connected wall-sized display. It is augmented by interaction aides such as slider cursors, markers, and embedded distribution visualizations. To infer the type and function of pen strokes, *SketchSliders* incorporates techniques from previous research such as stroke delimiters (Hinckley *et al.*, 2005), crossing-based selection (Aritz and Guimbretière, 2004), and simple gesture recognition (Wobbrock *et al.*, 2007). *iVoLVER* (Méndez *et al.*, 2016) is an interactive visualization authoring tool that does not require coding. It facilitates acquisition of various types of data (e.g., shapes, text, dates, colors) from multiple types of sources (e.g., bitmap charts, webpages, SVGs), and enables data transformations through a set of widgets.

While iVoLVER supports pen and touch interaction, the designers of iVoLVER decided not to take advantage of different input modalities so that the system can be used across a broad range of devices.

In 2016, Microsoft released another device that supports pen and touch, the Surface Studio, an all-in-one PC equipped with a 28-inch display. Several projects started to leverage the Studio's large tiltable display, again both for data presentation and for data exploration. DataInk (Xia *et al.*, 2018) presented pen and touch interfaces to create compelling infographics. DataToon (Kim *et al.*, 2019a) (Figure 2.7) is an authoring environment for data comics about dynamic network data. More recently, DataQuilt (Zhang *et al.*, 2020) (Figure 2.8) continued the efforts of leveraging sketching to enable people to create beautiful, custom-made visuals, while supporting data-driven visual authoring. DataQuilt allows authors to extract visual elements and stylistic features from existing images (e.g., paintings, photographs, art) and repurpose them to create pictorial visualizations. These systems illustrate how the added expressivity offered by pen and touch interactions can augment users' creativity, facilitating the use of visualization for communication purposes.

Besides data exploration and data-driven communication, recent examples have also shown how pen and touch interactions can aid sense-making and improve analytic provenance. For example, VoyagerInk (Kim *et al.*, 2019c) illustrates how pen and touch can support better note-taking during visual data exploration and help users better organize and track their findings from the data. ActiveInk (Romat *et al.*, 2019a) is another example of pen- and touch-based visualization tool that supports seamless switching between data exploration and externalization to facilitate sensemaking.

As discussed above, the recent InfoVis research pertaining to pen and/or touch has been initially carried out with tablet devices and then wall-sized displays, followed by large monitors. Very rare examples of research with a mobile phone form factor are the study of pan and zoom interactions (Schwab *et al.*, 2019a) and the Orchard (Eichmann *et al.*, 2020) mobile app. In crowd-sourced experiments, Schwab *et al.* compared a set of eight different pan and zoom interaction techniques with varying levels of task difficulties. Even though they conducted



Figure 2.7: DataToon (Kim *et al.*, 2019a) is a storyboarding tool that facilitates the creation of data comics focusing on network data by blending analysis and presentation in a unified environment supported by pen and touch interactions.



Figure 2.8: DataQuilt (Zhang *et al.*, 2020) helps authors create expressive custom visuals with sketching by allowing them to repurpose visual elements and stylistic features from existing images.

the study for both mobile and desktop devices, the goal of the study was not to compare the techniques on mobile versus desktop. As such, they did not compare methods across mobile and desktop in their analysis. Orchard is a touch-based interface for exploring multivariate heterogeneous networks on mobile phones. To enable casual browsing of data-rich network datasets, Orchard employs a simple swipe gesture with a list-based representation most people are familiar with. It is important to note that Orchard's simple representation and interaction design was enabled by a novel query model, called Pivot Trails.

2.2 Natural Language Interaction

Natural language interfaces (NLIs) have a rich history in HCI research dating back to the 1960s, with classic systems like ELIZA (Weizenbaum, 1966) and SHRDLU (Winograd, 1972). Although NLIs were developed in an InfoVis context decades after these systems, natural language has gained notable interest in the context of post-WIMP interaction with InfoVis systems, especially over the past five years or so. Below we discuss the key trends and themes of NLI research for InfoVis since their introduction in the early 2000s.

2.2.1 Early-stage research studies and prototypes until the early 2010s

Cox *et al.* (2001) presented an early prototype of an NLI for data visualization. With their system, to produce a visualization, users had to specify a well-structured query or a set of partial queries while engaging in a dialogue with the system. This project was followed by a gap of almost a decade before the appearance of a second NLI for data visualization. In 2010, the Articulate system (Sun *et al.*, 2010) introduced a model to create visualizations from NL queries by deriving mappings between tasks and data attributes in user queries. (Articulate2 (Kumar *et al.*, 2016) extended Articulate (Sun *et al.*, 2010) with the support for conversational queries.) Metoyer *et al.* (2012) conducted a study to observe how people verbally describe visualizations. Their study highlighted several system design considerations including the support

for semantic language to describe charts and providing a feedback mechanism to help users resolve ambiguities in input queries. Although these systems and studies were a promising start, the idea of using natural language was still restricted to simple research prototypes. Perhaps a reason for this limited interest was because natural language was not yet a prevalent mode of HCI. This is no longer the case, however. Particularly since Apple’s introduction of Siri (Apple Inc., 2021) as part of iPhones in 2011 and Amazon’s release of Alexa (Amazon.com Inc., 2021) in 2014, NLI has become not just a reality but an integral part of our daily lives.

2.2.2 Resurgent of commercial interest and supporting analytic conversations since mid the 2010s

The growing popularity of voice interfaces and the advent of natural language understanding and generation technology rekindled interest in NLI for visualization. Commercial tools including IBM Watson Analytics (IBM, 2018), Microsoft Power BI (Microsoft, 2021), and Tableau (Tableau Software, 2021) adopted natural language interaction. IBM announced Watson Analytics—a natural language-based cognitive service—in 2014, Microsoft Power BI made the ‘Ask A Question’ feature (later transformed to the Q&A visual) available in the Power BI Desktop in 2018, and Tableau launched its ‘Ask Data’ feature in 2019. These applications allowed people to query data or request visualizations through natural language, providing features like query auto-completion and question suggestions to help users formulate valid queries (Figure 2.9). The proliferating user base of natural language-based commercial visualization tools also led to a renewed series of research efforts in the past five years.

DataTone (Gao *et al.*, 2015) allowed specifying visualizations through natural language. In particular, DataTone focused on managing ambiguities in the input query and introduced the idea of using “ambiguity widgets” to let users resolve ambiguities through a GUI. Eviza (Setlur *et al.*, 2016) explored the idea of using natural language to interact with an active visualization. Eviza enabled a richer analytic dialog by allowing people to issue a sequence of utterances (as opposed to one-off

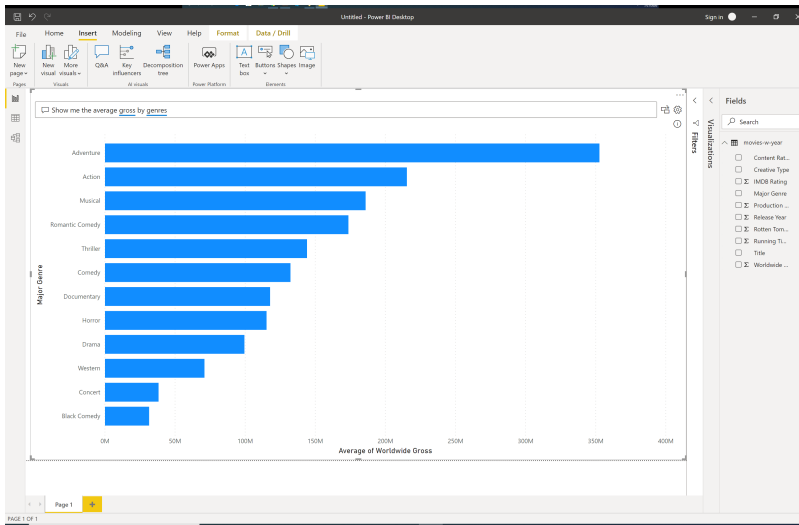


Figure 2.9: The Q&A visual in Microsoft’s Power BI allows users to specify charts using natural language commands. In this case, the system underlines (in blue) terms in the input query that were detected as data attributes.

commands) to iteratively refine an active visualization and perform visual analysis. A comparative study between Eviza and Tableau Desktop also showed that people were almost twice as fast with the natural language interface.

Eviza’s design and evaluation spurred a series of research efforts focusing on enabling a richer “analytic dialog” between users and visualization tools. For instance, extending Eviza’s capabilities and incorporating additional pragmatics concepts, Evizeon (Hoque *et al.*, 2018) enabled both specifying and interacting with visualizations through standalone and follow-up utterances (Figure 2.10). Following these initial research prototypes with a Wizard-of-Oz study, Tory and Setlur (2019) further explored how people naturally engage in an analytic dialogue with visualization tools, highlighting the challenges and importance of understanding user intent and query context. Building upon this work, Setlur *et al.* (2019) devised a set of inference techniques to tackle the challenge of query underspecification (e.g., queries with missing chart types, partial references to data attributes). Collectively,

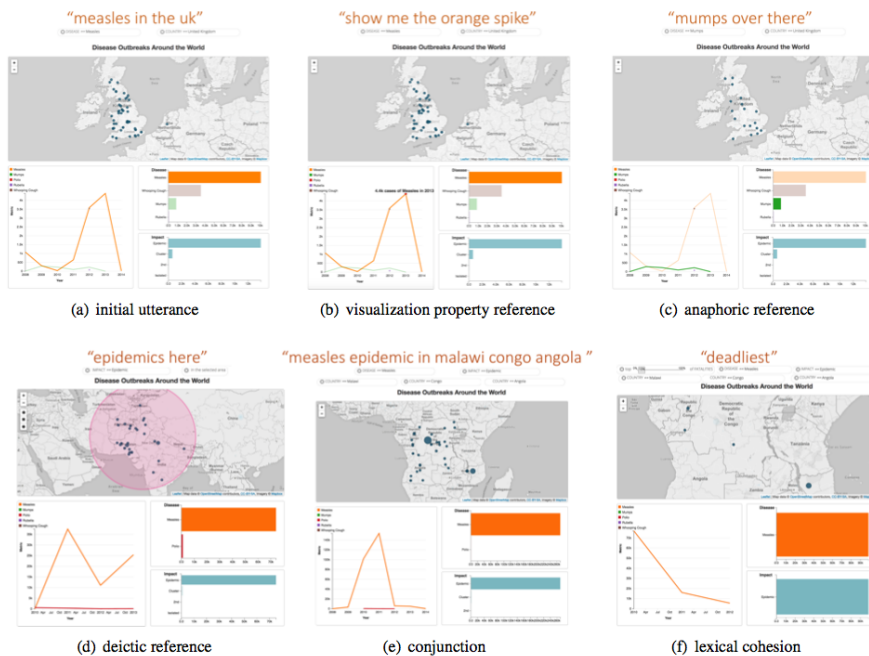


Figure 2.10: An natural language interaction sequence illustrating how Evizeon (Hoque *et al.*, 2018) supports a rich dialog between users and visualization systems during visual data analysis.

these efforts (Setlur *et al.*, 2016; Hoque *et al.*, 2018; Setlur *et al.*, 2019) also informed the design of Tableau’s Ask Data (Tableau Software, 2021) service launched in 2019. Although their core goal is to support computational data science tasks, systems like Ava (John *et al.*, 2017) and Iris (Fast *et al.*, 2018) also allow specifying visualizations in the context of a dialog about the underlying machine learning model or statistical tests.

2.2.3 Going beyond data exploration

Besides creating and interacting with visualizations during data exploration, recent systems have also begun to explore natural language interaction to support other tasks. For example, FlowSense (Yu and Silva, 2019) augmented a dataflow-based visualization system with natu-

ral language input. In doing so, FlowSense allowed people to specify and connect components in the underlying dataflow system without having to learn the intricacies of operating the interface. Text-to-Viz (Cui *et al.*, 2019) offered an interface to convert proportion-related NL statements (e.g., “*More than 20% of smartphone users are social network users.*”) into stylized infographics. Vis-Annotator (Lai *et al.*, 2020) let people enter textual descriptions about a given visualization (e.g., “*The large point above 250 represents China.*”) and annotated the visualization based on those descriptions. Kim *et al.* (2020) provided a dataset and model for answering questions about given visualizations. Specifically, the model takes NL questions about a given visualization (e.g., “*What country has the shortest orange bar?*”) as input and answers them along with an explanation of the steps performed to get to the response (e.g., “*China. I looked up ‘Country’ of the shortest orange bar.*”).

2.2.4 Voice-based visualization interfaces since late 2010s

Unlike the aforementioned systems that support natural language input with typing, speech-based systems require different interface and interaction design considerations due to the added complexity with potential speech-to-text recognition errors and the lack of assistive features such as autocomplete. (We discuss the main challenges for designing speech-based interactions further in Section 3.4.) Furthermore, typing is not an efficient input technique in post-WIMP settings such as touchscreens (especially on mobile devices with small screen) and AR/VR where speech is a more natural form of input. To this end, researchers have also investigated how speech can augment interaction with visualization tools in post-WIMP settings.

Aurisano *et al.* (2020) investigated how people create and manage multiple views during visual analysis on a wall-sized display using the Articulate2 system (Kumar *et al.*, 2016). Possibly the first visualization system investigating the combination of touch and speech input was Orko (Srinivasan and Stasko, 2018) developed in 2017. Orko focused on node-link diagrams and provided a multimodal interface for a suite of typical network visualization operations (e.g., finding nodes, finding paths). A study comparing the multimodal interactions in Orko (Sak-

theeswaran *et al.*, 2020) to their unimodal counterparts (touch-only, speech-only) also found that people prefer multimodal interfaces with this preference stemming from different factors: 1) the freedom of expression, 2) the complementary nature of speech and touch, and 3) integrated interactions afforded by the combination of the two modalities. Similar to Orko, Valletto (Kassel and Rohs, 2018) also explored the use of touch and speech for visual analysis but focused on tabular data and tablet devices.

Incorporating a pen as a third modality in addition to touch and speech, InChorus (Srinivasan *et al.*, 2020a) (Figure 2.11) illustrated how multimodal interaction can support a fluid and consistent interaction experience during visual analysis on tablets. In addition to working individually for the operations that match their innate characteristics, three input modalities can work together when they complement each other. The pen and touch enables direct and precise interaction while speech affords the high freedom of expression. InChorus combines multiple modalities to make it easy to perform a complex operation. For example, with a movie dataset, saying “Budget, running time, rotten tomatoes, and imdb rating” while pointing on the Y-axis with touch adds the four mentioned attributes to the Y-axis, creating a parallel coordinates. Furthermore, for many operations that can be achieved in multiple ways using different modalities, people can choose the way they feel most comfortable with.

Another recent example of speech-based multimodal visualization interfaces is DataBreeze (Srinivasan *et al.*, 2020b) (Figure 2.12). Unlike previous systems that explored how multimodal input can augment interaction experiences with canonical representations (e.g., bar charts, line charts, node-link diagrams), DataBreeze interweaved pen-, touch-, and speech-based multimodal interaction with “flexible unit visualizations.” Through this interweaving, DataBreeze allowed users to create and interact with both systematically-bound views (e.g., scatterplots, unit column charts) and customized views reflecting their mental model of a data space, in turn, enabling a freeform style of visual data exploration.

In addition to the tablets and large displays, multimodal interaction lends itself to visualization interfaces on smartphones that have limited screen space. Data@Hand (Kim *et al.*, 2021) (Figure 2.13) combined

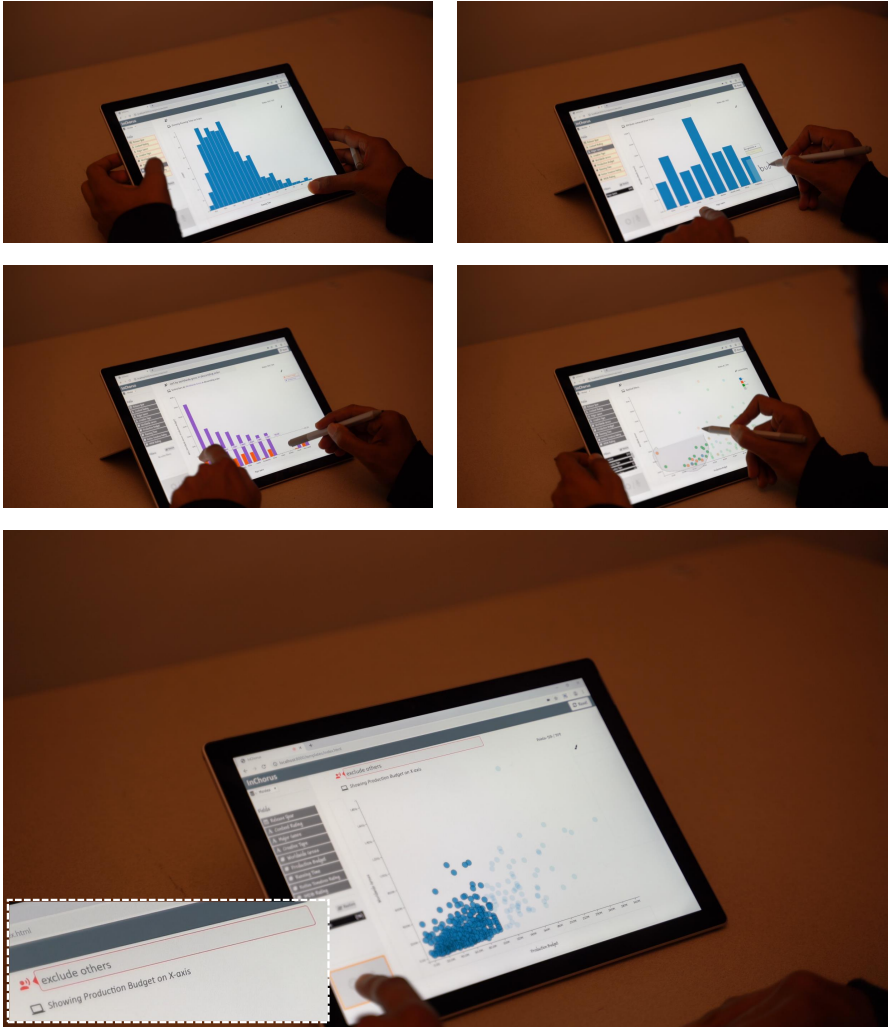


Figure 2.11: InChorus (Srinivasan *et al.*, 2020a) employs multimodal interactions that function consistently across different types of visualizations, supporting core visual data analysis operations on tablet devices.

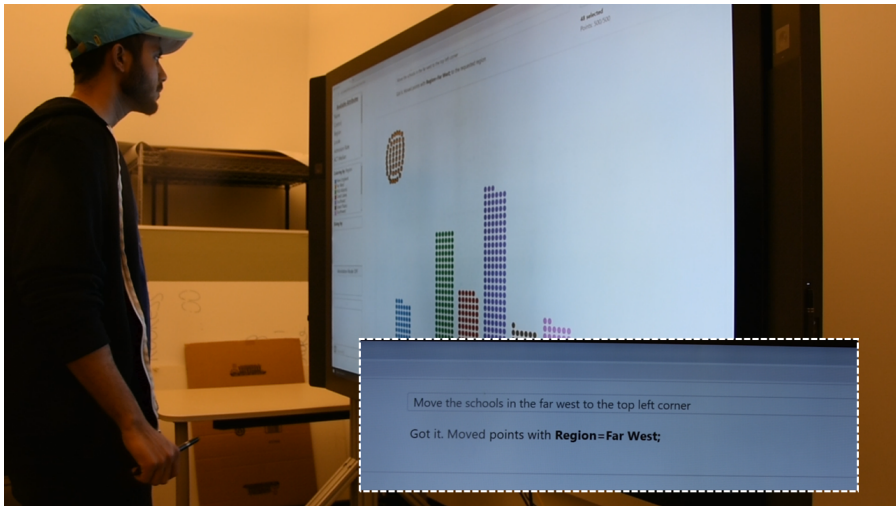


Figure 2.12: DataBreeze (Srinivasan *et al.*, 2020b) running on an 84” Microsoft Surface Hub. The issued voice command is also shown separately on the bottom-right to aid readability.



Figure 2.13: Data@Hand (Kim *et al.*, 2021) enables people to easily navigate and compare their personal health data on smartphones by leveraging two complementary modalities: speech and touch.

two complementary modalities—speech and touch—to enable people to easily navigate and compare their personal health data on smartphones. By carefully considering the smartphone form factor and personal data context together, Data@Hand’s multimodal interaction enables flexible time manipulation, while reducing the interface complexity.

2.3 Mid-air Gestures, Proxemics, Gaze, and Immersive Interfaces

As discussed above, pen/touch and natural language have been the most commonly explored input modalities. Although not many, other research efforts have investigated the use of alternative modalities or have leveraged the technological advances in immersive devices.

2.3.1 Adoption of Proxemics and Mid-air Gestures

Proxemic interaction (Ballendat *et al.*, 2010; Greenberg *et al.*, 2011) utilizes fine-grained knowledge of nearby people, objects, and digital devices, such as their position, identity, movement, and orientation to enable implicit and explicit interactions. Implicit actions are invoked based on the implied user actions (e.g., a person’s entrance to the room or approach to the display) while explicit actions are stated by the user (e.g., direct touch).

Spindler *et al.* (2010) explored mid-air gestures of tangible cardboard displays (called tangible views). These could be held above a tabletop displays showing data visualizations. Projectors on the top visualized additional views on the cardboard displays depending on their orientation and distance relative to the tabletop display as well as specific gestures (shaking, flipping, tilting) of the tangible views. This system is an example for the use of the concept of proxemics with specific props or devices. Others have explore proxemics through the movement of people in front of displays. Jakobsen *et al.* (2013), for example, explored how proxemics can be applied to InfoVis, identifying design opportunities based on movement and distance to wall-sized displays. They reported three user studies conducted to gather initial empirical data about the usefulness of three sets of interaction techniques they implemented: (1) pan & zoom navigation by physical movement; (2) adapting visual rep-

representations based on distance and movement; and (3) dynamic queries based on movement. Badam *et al.* (2016) demonstrated how gestures and proxemic distance from a large wall-sized display can be used for both individual and collaborative visual data exploration. With their system, users could select and control visualizations through explicit gestures while movements in the physical space are implicitly inferred to resize and move the selected visualizations.

Another input modality that has been considered for visualization interaction is eye-gaze. Streit *et al.* (2009) illustrated how simple gaze-tracking techniques can be used to support focus+context based interactions and zooming parts of a visualization on-the-fly based on user focus. Extending this idea, Okoe *et al.* (2014) demonstrated gaze-based interactions for visual network data exploration with node-link diagrams. Specifically, they showed how eye-gaze can help users focus on nodes and links of interest by dynamically filtering other parts of the network.

2.3.2 Arrival of Compelling AR/VR Devices in 2010s

The origin of AR/VR goes back to 1830s when the stereoscope was invented to create a 3D image for the viewer. Since then, AR/VR technology has evolved but with the level of interests ebbing and flowing. It has not gained traction from the InfoVis community until recently, when several compelling AR/VR devices were introduced. After being started as a crowd-funded Kickstarter campaign in 2013, Meta began to ship its Meta 1 Developer Kit in 2014 and announced the launch of the Meta 2 in 2016. In the meantime, HTC unveiled a VR headset called Vive during its Mobile World Congress keynote in 2015. In 2016, Oculus VR (a division of Facebook Inc.) released Oculus Rift, which is a line of VR headsets, while Microsoft shipped the pre-production version of HoloLens, mixed reality smartglasses, in the United States and Canada. The availability of these devices empowered some researchers to explore more post-WIMP interactions in the context of immersive visualization systems, beyond representing data and simple WIMP-style interactions with virtual menus in 3D.

While the interaction was facilitated via pointer-based tracked con-

trollers, the ImAxes system (Cordeil *et al.*, 2017) simulated the action of physically grabbing and moving lines that represent axes for data attributes in free space. (The authors demonstrated its use with the HTC Vive but it can be used with most common VR headset and tracked controllers.) By aligning the axes in different orientations, users can fluidly create different visualizations (e.g., arranging lines in an ‘L’ shape creates a scatterplot whereas aligning lines in-parallel creates a parallel coordinates plot) and perform basic filtering by adjusting slider-like widgets directly embedded on the axes. Using Microsoft HoloLens, FiberClay (Hurter *et al.*, 2018) introduced a set of bi-manual interactions to support creating and adjusting filter brushes while exploring trajectory data in VR. In doing so, FiberClay illustrated how enabling more fluid interactions for analytic operations like filtering (as opposed to more conventional menu-based filtering) can support better insight discovery in immersive settings. Using the HTC Vive Pro Head Mounted Display and HTC Vive Controllers, Drogemuller *et al.* (2018) compared four VR graph navigation techniques—Worlds-in-Miniature, Two-Handed Flying, One-Handed Flying, and Teleportation—focusing on node-link diagrams. Through a user study, they found that steering (i.e., one- and two-handed flying) was a more effective form of interaction than teleportation and worlds-in-miniature.

More recently, Embodied Axes (Cordeil *et al.*, 2020) (Figure 2.14) enabled direct selections for 3D visualizations in AR by incorporating a tangible object. As a visualization environment, it consists of three tangible and actuated sliders that form orthogonal axes, and an AR head-mounted display that aligns a visualization within the 3D data space formed by the three sliders. A controlled user study comparing Embodied Axes with conventional tracked controllers showed that Embodied Axes offers more precise selection with no reported fatigue.

In this work, the authors designed and developed four versions of prototypes, using different combinations of devices. The first prototype, they mounted a Leap Motion controller to the HTC Vive headset and used the video pass-through from the Leap Motion cameras. The second prototype of the Embodied Axes was integrated with the Microsoft HoloLens, which offered a less constrained environment for the user. Compared to the HoloLens, the Meta 2 see-through immersive AR

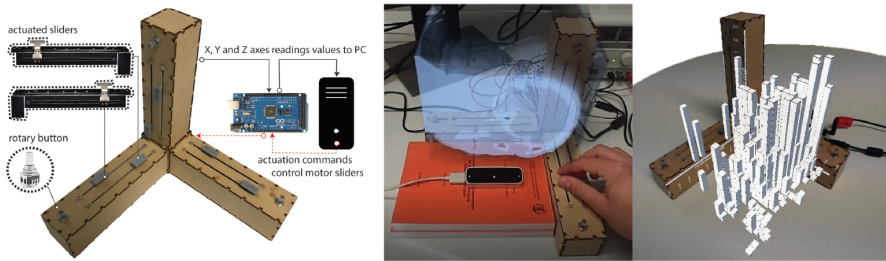


Figure 2.14: Embodied Axes consists of (left) a tangible and actuated controller that supports direct selections for 3D imagery and data visualizations in augmented reality (AR), and (right) an AR head-mounted display that shows the visualizations in its 3D data space.

headset used in the third prototype provided a bigger field of view and higher resolution. Finally, the fourth prototype used the stereo ZedMini mounted on the HTC Vive with an RGB stereo pass through AR. This provided much more stable head tracking than the third one. This demonstrates the benefits of having a range of AR/VR devices, each of which has different strengths and weaknesses.

Reipschläger *et al.* (2021) used a HoloLens to implement an AR+wall-sized display interaction concept that extends visualizations into the space in front of the wall-sized display. The system adjusts views of wall-sized display visualizations based on people’s position in front of the display or based on interactions with the visualizations. On the head-mounted display viewers see, for example, a curved view of the wall-sized visualization with curvature adjusted for their viewing angle.

2.4 Workshops

Besides research prototypes and manuscripts, the InfoVis community has also investigated post-WIMP interaction through focused workshops held as individual events or as part of academic conferences. By providing a chance for participants with shared interests to meet in a focused and interactive setting, workshops provide an opportunity to establish a research agenda, build a community, and move a field forward. They are a great place to find a group of people, disseminate work in progress, and promote an emerging topic. Through workshops, people generate

ideas from an active discussion about the topic and suggest promising directions for future work. While not many, there have been several workshops that treated post-WIMP interactions as a core aspect.

As mentioned above, the IEEE International Workshop on Horizontal Interactive Human-Computer Systems Tabletop¹ started in 2006 and evolved into the ACM ITS (Interactive Tabletops and Surfaces) Conference in 2009, which was later renamed to ISS (Interactive Surfaces and Spaces)² in 2016. In 2011, DEXIS (Workshop on Visual Data Exploration on Interactive Surfaces) (Isenberg *et al.*, 2013a), one of the first workshops (if not the first) with the focus on post-WIMP interaction, was held at ITS. Researchers and practitioners from the Visualization and HCI communities gathered to discuss and shape the field of visualization and analysis on interactive surfaces. A follow-up workshop (Isenberg *et al.*, 2015) was held with the same name in 2015 again at ITS, continuing the efforts to grow a community and refine research around the use of interactive surfaces for visual data exploration, discussing and reflecting on research dedicated to visualization systems for interactive surfaces.

Inspired and encouraged by the rapid advances in immersive technologies (e.g., VR/AR, large displays, tangible surfaces), the first workshop on Immersive Analytics (Bach *et al.*, 2016) was held at ISS in 2016. It was meant to explore the applicability and development of emerging user-interface technologies for creating more engaging experiences and seamless workflows for data analysis and presentations. Follow-up workshops were held one at IEEE VIS 2017 to directly approach the visualization community (Bach *et al.*, 2017) and the other at CHI 2019 to engage the CHI community (Bach *et al.*, 2019) in an effort to improve interaction design for data-centric, immersive systems. Most recently, another follow-up workshop (Ens *et al.*, 2020) was held at CHI 2020 with the focus on productivity. Specific goals of this workshop included understanding how to design and develop immersive analytics systems that effectively support visualization and analytics tasks in VR/AR, and identifying novel interactions to achieve productive use of these

¹<https://ieeexplore.ieee.org/xpl/conhome/10546/proceeding>

²<https://dl.acm.org/conference/its/proceedings>

immersive analytics systems in real world scenarios.

As mobile visualization becomes more common and new mobile device form factors and hardware capabilities emerge, Data Visualization on Mobile Devices workshop (Lee *et al.*, 2018a) was held at CHI 2018. Due to fundamental differences between desktop and mobile environments, including display size, input modalities, target audiences, and usage contexts, existing knowledge and guidelines for data visualization and interaction design do not readily transfer to mobile devices. In 2019, a five-day Dagstuhl seminar titled “Mobile Data Visualization” (Choe *et al.*, 2019) followed the footsteps of the CHI workshop, establishing a community around mobile data visualization. The seminar explored mobile data visualization in depth through speedy and intense research exchanges, interactive demos and tutorials, as well as active breakout group discussions.

A few months after the first mobile data visualization workshop, a workshop dedicated to multimodal interaction for data visualization (Lee *et al.*, 2018b) was held at AVI 2018. This workshop was motivated by the fact that prior research on visualization and interaction techniques has mostly explored a single input modality such as mouse, touch, pen, or more recently, natural language and speech. Therefore, the goal of the workshop was to discuss and identify the unique challenges and opportunities of synergistic multimodal interaction for data visualization, calling for further research efforts.

3

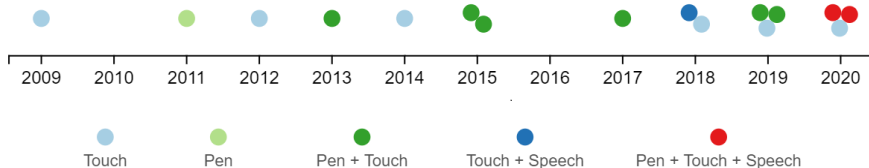
Reflections and Viewpoints

Over the course of almost a decade, we—the authors of this monograph—have been exploring novel ways to interact with data, which enable more natural interactions and engaging experiences for visualization systems, by leveraging different combinations of three input modalities—pen, touch, and speech (Figure 3.1). In this section, we reflect on the insights and knowledge we gained, mainly focusing on our own efforts.¹ We discuss challenges we have faced, misconceptions we initially had, and how our perspectives on post-WIMP for InfoVis have evolved over the course of different projects.

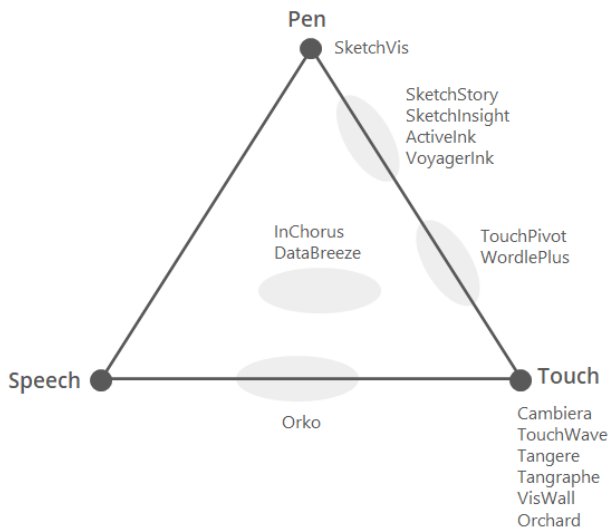
3.1 Towards Synergistic Multimodal Visualization Interfaces

Following the approach of multimodal interface frameworks (e.g., (Gourdol *et al.*, 1992; Martin, 1998)), we categorize our systems that employed multiple modalities along two dimensions: i) their supported use of modalities (Sequential vs. Parallel) and ii) the level of input integration that the system supports (Independent vs. Combined) (Figure 3.2). For

¹Here, usage of the term “we” or “our” refers to prior work that any one of the authors of this monograph participated in and may not reflect the opinion of other coauthors of the individual papers.



(a) Our projects over time colored by the supported input modalities: we have been trying to combine multiple modalities as technologies supporting different modalities have been introduced and stabilized.



(b) Overview of our projects on post-WIMP InfoVis systems grouped by combinations of modalities.

Figure 3.1: We—the authors of this monograph—have been investigating different combinations of three input modalities—pen, touch, and speech.

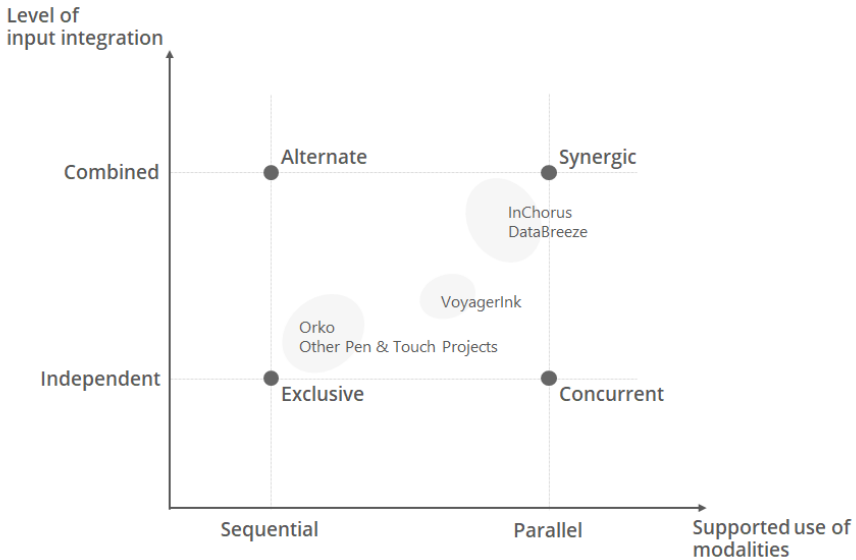


Figure 3.2: Overview of our mutlimodal InfoVis systems based on the design space of multimodal interactions as described by Gourdol et al. (Gourdol *et al.*, 1992).

example, systems like SketchStory (Lee *et al.*, 2013), SketchInsight (Lee *et al.*, 2015), ActiveInk (Romat *et al.*, 2019a), and Orko (Srinivasan and Stasko, 2018) extensively support *exclusive* interactions (modalities used independently) while minimally supporting *alternate* (modalities combined sequentially) or *concurrent* (modalities used in parallel for different operations) interactions. On the other hand, systems like InChorus (Srinivasan *et al.*, 2020a) and DataBreeze (Srinivasan *et al.*, 2020b) support a higher degree of *synergistic* interactions (modalities combined in parallel) while also supporting other styles of multimodal interactions when semantically appropriate.

In retrospect, we initially approached this line of research from a comparative lens, looking to find trade-offs between different input modalities, sometimes by comparing multimodal input to unimodal input. Ensuring a “fair” comparison requires designing interfaces and interactions that are equivalent (i.e., they support the same set of operations through different modalities). However, in our experience, designing multimodal interfaces from a comparative perspective and

stressing equivalence between modalities is likely unwarranted and raises some practical challenges.

For example, consider the study (Saktheeswaran *et al.*, 2020) we conducted with Orko to compare a touch- and speech-based multimodal visualization system to its unimodal counterparts. Since no readily available systems existed to enable this comparison, we had to derive touch-only and speech-only systems to serve in that role. While this served the purpose of the study and helped us verify and understand why people prefer multimodal input, such derived unimodal interfaces may offer suboptimal interactions relative to systems organically designed with one modality in mind. For example, if we were to replace all speech-based interactions in a particular system with touch, the touch-only system would likely have to be bloated with complex menus, essentially mimicking a mouse-based WIMP interface.

Another important point to consider regarding equivalence is that different modalities inherently transmit different content (Oviatt, 1999). One example of this is navigation (e.g., pan & zoom) using touch versus speech. Given its inherent directness, touch is a natural method to support navigation operations. Conversely, specifying (continuous) pan & zoom with speech is at best tedious. Thus, supporting an operation through a modality with the exclusive goal of supporting equivalence (e.g., supporting voice-based navigation in Orko) may make the interaction design unnecessarily complex. Furthermore, the resulting suboptimal user experience with an individual modality may even discourage people from using that style of input altogether.

One important insight we gained through the design and development of multimodal InfoVis systems concerns people's expectation about pen and touch input. For the DataBreeze project that strove to support freeform exploration with unit visualizations, we initially separated the roles of pen and touch following Hinckley *et al.*'s guideline (Hinckley *et al.*, 2010). For example, we let people draw lassos (for selection) and annotations using the pen, and move the points with a finger. However, during our design sessions, we observed that participants often confused the roles of pen and touch, trying to use the two interchangeably for a few operations. Therefore, for simple interactions where semantic differences are not clear between pen and touch (e.g., tap for selection,

drag for move), we decided to support the same operations with both pen and touch. We suspect this expectation is partly because the pen was initially introduced as a mouse replacement, also addressing the fat finger problem, for WIMP-based applications. (Some existing applications like Microsoft OneNote and PowerPoint also let people draw with their finger in the “Draw with Touch” mode.)

We, however, note that sketch affords high freedom of expression, for example, allowing people to draw arbitrary shapes. As we discussed earlier, in the context of expressive visualization authoring, DataInk (Xia *et al.*, 2018) and DataSelfie (Kim *et al.*, 2019b) allow people to draw data glyphs while DataQuilt (Zhang *et al.*, 2020) uses freeform pen strokes to extract elements from images. In addition, QuerySketch (Wattenberg, 2001) and QueryLines (Ryall *et al.*, 2005) allow people to sketch a target graph for querying time-series data (even though they supported it with mouse input), while SketchSliders (Tsandilas *et al.*, 2015) let people sketch the controllers for exploring multidimensional data. Furthermore, in the context of editing or reviewing a document, SpaceInk (Romat *et al.*, 2019b) investigated how to streamline in-context annotations by supporting both space-making and inking actions with pen and touch. Therefore, it is important to continue to understand pen’s unique affordances and strengths, and explore ways to leverage them.

In our experience, synergistic multimodal interfaces can improve user workflows and enable new usage scenarios. A clear benefit of using multiple modalities is minimizing the need for “modes”: because mode-switch involves an additional level of indirection in an interface, it usually hinders users from focusing on their tasks. SketchVis (Browne *et al.*, 2011), our initial exploration of sketch-based charting, employed pen-only interaction. It thus had to limit interaction with charts to creation and erasure of strokes to avoid the need of mode switch. In contrast, our other projects employing pen and touch interaction (e.g., SketchInsight, SketchStory, VoyagerInk (Kim *et al.*, 2019c)) removed the need of explicit mode switch for free: modern pen- and touch-enabled devices can easily and accurately distinguish pen and touch input. In addition, synergistic interactions in DataBreeze (e.g., swipe-and-speak, point-and-speak) enhanced the fluidity in performing freeform data exploration. Similarly, the expressive range of pen + speech interactions

for selection + filtering in InChorus made it possible to eliminate the WIMP-style control panel that would otherwise be needed to support filtering via touch in an equivalent setting.

While we see merit in synergistic interactions, one critical factor to consider when designing such interactions is the complementarity between modalities. For example, the freeform nature of speech makes it a strong complement to more focus-oriented modalities like pen and touch. This complementarity makes it possible to synergistically combine the modalities in various patterns (e.g., pointing on an axis and updating the mapped data attribute through speech) and also correct errors of one modality with another (e.g., using pen to modify an incorrectly transcribed speech command). On the other hand, combining two modalities that are inherently ambiguous (e.g., speech and gaze) may compound their errors and lead to unreliable system performance (Oviatt, 1999).

To summarize, as we explore the benefits and strengths of multi-modal interactions, we should avoid overemphasizing equivalence between modalities since it can result in suboptimal or contrived interactions (e.g., highly customized and complex gestures, heavy reliance on widgets). Instead, for a given set of modalities, we must strive to find the most suitable operation-to-modality mappings and synergistically combine modalities when an operation logically and semantically affords it. We, however, note that further considerations and attentions are necessary for designing visualization interfaces that support people with disabilities (Lee *et al.*, 2020; Lundgard *et al.*, 2019).

3.2 Towards the Ultimate Synergy: Post-WIMP + WIMP

In initial projects such as TouchWave and SketchVis, we strove to eliminate WIMP interactions altogether because we believed post-WIMP was the ideal and more natural way for people to interact with data and visualization. However, as we explored this idea in more depth, we realized that post-WIMP interaction may not be able to fully replace traditional WIMP interaction. We learned that WIMP has its own benefits and can complement post-WIMP. We can create synergy between WIMP and post-WIMP, addressing issues with each and improving the overall user experience.

As we have demonstrated with several projects, post-WIMP interfaces have tremendous potential to support fluid and natural interactions. However, these interfaces are inherently invisible, and they suffer from usability challenges like lack of discoverability, affordances, and feedback. For example, users may not know which speech commands are possible or which gestures can be performed, and it is not easy to provide meaningful feedback when a system recognizes or misinterprets a user intention. WIMP interfaces, on the other hand, provide a higher-degree of affordance and feedback, and are more familiar to end-users. More specifically, menus often suggest the set of possible operations that a person can perform, while the widgets that enable dynamic queries (Shneiderman, 1994), such as checkboxes, double-range sliders, and interactive legends (Riche *et al.*, 2010) also provide feedback on the current state of the visualization (e.g., checkboxes and interactive legends depict the active data and graphical encodings).

Therefore, completely discarding widgets (especially before people become familiar with the novel system) may result in user frustration. An effective post-WIMP interface might, in practice, incorporate such elements of traditional WIMP interfaces. For instance, from the two pilot studies conducted for the TouchPivot project (Jo *et al.*, 2017), we learned that the study participants (who were visualization novices) preferred traditional WIMP-style widgets for specific tasks (e.g., novices preferred checkboxes over a strikethrough gesture for filtering data categories). Correspondingly, in the final implementation, we blended both post-WIMP gestures and WIMP components to support fluid interactions while still catering to novice users and helping them bridge the gap between WIMP and post-WIMP systems (Figure 2.5).

Besides the learnability and familiarity challenges mentioned above, post-WIMP interactions often rely on system inference (e.g., speech-to-text transcription, gesture recognition), which is not always reliable. In such scenarios where post-WIMP interactions fail due to system errors or limitations, it can be beneficial to leverage the support of WIMP elements as a backup mechanism.

An overarching goal for a series of our projects—SketchVis, SketchInsight, and SketchStory—was to create a radically different way for people to explore and present their data using post-WIMP interactions lever-

aging pen and touch. To this end, we envisioned a data-enabled digital canvas that allows people to create and manipulate data charts through simple interactive sketches, with an aim to extend the advantages of a physical whiteboard in the real world with the digital features of an interactive display.

To understand the ideal and most natural ways for people to create and manipulate charts, we chose WYDIWYG (What You Draw Is What You Get) as our first design goal when designing the SketchInsight system (with a person recognizer) for a Wizard-of-Oz study. The system allowed study participants to draw the charts they wanted to see, enabling a compelling chart creation experience (e.g., after drawing axes, people could draw a bar shape and then write a data column name in-place to map the specified data field to the bars). However, supporting this concept in practice was prone to recognition errors, which led to user frustration. Hence, we eventually adapted to a more realistic goal, transforming WYDIWYG into a more fluid interaction experience. People could create and manipulate charts with high probability of recognition. However, when recognition accuracy was difficult to ensure, such as with chart type changes they could resort to alternative interaction options, such as a radial popup menu.

Another example is the use of ambiguity widgets (Gao *et al.*, 2015) to overcome issues with natural language input. In both Orko and DataBreeze, if there are words in an input query that the system maps to multiple data fields or values (e.g., in the context of a movies dataset, the word “rating” can map to one of the three attributes: ‘IMDb Rating,’ ‘Rotten Tomatoes Rating,’ and ‘Content Rating’), the system presents a dropdown menu with the possible matches so that users can select their desired field (Figure 3.3).

3.3 Towards Consistent Post-WIMP Interactions for InfoVis

It is possible to design a large number of different gestures for pen and/or touch interactions given their high freedom of expression. In theory, we can assign one distinctive gesture to each action. However, in practice, a challenge with this approach is to maintain consistency in interactions across different visualizations (even within one system).

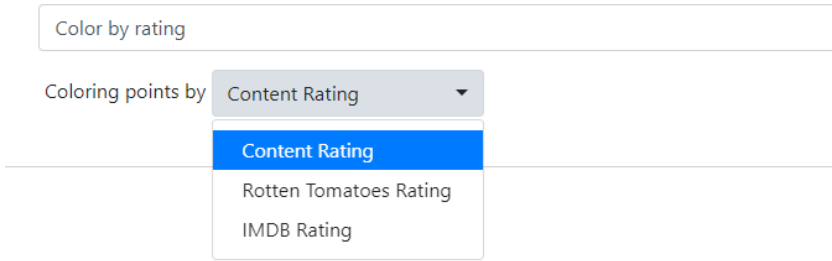


Figure 3.3: Illustration of an attribute-level ambiguity widget in DataBreeze.

For TouchWave, our very first project with the (touch-only) iPad, we created a set of touch gestures for only one type of visualization (i.e., stacked graphs), which did neither rely on any WIMP elements nor introduce modes. We first generated a list of manipulations (operations) that would be available. We then made two lists: one for available visual elements (e.g., single layers, stacked graphs, background) and another for available elementary interactions (e.g., tap, pinch, tap & hold). Finally, we determined the mapping between manipulation and a combination of a visual category and an elementary interaction (e.g., zoom == background + pinch). This process allowed us to pick suitable interactions for each manipulation while avoiding conflicts. However, this interaction set is not likely to scale to other visualizations that provide a broader set of operations.

The evolution of Tangere system also highlights the challenges in maintaining interaction consistency across multiple visualizations. Initially focusing on one visualization type (i.e., scatterplots), we designed a gesture set that supported frequently used operations (e.g., changing encodings, filtering, selection) with the scatterplot (Sadana and Stasko, 2014). However, when different visualizations were incorporated into the system to enable richer analysis, it was challenging to support the same interactions for an operation across different charts. For instance, in the initial scatterplot prototype, we used the action of dragging a finger along an axis to select points in that range of values. The same interaction gesture on a bar chart axis, however, sorted the bars along that axis. Thus, when the two chart types were supported in the same

interface, there was a conflict in interactions. In this case, to maintain global consistency across different charts, we eventually used the dragging-along-axis gesture for selection and introduced a specialized gesture to sort a bar chart (pressing/holding a finger on the axis of a bar chart to enter a transient “sort mode” and then swipe to sort) even though this was locally suboptimal.

In our experience, two possible approaches can help address this issue: (1) incorporate elements of WIMP interfaces and (2) integrate other modalities. Tangere as well as other systems such as Kinetica (Rzeszutarski and Kittur, 2014) adopt the former approach and introduce basic WIMP features such as tab-like menus for switching interaction modes, “remove/keep-only” buttons specifying selections, or control panels with sliders and dropdown menus to support filtering. Although this approach helps to improve consistency in interactions, heavily relying on WIMP-style widgets may limit the fluidity of interactions and, in some cases, may simply be infeasible due to practical constraints (e.g., a smartphone screen may not provide enough space to add menus).

In contrast, with systems like SketchInsight (Lee *et al.*, 2015) and InChorus (Srinivasan *et al.*, 2020a), we incorporated additional input modalities such as pen and/or speech. We found that the addition of new modalities makes it easier to retain locally optimal interactions with individual charts while still preserving global consistency in interactions across charts. For instance, with InChorus, we could support the dragging action for selection and sorting by allowing users to drag along an axis with a pen to select or with a finger to sort, respectively. Furthermore, by complementing similar simplistic pen and touch interactions with speech, we could support additional capabilities like sorting by an attribute or criteria-based filtering of selected data marks.

While an important aspect of designing post-WIMP visualization interfaces is devising natural or even novel interactions, to ensure the resulting systems’ practical viability, it is important to support consistency in interactions across different views in an interface. As discussed above, we have previously designed sets of interactions by reviewing existing systems and following their lead. However, while doing so, we found that we lacked coherent terminology to describe and compare alternative interactions.

Operation	Parameters (I , D)	Targets (T)	Instruments (DI or PI)	Sample keywords (K)	Interaction patterns	Examples	Modalities T P S
O1	Bind multiple attributes to an encoding	X/Y axes Color legend Attribute pills: II	Axis title region, Color legend title, Attribute pills: II	<i>Add, include, group by, split by</i>	I1: Drag II to DI	Drag worldwide gross to X-axis title	✓
					I2: Point on DI and tap II	Point on X-axis title and tap major genre	✓
					I3: Point on DI and write P	Point on X-axis title and write running time	✓
					I4: Point on DI and speak P	Point on color legend title + "Content Rating"	✓
					I5: Speak < T , P >	"Show creative type on the X-axis"	✓
O1	Bind multiple attributes to an encoding	X/Y axes Attribute pills: II	Axis title region, Attribute pills: II	<i>Add, include, group by, split by</i>	I7: Speak < T , K , P >	Point on X-axis title + "Add budget and gross"	✓
					w/ modifier active:	"Group X-axis by content rating"	✓
					I8: Drag II to DI	w/ modifier active: Drag major genre to X-axis title	✓
					I9: Point on DI and tap II	Point on X-axis title and tap IMDb Rating	✓
					I10: Point on DI and write P	Point on X-axis title and write Running Time	✓
O1	Remove attributes from an encoding	X/Y axes Color legend Color legend title	Axis title region, Color legend title	<i>Remove, unbind, color, delete</i>	I11: Erase Px from DI	Erase X-axis title to remove attributes mapped to X-axis	✓
					I12: Point on DI and speak < K , P >	Point on color legend title + "Genre"	✓
					I13: Speak < T , K , P >	"Remove budget from the Y-axis"	✓
					I14: Point on DI and write P	Point on Y-axis title and write sum	✓
					I15: Point on DI and speak P	Point on Y-axis title + "max"	✓
O2	Change data aggregation level	X/Y axes	Axis title region	<i>Show the average values on Y"</i>	I16: Speak < T , P >	Swipe downwards on Y-axis title to sort in descending order by Y-axis attribute values	✓
					I17: Swipe on DI , direction determines P	"Show the average values on Y"	✓
					I18: Point on DI and speak < K , P >	Point on Y-axis title + "Sort" to sort in ascending order (default system value) by Y-axis attribute values	✓
					I19: Speak < T , K , P >	"Group X-axis in decreasing order"	✓
					I20: (select 1) Erase DI	Erase bar in a bar chart to remove a data category	✓
O3	Filter by criteria values	Marks	Color legend rows remove, keep only	<i>Filter, exclude others remove, keep only</i>	I21: w/ DI selected, speak < K >	Select points in a scatter plot + "remove other" to filter unselected points	✓
					I22: Erase P from DI	Erase Action row from color legend to remove action movies	✓
					I23: Speak < K , P >	"Exclude horror movies"	✓
					I24: Long press DI	Long press on a bar in a bar chart to see its value	✓
					I25: Drag along DI	Drag along Y-axis ticks in a horizontal line chart to see values for a specific timestamp	✓
O4	Get mark details by value	Attribute values	Marks	Axis scales	I26: Speak < P >	"Switch to line chart"	✓
					I27: Drag II to DI	Switch to line chart	✓
O5	Change chart type	Chart type	Canvas		I28: Drag II to DI	Switch to line chart	✓
					I29: Drag II to DI	Switch to line chart	✓

Table 3.1: Multimodal interactions for low-level operations during visual analysis with InChorus (Table from (Srinivasan *et al.*, 2020a)). Operations categories are O1: Bind/unbind visual encodings, O2: Modify axes, O3: Filter, O4: Get details, and O5: Change chart type. Unless explicitly specified as an indirect instrument (II), all instruments are direct instruments (DI). An asterisk (*) indicates a parameter is optional. The rightmost column displays modalities (T: Touch, P: Pen, S: Speech) used in an interaction pattern.

To fill this gap, when designing InChorus (to consistently support a wide array of canonical visual analysis operations in a multimodal setting), we proposed an approach of thinking about interactions with post-WIMP visualization systems in terms of *operations*, *parameters*, *targets*, and *instruments*. Specifically, people interact with visualization systems through a set of low-level operations (e.g., binding an attribute to an encoding, sorting) to accomplish their high-level tasks (e.g., answering data-driven questions, creating specific visualizations). These operations typically require parameters (e.g., sorting order, attributes, encodings) and operate on one or more targets (e.g., selected marks, axis, canvas). Finally, operations are mediated through instruments in the interface (e.g., marks, axes scales). These instruments can be direct (i.e., when the target itself mediates an operation) or indirect (i.e., when an operation is performed on a target through a separate instrument).

Table 3.1 showcases how these concepts can be used to model interactions in the context of a minimalistic visualization tool with basic interface elements such as axes, marks, and legends. The interactions in Table 3.1 are not an exhaustive or definitive set, but they can serve as an initial resource to design and track interactions in post-WIMP visualization interfaces. Even though it may not be possible (or even desirable) to design a universal interaction set, it might be beneficial to develop an interaction model that can characterize and describe interactions that can be commonly used across systems. This can help to improve learnability and discoverability of post-WIMP interactions employed in visualization systems.

3.4 Designing Speech-driven Interactions with Visualizations

A defining characteristic of our work involving natural language interaction with visualizations is that we focus on speech as an input modality (as opposed to typed input in other systems). Although the two styles of input share many commonalities in terms of interpretation techniques, some important distinctions exist between the two from a system design and development standpoint. In this section, we reflect on our experiences related to these differences and hope this discussion can guide future work exploring speech-based interactions with InfoVis.

A key challenge with speech input from a system standpoint is simply to know when to trigger speech recording. Ideally, a system “listens” all the time² and accurately extracts commands meant for the system. This is extremely difficult in practice, however. Thus, most systems use a “wake” word to trigger command listening.

In the initial version of Orko, we allowed people to trigger speech by using the wake-word “*Hey, System*” (similar to “*Ok, Google*,” “*Hey, Siri*,” “*Alexa*”). However, during user study sessions we observed that participants would forget to trigger speech recording and issue a command without noticing that the system is not receiving their input. Furthermore, even when participants would utter the wake-word, the system would sometimes fail to recognize it and thus not record the subsequent command. Particularly because Orko was a multimodal system, we also noticed that these issues with speech triggering were detrimental to naturally occurring interactions involving both touch and speech. Based on these observations, when designing latter systems like InChorus and DataBreeze, we incorporated additional mechanisms to trigger speech recording when users performed actions like pointing (or holding) and selections.

In addition, unlike its typing counterpart, speech is not suitable for incorporating assistive features like autocomplete or command recommendations (e.g., Figure 3.4). Such suggestions not only help people complete sentences faster, but can also provide hints about what can be stated. This means that speech-based systems need to incorporate additional techniques to convey what can be said.

Furthermore, speech-based systems also suffer from the unique challenge of speech-to-text transcription errors, which remains the prominent source of errors in any voice interface (Myers *et al.*, 2018). Hence, natural language interpretation modules in speech-based systems need to be designed with recognition errors in mind and ideally also provide appropriate mechanisms to adjust previously input queries (e.g., by allowing users to point on an incorrectly recognized word and change just that one word as in Voice Typing (Kumar *et al.*, 2012)).

²To focus our discussion on interaction, we do not consider privacy concerns that may arise from an “always on” voice interface.

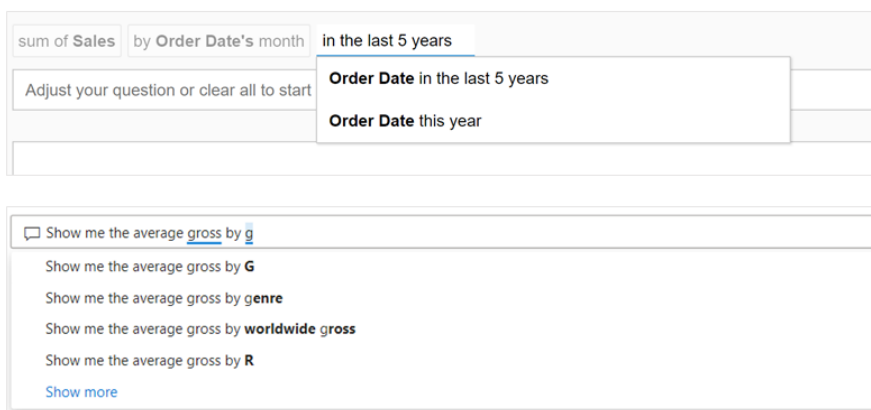


Figure 3.4: Examples of query suggestions when entering queries in (top) Tableau Ask Data and (bottom) Microsoft Power BI Q&A.

Finally, an additional challenge in designing speech-based interactions is to ensure the delivery of system feedback. Following the feedback techniques in other desktop-based natural language interfaces for visualization, in our systems we provided textual feedback in response to spoken queries (e.g., Figures 2.11, 2.12). However, since users may not be looking at the text box while providing spoken input and especially right after they complete a speech command, we observed that users often fail to notice the feedback text. This, in turn, leads to users repeating their utterances when they do not realize the subtle changes the system made, or when they do not understand what caused the error (even if the feedback text explains this error). Thus, when designing speech-based visualization systems, it is important to consider more effective feedback techniques that can overcome these issues.

3.5 Evaluating InfoVis Systems with Post-WIMP Interactions

Evaluation practices in InfoVis have always been heavily inspired by HCI research. Evaluations began with a focus on controlled experiments and usability studies (Plaisant, 2004) and this focus is still prevalent today (Isenberg *et al.*, 2013b; Lam *et al.*, 2012). Plaisant argued that evaluation in visualization should place more emphasis on understanding users,

tasks, and their real problems. Plaisant also asserted that, in contrast to common UI evaluations, InfoVis faces several challenges. The process of data analysis and discovery is akin to many creative tasks as it requires taking on multiple viewpoints as well as long-term engagement with visualizations, deriving new questions and changing analysis approaches. In addition, if one focuses on insight discoveries as an evaluation metric, these instances of deep insight are rare, and thus it is difficult to measure the success of an analysis interface. Several approaches and methodologies have since then been proposed specifically targeted at InfoVis evaluations, and a methodology overview (Carpendale, 2008) and survey manuscripts (e.g., (Lam *et al.*, 2012; Isenberg *et al.*, 2013b)) give general advice on how and which kind of evaluation goals to choose or which methodologies to pick.

However, as the application domains and usage contexts of InfoVis broaden, several researchers (Stasko, 2014; Thudt *et al.*, 2017; Ren *et al.*, 2018) acknowledge the limitations of traditional comparative studies and suggest the use of more appropriate evaluation methods and metrics by properly adapting existing methods or devising new ones.

We face similar issues and challenges in evaluating post-WIMP visualization systems. While it is tempting to measure “improvement” by comparison, traditional comparative studies usually are not appropriate for several reasons. As we discussed earlier in Section 3.1, a comparative lens may not be appropriate when the improvement comes from the synergy of factors (e.g., different modalities, WIMP vs. post-WIMP) that have their complementary characteristics. Unlike comparing a few factors with low-level tasks, designing a comparative study to holistically evaluate a system that supports high-level tasks poses practical challenges, such as the lack of comparable baseline systems and difficulty in preparing different but equivalent systems for a fair comparison. Typical evaluation metrics (i.e., time and accuracy) may not be suitable for evaluating these systems if the goal is to offer new experiences that were not possible before (Wang *et al.*, 2019). Furthermore, such metrics may be insufficient if we want to gain a deeper understanding of if and how people can use and benefit from the proposed interaction.

In our own work, we thus strove to select and adapt evaluation methods to achieve their study goals, supporting their intended con-

tributions. Table 3.2 shows a summary of the evaluation methods we applied to our projects. Below, we reflect on a few projects discussing the challenges, limitations, and benefits of different methods and potential improvements we can make.

We often used a qualitative approach to investigate how people adopt and react to a system. For example, a study on Cambiera (Isenberg *et al.*, 2010), a tabletop system for analyzing document collections, adopted an observational methodology. Through an open coding-based analysis of session videos, we characterized a range of collaboration styles from close to loose collaboration, and described how the workspace was used and how people worked off the digital tabletop interface (taking notes and communicating). These observations led to discussions on the value of designing interfaces to encourage close collaboration and how to design interfaces that allow groups to switch collaboration styles.

A study of SketchInsight (Walny *et al.*, 2012) also took a qualitative approach, observing how people respond to a pen and touch whiteboard for creating and manipulating three types of charts—bar charts, line charts, and scatterplots. Furthermore, to ensure robust and flexible experiences while simulating a realistic context before implementing a wide range of expensive recognizers, we employed a Wizard-of-Oz study (Dahlbäck *et al.*, 1993). Through the qualitative analysis of the interaction our study participants used, we gained valuable insights about pen and touch interactions in the context of learnability and the interplay between pen and touch gestures, which in turn helped us design and develop the SketchInsight system (Lee *et al.*, 2015). In contrast to the Cambiera study mentioned above, we did not perform a thorough qualitative analysis using the recorded session videos. Instead, we focused on a low-cost evaluation approach that allowed us to identify common usage patterns as well as clear system improvements and remaining usability issues quickly.

We performed low-cost and light-weight qualitative analyses in several projects, including Orko, DataBreeze, and InChorus. We considered the design and implementation of the system itself as a main contribution and the user study as a means to assess the general usability of the proposed interactions and to gather subjective feedback, while collecting the most obvious usage patterns. We note that for the DataBreeze

Evaluation Method	Project	Notes
Qualitative study	Cambiera (Isenberg <i>et al.</i> , 2010)	intelligence analysis task for pairs
	SketchInsight Study (Walny <i>et al.</i> , 2012)	a Wizard-of-Oz study with eight high-level tasks
Exploratory study	SketchInsight (Lee <i>et al.</i> , 2015)	six tasks with three levels of difficulties
	Orko (Srinivasan and Stasko, 2018)	10 tasks: a mix of fact verification (Gao <i>et al.</i> , 2015) and entity identification
	DataBreeze (Srinivasan <i>et al.</i> , 2020b)	an open-ended task with a colleges dataset
	InChorus (Srinivasan <i>et al.</i> , 2020a)	two task types: (1) replication & value identification and (2) fact verification (Gao <i>et al.</i> , 2015)
	Orchard (Eichmann <i>et al.</i> , 2020)	eight tasks covering 10 low-level tasks for multivariate network data (Pretorius <i>et al.</i> , 2014)
“Comparative” exploratory study	Multitouch Study (North <i>et al.</i> , 2009)	compared surface, physical, and mouse
	ActiveInk (Romat <i>et al.</i> , 2019a)	compared ink and Prefix & Postfix of ActiveInk; high-level tasks involving hypothesis generation, browsing, and data investigation
	VoyagerInk (Kim <i>et al.</i> , 2019c)	indirect comparison
	Study using Orko (Saktheeswaran <i>et al.</i> , 2020)*	mixed-design: (1) touch vs. touch & speech and (2) speech vs. touch & speech
Comparative controlled experiment	TouchPivot (Jo <i>et al.</i> , 2017)	a between-subjects design with two types of tasks
	SketchStory (Lee <i>et al.</i> , 2013)	combined it with a usability study for the authoring experience
	Tangere Study (Sadana <i>et al.</i> , 2018)	compared against Tableau’s Vizable focusing on discoverability
Other	TouchWave (Baur <i>et al.</i> , 2012)	presented usage scenarios with two datasets

Table 3.2: Summary of evaluation methods employed in our research on post-WIMP interaction for InfoVis. We did not perform an evaluation for the Tangere (Sadana and Stasko, 2014), WordlePlus (Jo *et al.*, 2015), Tangraphe (Thompson *et al.*, 2018), and VisWall (Agarwal *et al.*, 2019) systems. *This study did not evaluate the Orko system but used the system to compare different input modalities.

project, we followed an iterative design process that helped us identify and refine a set of design goals we needed to accomplish. We iterated on DataBreeze’s design and implementation with six design sessions within the project team as well as with external people (i. e., graduate students at Georgia Tech). It would still be valuable to conduct more robust qualitative studies with these systems preferably in a more realistic setting as such studies would deepen our understanding on how people work with or would like to work with our systems.

When appropriate, we also performed comparative controlled studies because they could help us evaluate a specific aspect of a post-WIMP interface. For example, to evaluate TouchPivot against the widely-used pivot interfaces, we compared the commonly used performance metrics—*time* to complete a task and number of *errors* made during task completion—as well as the SUS (System Usability Scale) (Brooke, 1996), a low-cost usability scale that is increasingly used for global assessments of systems’ usability.

On the other hand, for the SketchStory project, our goal was to examine the subjective level of engagement of SketchStory for both the audience and presenters. Therefore, we asked our study participants to answer questions intended to measure the subjective level of engagement using a Likert scale: our presenter participants answered one more question about the ease of learning of the tool. The use of metrics such as engagement, enjoyment, and memorability that go beyond notions of utility are gaining momentum in the community (Saket *et al.*, 2016; Wang *et al.*, 2019), and they seem to be particularly useful when evaluating post-WIMP systems.

In post-WIMP systems that are designed to be more fluid and natural to use, one may be curious how new users will learn and adapt to a system, perhaps even with very little training. Today’s users frequently attempt to use a system without reviewing a tutorial or manual. To simulate such a scenario, our evaluation of the Tangere system took a discoverability-focused approach (Sadana *et al.*, 2018). Participants in the study were only briefly introduced to the system’s functionality and operations, and a key metric was how many of the operations they discovered and eventually used.

3.6 Towards Technology-Inspired Research

As described in Section 2, much of our research on post-WIMP interaction (Figure 3.1) has been technology-inspired research that was sparked and enabled by the advancements in both hardware and software technologies. Technology-inspired research stands in contrast to problem-oriented research that is motivated by solving specific end-user problems and is thus also highly valuable (Shneiderman, 2016). Using a problem-oriented research approach, the visualization research community has collectively made considerable advances with solutions that are different (or new) in significant respects and better than existing (so-called state-of-the art) solutions. Problem-oriented research is an accepted, advocated, and often expected approach in the visualization research community.

However, with the prevalence of and expectations around problem-oriented research, it is often difficult to publish technology-inspired work (especially in the visualization community). While pursuing the publication of our post-WIMP visualization systems, we sometimes received reviews and feedback, asking for the clarification of the end-user problems we are addressing or to specify and characterize the exact end-user group we target. These are problematic requests for technology-inspired research because this type of work is often meant to inspire *new* uses and experiences: restricting to a specific user group or tasks may hinder innovation and the broad applicability of techniques later. Requests for defining end users and their tasks and problems also mean that future user studies on novel post-WIMP interfaces should consider these. However, sometimes the right group of end users may not yet exist. For example, if we test speech-based systems with people who have been accustomed to the current tools designed for the desktop environment now, we may find that they are not comfortable talking to a computer system. We might abandon the idea based on a study suggesting that using speech may not be appropriate. Yet, a new generation who is exposed to speech-enabled technologies at an earlier age might be very receptive to speech-incorporated multimodal systems in the future.

Researchers in HCI have acknowledged the challenges in evaluating problem-driven research with existing methods. For example, Olsen Jr

(2007) articulated a set of criteria for evaluating new user interface systems in addition to discussing the limitations of simple usability testing. Greenberg and Buxton (2008) discussed the danger of an unquestioning adoption of the doctrine of usability evaluation. They argued that traditional usability evaluation of innovations does not provide accurate indication about its adoption over time. In addition, as argued by Lee *et al.* (2019) we also need to embrace assessments of failed attempts at creating novel interfaces and technology as these failed attempts can provide incredibly helpful lessons for future design and research. In addition to diligently documenting the rationales behind the research and the justification of why it could have been expected to succeed, an assessment must include the analysis of reasons behind the failure and lessons for future projects.

In summary, we fear that a rigid expectation of problem-driven research will lead the visualization research community to tackle challenges that may not necessarily spark true innovations on post-WIMP interactions. If we instead embrace technology-inspired research, we can use the creative innovations that post-WIMP enables to move towards grand high-level visions, such as van Dam's "disappearing" post-WIMP interfaces (van Dam, 1997) or Weiser's ubiquitous computing (Weiser, 1999), for and with visualization systems. We also assert that novel InfoVis research with post-WIMP interaction should consider (or even create) *new* user groups and *innovative* experiences.

4

Looking Ahead

As we discussed earlier, research on post-WIMP interaction for InfoVis has recently made great progress. However, we see post-WIMP interaction as still far from realizing its full potential, holding great opportunities for researchers and practitioners in the InfoVis and broader HCI communities, and thus calls for further research endeavors. In this section, we discuss research that can facilitate post-WIMP interaction in InfoVis and the opportunities post-WIMP interaction can bring, expanding the boundaries of InfoVis and enabling new experiences that were not possible with WIMP.

4.1 Creative Adaptations from Broader HCI Research

While post-WIMP interaction has received increased attention within visualization research over the past decade, the broader HCI communities have explored such interfaces for several decades. HCI researchers have extensively investigated post-WIMP systems for a wide range of domains including graphics & image editing (e.g., (Hauptmann, 1989; Pausch and Leatherby, 1991; Laput *et al.*, 2013)), 3D modeling (see (Olsen *et al.*, 2009) for a survey), education (e.g., (LaViola Jr and Zeleznik, 2006; Bott and LaViola Jr, 2010; Yoon *et al.*, 2014; Yoon and Mitros,

2015; Yoon *et al.*, 2016)), collaboration & productivity (e.g., (Mynatt *et al.*, 1999; Klemmer *et al.*, 2001; Ju *et al.*, 2008)), and military simulations & crisis planning (e.g., (Cohen *et al.*, 1997; Cohen and McGee, 2004; Sharma *et al.*, 2003)), just to name a few. This wealth of research presents an opportunity to leverage interfaces and interactions that have demonstrated their potential benefits in other domains and apply them to visualization systems in a creative and compelling manner.

As shown by some examples from existing visualization research, such application can enable novel visualization tools and workflows that amplify data exploration and data-driven communication capabilities offered by current systems. The Zooids platform (Le Goc *et al.*, 2016), for instance, was initially developed to enable the creation of swarm user interfaces, a class of interfaces comprised of many autonomous robots that handle both display and interaction. It was later extended to support data analysis scenarios (Le Goc *et al.*, 2019), leading to the idea of dynamic composite physicalizations, where data are represented as actuated tangible tokens (specifically, mini-robots). Although they were largely an extension of the initial idea, dynamic composite physicalizations enable new opportunities to design tangible visualizations and support collaborative analysis scenarios in household settings.

Another example is the object-oriented drawing (OOD) idea (Xia *et al.*, 2016) that allows direct editing of graphical objects through touch interaction (as opposed to indirect editing via menus). By applying the idea of OOD to visualization authoring, DataInk (Xia *et al.*, 2018) facilitates the fluid creation of expressive infographics through pen and touch interaction that enables drawing and direct manipulation on a digital canvas. Along this line, one promising area where post-WIMP interaction could enable natural interaction is visualization authoring for communication purposes. A new generation of visualization authoring systems have recently emerged to support expressive InfoVis without textual programming (Satyanarayan *et al.*, 2019), including Lyra (Satyanarayan and Heer, 2014), iVisDesigner (Ren *et al.*, 2014), Data Illustrator (Liu *et al.*, 2018), and Charticulator (Ren *et al.*, 2019). However, these systems all still employ the WIMP metaphor. Conversely, systems like DataInk (Xia *et al.*, 2018), DataToon (Kim *et al.*, 2019a), and DataQuilt (Zhang *et al.*, 2020) (Figure 2.8) demonstrate

the creative potential of post-WIMP interaction leveraging pen and touch. Specifically, DataToon offers a fluid storyboarding experience for data comics by blending analysis and presentation of dynamic networks in a unified environment supported with pen and touch interactions. DataQuilt allows authors to extract components and stylistic features from raster images (e.g., sketches, paintings, photographs) and repurpose them as elements of a graphical language, to encode and decorate data. Extending this line of work, it would be worthwhile to design visualization authoring interfaces that support more natural interactions, drawing from the rich body of research on multimodal interfaces for graphics applications (e.g., (Gourdol *et al.*, 1992; Pausch and Leatherby, 1991; Subramonyam *et al.*, 2018)).

Inspired by active reading (Adler and Doren, 1972) of text, which is extensively studied in HCI (for a summary, see Table 1 in Chen *et al.* (Chen *et al.*, 2012)), Walny *et al.* (2017) proposed and explored the concept of active reading of visualizations. They first observed whether and how people actively read visualizations using paper-based node-link visualizations. They then studied the marking and creation actions (i.e., drawing and erasing marks) identified from their observational study because these actions are involved in the higher-level goals of decoding and analyzing a visualization. One of the design opportunities discussed by the authors is how to best support active reading in digital environments, suggesting support for freeform annotation and additional active reading actions. For example, ActiveInk (Romat *et al.*, 2019a) enables people to seamlessly transition between exploring data and externalizing their thoughts using pen and touch, i.e., supporting active reading behaviors, during data exploration. Kim *et al.* (2019c) also recently investigated the types and use of digital externalization during the data analysis process via two exploratory studies. Further research is needed because these projects only scratch the surface of the potential benefits of post-WIMP interaction for active reading and deeper understanding of visualizations.

In summary, the aforementioned examples highlight how adapting interface and interaction techniques demonstrated in the broader HCI literature to visualization research presents twofold benefits. First, repurposing interfaces and interactions to data visualization contexts

helps validate and extend the software and hardware architecture accompanying the initial idea. For example, the Zooids platform (Le Goc *et al.*, 2016) was extended to support cross-device interaction and configuration via a tablet to enable collaborative data exploration. Second, from a visualization research standpoint, this repurposing promotes exploring and creating new interfaces and experiences that may not have been conceived by solely focusing on visualization-specific tasks and interfaces. For example, interface-level operations such as moving points in DataBreeze (Srinivasan *et al.*, 2020b) are not typical for visualization tools but stem from scenarios common in graphics manipulation. Therefore, going forward, an interesting opportunity for continued research lies in exploring how more novel ideas presented in broader HCI and UI communities can be brought to data visualization to enhance human-data interaction.

4.2 Toolkits and Frameworks for Prototyping Post-WIMP Visualization Interactions

A number of programming toolkits and libraries (e.g., Prefuse (Heer *et al.*, 2005), Protovis (Bostock and Heer, 2009), D3.js (Bostock *et al.*, 2011), Vega (Satyanarayan *et al.*, 2015), Vega-Lite (Satyanarayan *et al.*, 2016), Stardust (Ren *et al.*, 2017)) have been developed to aid data visualization. These toolkits facilitate developers with varying backgrounds and programming expertise to create and share visualizations. Besides mapping data to visual channels (e.g., size, position, color), with some additional programming or configuration, some of these toolkits also provide support for implementing simple interactions such as selections and brushing-and-linking. However, implementing post-WIMP interactions involving different modalities such as pen, touch, speech, or gestures, remains a massive implementation effort imposed entirely upon visualization developers.

To implement post-WIMP interaction with visualizations, developers currently typically need to 1) write the low-level code to capture different input events (e.g., active touches on the screen, voice input), 2) implement custom modules to parse these events (e.g., infer gestures based on touch coordinates, parse natural language queries using NLP

techniques), and then 3) write the code to propagate the result of the input parsing module back to the user interface and change the state of the view. While some toolkits such as hammer.js (Schmitz *et al.*, 2020), NLTK (Loper and Bird, 2002), WebSpeech API (*Web Speech API* 2019) assist capturing and parsing input from different modalities, unlike with visualization-focused toolkits mentioned above, these interactions are not implicitly bound to the visualization. In other words, developers still need to write the code to map the output from these toolkits onto the view, which can be both challenging and tedious. This added effort of writing custom code to capture, interpret, and propagate post-WIMP interactions makes it difficult to rapidly prototype novel interactions and interfaces, hindering research progress and widespread deployment of research prototypes.

Some projects have recently begun to address these challenges: example toolkits include EasyPZ.js (Schwab *et al.*, 2019b), DXR (Sicat *et al.*, 2018), and NL4DV (Narechania *et al.*, 2021). EasyPZ.js allows developers to add basic (pan and zoom) navigation interactions in otherwise static web-based visualizations on mobile devices with as little as one line of code. DXR enables the development of expressive and interactive visualizations in Unity that can be deployed in AR/VR environments. This enables the creation of human-data interaction experiences in a breadth of analytical (e.g., sports data analysis, flight trajectory planning) and casual (e.g., shopping, museum tours) scenarios. The NL4DV toolkit helps prototype natural language interfaces for visualization. Given a dataset and corresponding natural language queries, NL4DV infers relevant information, such as data attributes and analytic tasks as well as visualizations that are relevant to the query. Visualization developers can parse this inferred information to augment their existing tools with natural language or create novel natural language-driven interaction experiences. While these toolkits are promising initial efforts, they currently cater to a small set of modalities and interactions and also require programming know-how to implement and test novel interactions.

Going forward, it is essential to develop more toolkits that specifically make it easier to implement and test interactive visualizations in post-WIMP settings. In fact, following work on applications for in-

teraction prototyping in the broader HCI community (e.g., (Parnami *et al.*, 2016; Ledo *et al.*, 2019)), an interesting line of research also involves exploring interfaces for non-programmers (e.g., designers and practitioners) to prototype novel interaction experiences with visualizations. Besides implementing the toolkits and systems, theoretically formulating frameworks and grammars that can seed their development is also an important research topic.

4.3 Context-Aware Visualization Interfaces

In Section 4.1, we discussed the potential benefits of incorporating learnings and advancements from broader HCI and UI research within the domain of InfoVis. Along this line, one topic that has gained significant interest in UI systems but little consideration in InfoVis systems so far is the use of sensing technology. In particular, an extensive amount of research has been carried out in UI research to investigate the idea of context-aware sensing (e.g., (Hinckley *et al.*, 2000; Hinckley *et al.*, 2016; Zhang *et al.*, 2019; Laput, 2019)). Applications can be cognizant of the context in which a device is being used (e.g., the orientation of a device, the angle at which a device is positioned, the noise level surrounding a device). Especially with the growing interest for visualizations on mobile devices (Lee *et al.*, 2018a; Choe *et al.*, 2019; Lee *et al.*, 2020), exploring context-aware interactions with visualization systems is a fertile research topic with many interesting avenues for investigation.

In its simplest form, as envisioned by Hinckley *et al.* (2000), device context such as screen size and orientation can be used to create responsive visualizations that automatically reconfigure themselves to adapt to the active viewport (e.g., (Hoffswell *et al.*, 2020; Wu *et al.*, 2021)). In addition, explicit movements like tilting can also be used to interact with visualizations (e.g., sideways tilting can be used to navigate a timeline view).

A more advanced example of a context-aware visualization tool can be one where both the interface and interaction adapt to the available input modalities. Consider the case of a tablet-based visualization system like InChorus (Srinivasan *et al.*, 2020a) that supports multimodal interactions involving pen, touch, and speech. Users may not always

have access to all three modalities at all times, however (e.g., one may be in a noisy environment that is unsuitable for voice interaction). While one approach to overcome this issue is to design for equivalence in interactions across modalities, it has been shown that equivalent multimodal systems can limit the range of interactions that could be supported (Oviatt, 1999; Srinivasan *et al.*, 2020a). As an alternative, the system could leverage the tablet's sensors to detect the available modalities and adapt the interactions accordingly. For example, if pen and touch are the only available modalities, the system can provide a WIMP-style panel for filtering. On the other hand, if speech input can be supported this panel can be hidden (since filters can be applied through speech) to maximize the screen space allotted to the visualization itself.

A related, but likely to be more challenging, research direction is to leverage implicit user interactions. The systems and interactions discussed in this monograph cover a breath of examples highlighting scenarios where people explicitly interact with visualization through different modalities such as pen, touch, and speech, among others. However, besides these explicit actions, when working with visualization systems, people also tend to implicitly use other modalities like eye-gaze (e.g., visually inspecting a specific region of a chart before interacting with it via touch). Yet, much of the work on gaze-based interactions with InfoVis (e.g., (Streit *et al.*, 2009; Okoe *et al.*, 2014)) has only considered explicit eye-gaze input instead of how it could implicitly augment interactions with other forms of input. This is in stark contrast to HCI research where many examples have been presented to illustrate the potential of combining gaze with more explicit modalities like pen, touch, and VR controllers (Pfeuffer *et al.*, 2015; Pfeuffer and Gellersen, 2016; Pfeuffer *et al.*, 2017b). Thus, going forward, it would be interesting to explore how InfoVis systems could model user interest from implicit interactions via modalities like eye-gaze and leverage them to guide subsequent user actions.

To summarize, as devices become smarter with more sensors and input recognition technologies improve, visualization systems can detect richer interaction contexts from both devices and implicit user actions. These contexts can, in turn, be used to improve the user experience and design new interaction techniques.

4.4 Post-WIMP Interfaces for Data Manipulation

Following the visualization reference model by Card *et al.* (1999), visualizations can be thought of essentially as adjustable mappings from data to a visual form (Figure 4.1). Thus far, post-WIMP visualization interfaces have predominantly focused on the second half of this model (the ‘visual form’), exploring novel ways to specify visual mappings and perform view transformations. However, a major challenge faced by data scientists and data enthusiasts (Hanrahan, 2012) is getting the data into a format that can be visualized (Kandel *et al.*, 2012; Kandel *et al.*, 2011). As pen- and/or touch-enabled devices become more prevalent, an exciting topic for future research is to explore how we can support an entire analysis workflow by providing better data manipulation with post-WIMP.

Initial work in this space has dabbled into supporting data preparation and cleaning by enabling novel interactions with tables. For example, Tableau (Zraggen *et al.*, 2016) illustrates how pen and touch interactions can be used to perform operations like editing data columns by sketching. Srinivasan *et al.* (2018) also envision a scenario where multimodal interactions combining touch and speech support frequently used spreadsheet operations including selecting and editing cells. Hoffswell and Liu (2019) present a set of touch gestures to support table repair operations such as merging, splitting, or inserting rows/columns. While these initial efforts hint at the potential benefits of using alter-

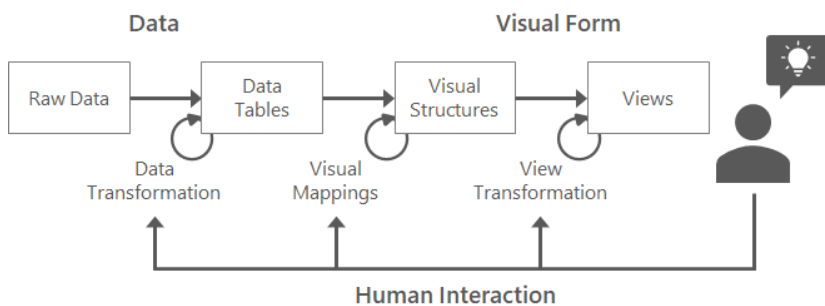


Figure 4.1: Reference model for visualization (redrawn from Card *et al.* (Card *et al.*, 1999)).

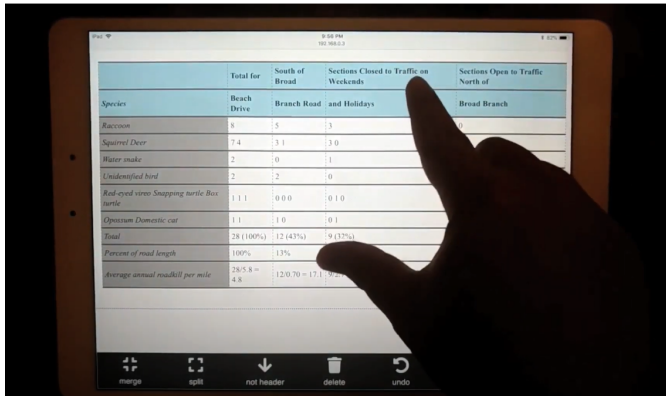


Figure 4.2: An example from Hoffswell and Liu’s tablet-based system for editing incorrectly scanned data tables from PDF documents (Hoffswell and Liu, 2019). In this case, the user is performing a pinch gesture to merge rows that were incorrectly split while reading a PDF table.

native modalities for data manipulation, exploring the role of different input modalities during data cleaning and transformation remains an exciting open topic for future work.

4.5 Interacting with Data Physicalizations

Data physicalizations are “physical artifacts whose geometry or material properties encode data” (Jansen *et al.*, 2015). They stand in contrast to purely virtual visualizations we can see and interact with on displays (e.g., monitors, mobile phones) and in virtual reality with head-mounted displays. For centuries people have built and used data physicalizations to keep track of and explain their environment as well as to explore scientific phenomena without any computer support. The extensive data physicalization list maintained by Dragicevic & Jansen (Dragicevic and Jansen, 2012) shows examples including the Mesopotamian Clay Tokens (5500 BC), Maxwell’s “thermodynamic surface” (1871), and John Kendrew’s molecular representations (1951).

Starting in early childhood, humans learn how to manipulate physical objects with their hands. These abilities transfer to interactions with data physicalizations as 3D physical objects that can be manip-

ulated and experienced with one's hands. As such, interactions with data physicalizations are beyond WIMP and could be considered more “natural” than interactions with virtual objects. With the advent of computer-supported maker equipment, such as laser cutters and 3D printers, the automated creation of data physicalizations are not only possible but also becoming easier. For example, MakerVis (Swaminathan *et al.*, 2014) is a data physicalization design tool that allows people to enter data tables and it produces digital fabrication-ready design files that could be loaded into a laser cutter. Figure 4.3 shows an example visualization built with MakerVis, a laser cutter, and some spray paint. However, research on interactions with data physicalizations is sparse.

Considering the unique characteristics of data physicalization, post-WIMP interaction and data physicalization can have a synergy in making it easy for people to understand data while making the human-data interaction experience more enjoyable. For example, “constructive visualization” (Huron *et al.*, 2014a) demonstrates how letting people use and manipulate simple elements to create a visualization facilitates the understanding of data. Different types of interactions that are part of the construction of token-based visualizations have been studied (Huron *et al.*, 2014b) and showed the wide variety of interaction tasks people adopted with tokens including construction, computation, and storytelling-related interactions. Others have explored possibilities to create and interact with dynamic data physicalizations such as the

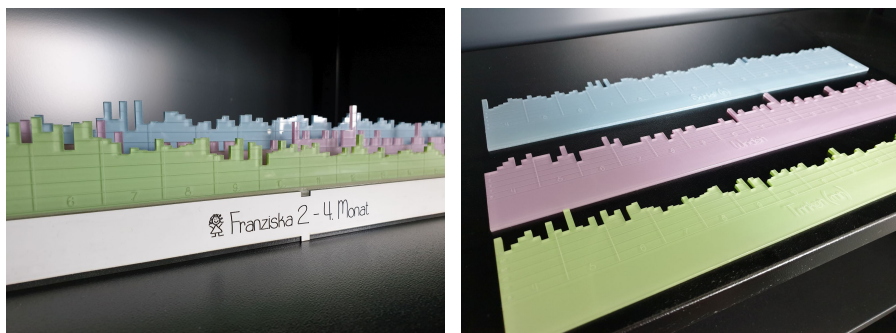


Figure 4.3: A visualization of milk drinking, diapers changed, and sleep times for a 2-4 month old baby built with MakerVis (Swaminathan *et al.*, 2014).

wheeled micro robots mentioned earlier (Le Goc *et al.*, 2019) or dynamic bar charts as created by Taher *et al.* (2015). With a 10×10 interactive bar chart users could annotate, navigate, sort, compare, and organize data values. Dynamic data physicalizations like these are currently very difficult to build, but they hold significant potential.

Data physicalization research is still in a nascent state with many unsolved research questions, including the appropriate evaluation of data physicalizations (Wang *et al.*, 2019) and the design of interactive and dynamic physicalizations.

4.6 Reaching Broader Audiences

In contrast to the 1990s when InfoVis started, today a broad range of people (beyond data-savvy experts) have access to diverse data and data-driven stories. In their recent Visualization Viewpoints article, Lee *et al.* (2020) argued that the visualization research community can and should reach broader audiences beyond data-savvy experts. They highlighted four research topics—personal data visualization, data visualization on mobile devices, inclusive data visualization, and multimodal interaction for data visualization. We assert that post-WIMP interaction can play a central role in achieving this goal.

Lay individuals today can collect and consume their personal data on wearable and mobile devices, where the traditional interaction techniques are not available. These people may have various types of data, including health (e.g., sleep, diet, exercise, mood), productivity, finance, and habit (Choe *et al.*, 2014). Furthermore, these people will exhibit a wide range of visualization literacy: while some might be familiar with more sophisticated charts, such as scatterplots and node-link diagrams, a majority may only have a basic knowledge (if any) of simple charts, such as bar charts, line charts, and pie charts. In addition, they may mainly use data visualization for casual purposes, such as for fun and curiosity, and for simple data access. In such settings, speech can replace visualization operations typically performed through a keyboard and mouse on a desktop/laptop. Besides removing the need of control panels that take up valuable display real estate, voice input can also allow people to simply express their intended goals and questions.

For example, Choe *et al.* (2018) envisioned a scenario in which people can use speech interaction to easily perform data entry (e.g., date, time, date ranges) and to explore their time-oriented data (e.g., self-tracking data), which are often tedious or difficult to do on mobile devices. More generally, multimodal interaction could help us bring visualization to novices. Natural language interaction, especially when combined with direct manipulation, can lower the barriers to access data by making it easy to express data-oriented questions without having to learn native database query languages or specific user interfaces (Setlur *et al.*, 2016). Such interfaces could allow people to express questions about data more naturally, as they would in everyday conversation.

Another group of people who could benefit from novel multimodal interactions, a type of post-WIMP, are people with disabilities. A large number of people—approximately 36 million blind people in the world plus an additional 217 million with moderate to severe vision impairment—are currently excluded from the benefits of data visualization. In addition to blindness or low-vision (BLV), people with intellectual and developmental disabilities (IDDs) also may find it challenging to use existing visualization tools because of impaired cognitive processing (Wu *et al.*, 2019). Unfortunately, the data visualization community has paid little attention to this accessibility problem except for color blindness. This raises a serious equity issue because data visualization is widely used for data-driven reasoning, decision making, and communication both for daily lives and for work: being able to effectively interact with data is becoming a desirable skill. InfoVis and HCI researchers can make visualization tools and systems more accessible by providing additional flexibility leveraging appropriate post-WIMP interaction. We note that improving accessibility will benefit people without disabilities as well. For example, flexible speech-incorporated solutions could allow people to interact with data visualization in a hands-free mode while cooking or running.

As tactile representations are a commonly used form of data graphics for the blind, some researches investigated the effectiveness of such representations either by comparing them against others (e.g., tactile scatterplot vs. tactile table vs. electronic file (Watanabe and Mizukami, 2018)) or by comparing multiple tactile representations (e.g., tactile table

vs. tactile bar chart (Goncu *et al.*, 2010), four tactile representations of network data (Yang *et al.*, 2020)). One approach to support *interactive* exploration or consumption of data by BLV people is, as illustrated by accessible graphing calculators (e.g., GraCALC (Goncu and Marriott, 2015)), to present visualization on a touchscreen device using a mix of audio and haptic feedback where the audio feedback includes both sonification and audio description. Speech input also offers a tremendous opportunity to assist people with motor disabilities, who may find it challenging or impossible to use traditional input devices (e.g., mouse, keyboard, pen). For example, systems like VoiceDraw (Harada *et al.*, 2007) have shown how people with motor impairments can use voice input alone to operate a drawing tool using a combination of spoken commands and sound-based pointer control. Taking inspiration from such systems, we can contemplate conversational interfaces for data exploration and even visualization authoring, thus including additional classes of users, such as designers with disabilities.

In summary, post-WIMP interfaces (perhaps in the form of conversational interfaces) that leverage not only speech but also pen, touch, and gestures are a promising direction to pursue. They will help to expand the boundaries of InfoVis, while improving existing user experiences and introducing new ones.

5

Conclusions

As explorations into post-WIMP interactions for InfoVis, we have provided a personal yet in-depth reflections on a series of projects along with the related work that influenced us. We hope this retrospective will help readers gain an initial understanding of how post-WIMP interaction techniques have evolved in the context of InfoVis, where the field currently stands, and more importantly where it is possible to go in the future. In this section, we conclude with a brief description of some limitations of our personal retrospective.

5.1 Limitations

As stated earlier, our goal with this monograph was to provide a personal retrospective of our own projects and experiences, instead of conducting a comprehensive survey that covers all possible related research on post-WIMP interaction and natural interfaces for InfoVis. As what we cover in this monograph is, therefore, by no means exhaustive or comprehensive, it invites a further, more systematic review incorporating research on related topics. For example, although our discussions on pen- and touch-interaction in Section 2.1 covered different devices (e.g., large displays, tablets), they were centered on a single device setup. However,

researchers have investigated the combination of multiple devices (often with different form factors), including phones and tablets (e.g., Langner *et al.*, 2017), phones and large displays (e.g., Kister *et al.*, 2017; Langner *et al.*, 2018), as well as smartwatches and large displays (e.g., Horak *et al.*, 2018). We discussed interaction with data physicalizations as future research opportunities in Section 4.5. Nonetheless, tangible user interfaces (e.g., Klum *et al.*, 2012; Langner *et al.*, 2014) that can help to manipulate or query data are a related class of systems that can emulate and amplify interactions with data. We also included only a limited number of augmented reality (AR), virtual reality (VR), and mixed reality (MR) applications because much of their interaction relied solely on a WIMP metaphor. Yet, we expect to see more research in the AR, VR, and MR environments, with the focus on innovative interaction with visualization systems. As these are exciting research directions that can help us—InfoVis and HCI researchers—investigate and innovate novel interactions with InfoVis going forward, we hope that future work will further explore these research efforts.

5.2 Concluding Remarks

Over two decades ago, van Dam envisioned post-WIMP user interfaces and called for new thinking about this fourth-generation of UIs (van Dam, 1997). As he conceived, the technologies we need to pursue this vision, such as touch sensors, gesture recognizers, haptic displays, and natural language understanding along with voice recognition, have become more reliable and readily available. HCI researchers have extensively investigated post-WIMP systems that were enabled by (and also promoted) these technological breakthroughs. InfoVis researchers have also recognized the potential of leveraging advancements in these input and interaction technologies in the context of visualization systems. They have been investigating post-WIMP interaction to support more natural, flexible, and engaging user experiences, going beyond a traditional desktop environment. As we discussed in this monograph, research on post-WIMP interaction for InfoVis has made great strides.

On the other hand, we assert that post-WIMP interactions still hold great opportunities for InfoVis. Both the group of people who would

use and need visualizations, and the context in which people engage with visualizations expand. This expansion invites new user experiences and thus requires further research efforts in facilitating and leveraging post-WIMP interaction in InfoVis. To promote further research on post-WIMP InfoVis systems, we sought to share not only achievements and successes (which were already published as papers), but also failures and struggles we had so that others can avoid them. To that end, we reflected on our own projects with which the authors of this article were intimately involved. We discussed the main challenges we have faced as well as insights and lessons we have gained while developing several post-WIMP interactions for InfoVis. We also reflected on how our perspectives and viewpoints on post-WIMP interaction for InfoVis have evolved over the course of different projects. Furthermore, we highlighted several research directions that can facilitate post-WIMP interaction in InfoVis and the opportunities emerging from post-WIMP interfaces.

We expect researchers and industry will make further breakthroughs in input and interaction technologies at a faster pace. The InfoVis and broad HCI communities are fully empowered to make important contributions to post-WIMP InfoVis, expanding the boundaries of InfoVis and enabling new and improved experiences.

Acknowledgements

This monograph would not have been possible without the collaborations and discussions with our collaborators, colleagues, and mentors, as well as the comments and thoughts from them, our study participants, and anonymous reviewers of our paper submissions. We thank the co-authors of research papers, co-organizers of workshops, and team members of our research projects pertaining to post-WIMP interaction for InfoVis: Mallika Agarwal, Meeshu Agnihotri, Caroline Appert, Benjamin Bach, Dominikus Baur, Matthew Brehmer, Jeff Browne, Sheelagh Carpendale, Eun Kyoung Choe, Christopher Collins, Maxime Cordeil, Mary Czerwinski, Raimund Dachselt, Steven Drucker, Tim Dwyer, Darren Edge, Philipp Eichmann, Alex Endert, Nathan Evans, Danyel Fisher, Nathalie Henry Riche, Tobias Hesselmann, Ken Hinckley, Jessica Hullman, Seung-Won Hwang, Kori Inkpen, Tobias Isenberg, Jaemin Jo, Paul Johns, Alark Joshi, Amy Karlson, Rubaiat Habib Kazi, Yea-Seul Kim, Young-Ho Kim, Sehi L'Yi, Ricardo Langner, Kim Marriott, Ron Metoyer, Meredith Ringle Morris, Arpit Narechania, Chris North, Emmanuel Pietriga, George Robertson, Hugo Romat, Ramik Sadana, Bahador Saket, Ayshwarya Saktheeswaran, Jinwook Seo, Vidya Setlur, Tim Sherwood, Greg Smith, Anthony Tang, John Thompson, Melanie Tory, Jagoda Walny, Ulrich von Zadow, and Christopher White.

References

- Adler, M. J. and C. van Doren. 1972. “How to read a book: The classic guide to intelligent reading”.
- Agarwal, M., A. Srinivasan, and J. Stasko. 2019. “VisWall: Visual Data Exploration Using Direct Combination on Large Touch Displays”. In: *2019 IEEE Visualization Conference (VIS)*. 26–30.
- Amazon.com Inc. 2021. “Amazon Alexa”. Product. URL: https://en.wikipedia.org/wiki/Amazon_Alexa.
- Apitz, G. and F. Guimbretière. 2004. “CrossY: a crossing-based drawing application”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 3–12.
- Apple Inc. 2021. “Apple Siri”. Product. URL: <https://en.wikipedia.org/wiki/Siri>.
- Aurisano, J., A. Kumar, A. Alsaïari, B. D. Eugenio, and A. Johnson. 2020. “Many At Once: Capturing Intentions to Create And Use Many Views At Once In Large Display Environments”. *Computer Graphics Forum*. 39(3): 229–240.
- Bach, B., M. Cordeil, T. Dwyer, B. Lee, B. Saket, A. Endert, C. Collins, and S. Carpendale. 2017. “Immersive analytics: Exploring future visualization and interaction technologies for data analytics”. In: *2017 IEEE Visualization Conference (VIS), Accepted Workshop*. Vol. 2.

- Bach, B., M. Cordeil, U. Engelke, B. Ens, M. Serrano, and W. Willett. 2019. "Interaction Design & Prototyping for Immersive Analytics". In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. 1–8.
- Bach, B., R. Dachselt, S. Carpendale, T. Dwyer, C. Collins, and B. Lee. 2016. "Immersive analytics: Exploring future interaction and visualization technologies for data analytics". In: *Proceedings of the ACM Conference on Interactive Surfaces and Spaces*. 529–533.
- Badam, S. K., F. Amini, N. Elmqvist, and P. Irani. 2016. "Supporting visual exploration for multiple users in large display environments". In: *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*. 1–10.
- Ballendat, T., N. Marquardt, and S. Greenberg. 2010. "Proxemic interaction: designing for a proximity and orientation-aware environment". In: *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*. 121–130.
- Baur, D., B. Lee, and S. Carpendale. 2012. "TouchWave: kinetic multi-touch manipulation for hierarchical stacked graphs". In: *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*. 255–264.
- Beaudouin-Lafon, M. 2000. "Instrumental interaction: an interaction model for designing post-WIMP user interfaces". In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 446–453.
- Bolt, R. A. 1980. "Put-that-there" Voice and gesture at the graphics interface". In: *Proceedings of the Conference on Computer Graphics and Interactive Techniques*. 262–270.
- Bostock, M. and J. Heer. 2009. "Protovis: A graphical toolkit for visualization". *IEEE Transactions on Visualization and Computer Graphics*. 15(6): 1121–1128.
- Bostock, M., V. Ogievetsky, and J. Heer. 2011. "D³ data-driven documents". *IEEE Transactions on Visualization and Computer Graphics*. 17(12): 2301–2309.
- Bott, J. N. and J. J. LaViola Jr. 2010. "A pen-based tool for visualizing vector mathematics". In: *Proceedings of the Seventh Sketch-Based Interfaces and Modeling Symposium*. 103–110.

- Brooke, J. 1996. "SUS: a 'quick and dirty' usability". *Usability evaluation in industry*. 189(104): 1–7.
- Browne, J., B. Lee, S. Carpendale, N. Riche, and T. Sherwood. 2011. "Data analysis on interactive whiteboards through sketch-based interaction". In: *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*. 154–157.
- Card, S. K., J. Mackinlay, and B. Shneiderman. 1999. "Readings in information visualization: using vision to think. 1999". *San Francisco: Morgan Kaufman*.
- Carpendale, S. 2008. "Evaluating Information Visualizations". In: *Information Visualization: Human-Centered Issues and Perspectives*. Vol. 4950. *LNCS*. Berlin: Springer. 19–45.
- Chen, N., F. Guimbretiere, and A. Sellen. 2012. "Designing a multi-slate reading environment to support active reading activities". *ACM Transactions on Computer-Human Interaction*. 19(3): 1–35.
- Choe, E. K., B. Lee, and S.-W. Hwang. 2018. "Personal Data Exploration with Speech on Mobile Devices". In: *AVI 2018 Workshop on Multimodal Interaction for Data Visualization*.
- Choe, E. K., N. B. Lee, B. Lee, W. Pratt, and J. A. Kientz. 2014. "Understanding quantified-selfers' practices in collecting and exploring personal data". In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1143–1152.
- Choe, E., R. Dachselt, P. Isenberg, and B. Lee. 2019. "Mobile data visualization (dagstuhl seminar 19292)". *Dagstuhl Reports*.
- Cohen, P. R., M. Dalrymple, D. B. Moran, F. Pereira, and J. W. Sullivan. 1989. "Synergistic use of direct manipulation and natural language". In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 227–233.
- Cohen, P. R., M. Johnston, D. McGee, S. Oviatt, J. Pittman, I. Smith, L. Chen, and J. Clow. 1997. "QuickSet: Multimodal interaction for distributed applications". In: *Proceedings of the ACM Conference on Multimedia*. 31–40.
- Cohen, P. R. and D. R. McGee. 2004. "Tangible multimodal interfaces for safety-critical applications". *Communications of the ACM*. 47(1): 41–46.

- Cordeil, M., B. Bach, A. Cunningham, B. Montoya, R. T. Smith, B. H. Thomas, and T. Dwyer. 2020. “Embodied Axes: Tangible, Actuated Interaction for 3D Augmented Reality Data Spaces”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1–12.
- Cordeil, M., A. Cunningham, T. Dwyer, B. H. Thomas, and K. Marriott. 2017. “ImAxes: Immersive axes as embodied affordances for interactive multivariate data visualisation”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 71–83.
- Cox, K., R. E. Grinter, S. L. Hibino, L. J. Jagadeesan, and D. Mantilla. 2001. “A multi-modal natural language interface to an information visualization environment”. *International Journal of Speech Technology*. 4(3-4): 297–314.
- Cui, W., X. Zhang, Y. Wang, H. Huang, B. Chen, L. Fang, H. Zhang, J.-G. Lou, and D. Zhang. 2019. “Text-to-viz: Automatic generation of infographics from proportion-related natural language statements”. *IEEE Transactions on Visualization and Computer Graphics*. 26(1): 906–916.
- Dahlbäck, N., A. Jönsson, and L. Ahrenberg. 1993. “Wizard of Oz studies: why and how”. In: *Proceedings of the ACM Conference on Intelligent User Interfaces*. 193–200.
- Dietz, P. and D. Leigh. 2001. “DiamondTouch: a multi-user touch technology”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 219–226.
- Dragicevic, P. and Y. Jansen. 2012. “List of Physical Visualizations”. URL: <http://dataphys.org/list>.
- Drogemuller, A., A. Cunningham, J. Walsh, M. Cordeil, W. Ross, and B. Thomas. 2018. “Evaluating navigation techniques for 3d graph visualizations in virtual reality”. In: *2018 International Symposium on Big Data Visual and Immersive Analytics (BDVA)*. IEEE. 1–10.
- Drucker, S. M., D. Fisher, R. Sadana, J. Herron, and M. Schraefel. 2013. “TouchViz: a case study comparing two interfaces for data analytics on tablets”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 2301–2310.

- Dunne, C., N. Henry Riche, B. Lee, R. Metoyer, and G. Robertson. 2012. “GraphTrail: Analyzing large multivariate, heterogeneous networks while supporting exploration history”. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 1663–1672.
- Dwyer, T., B. Lee, D. Fisher, K. I. Quinn, P. Isenberg, G. Robertson, and C. North. 2009. “A comparison of user-generated and automatic graph layouts”. *IEEE Transactions on Visualization and Computer Graphics*. 15(6): 961–968.
- Eichmann, P., D. Edge, N. Evans, B. Lee, M. Brehmer, and C. White. 2020. “Orchard: Exploring Multivariate Heterogeneous Networks on Mobile Phones”. *Computer Graphics Forum*. 39(3): 115–126.
- Ens, B., B. Bach, M. Cordeil, U. Engelke, M. Serrano, and W. Willett. 2020. “Envisioning Future Productivity for Immersive Analytics”. In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. 1–7.
- Fast, E., B. Chen, J. Mendelsohn, J. Bassen, and M. S. Bernstein. 2018. “Iris: A conversational agent for complex tasks”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 473.
- Forlines, C. and C. Shen. 2005. “DTLens: multi-user tabletop spatial data exploration”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 119–122.
- Frisch, M., J. Heydekorn, and R. Dachsel. 2009. “Investigating multi-touch and pen gestures for diagram editing on interactive surfaces”. In: *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*. 149–156.
- Gao, T., M. Dontcheva, E. Adar, Z. Liu, and K. G. Karahalios. 2015. “DataTone: Managing Ambiguity in Natural Language Interfaces for Data Visualization”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 489–500.
- Goncu, C. and K. Marriott. 2015. “GraCALC: An accessible graphing calculator”. In: *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility*. 311–312.
- Goncu, C., K. Marriott, and J. Hurst. 2010. “Usability of accessible bar charts”. In: *International Conference on Theory and Application of Diagrams*. Springer. 167–181.

- Gourdol, A. P., L. Nigay, D. Salber, and J. Coutaz. 1992. “Two Case Studies of Software Architecture for Multimodal Interactive Systems: VoicePaint and a Voice-enabled Graphical Notebook.” *Engineering for Human-Computer Interaction*. 92: 271–84.
- Greenberg, S. and B. Buxton. 2008. “Usability evaluation considered harmful (some of the time)”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 111–120.
- Greenberg, S., N. Marquardt, T. Ballendat, R. Diaz-Marino, and M. Wang. 2011. “Proxemic interactions: the new ubicomp?” *Interactions*. 18(1): 42–50.
- Han, J. Y. 2005. “Low-cost multi-touch sensing through frustrated total internal reflection”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 115–118.
- Hanrahan, P. 2012. “Analytic database technologies for a new kind of user: the data enthusiast”. In: *Proceedings of the ACM SIGMOD Conference on Management of Data*. 577–578.
- Harada, S., J. O. Wobbrock, and J. A. Landay. 2007. “VoiceDraw: a hands-free voice-driven drawing application for people with motor impairments”. In: *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility*. 27–34.
- Hauptmann, A. G. 1989. “Speech and gestures for graphic image manipulation”. *ACM SIGCHI Bulletin*. 20(SI): 241–245.
- Heer, J., S. K. Card, and J. A. Landay. 2005. “Prefuse: a toolkit for interactive information visualization”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 421–430.
- Hinckley, K., P. Baudisch, G. Ramos, and F. Guimbretiere. 2005. “Design and analysis of delimiters for selection-action pen gesture phrases in scriboli”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 451–460.
- Hinckley, K., S. Heo, M. Pahud, C. Holz, H. Benko, A. Sellen, R. Banks, K. O’Hara, G. Smyth, and W. Buxton. 2016. “Pre-touch sensing for mobile interaction”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 2869–2881.
- Hinckley, K., J. Pierce, M. Sinclair, and E. Horvitz. 2000. “Sensing techniques for mobile interaction”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 91–100.

- Hinckley, K., K. Yatani, M. Pahud, N. Coddington, J. Rodenhouse, A. Wilson, H. Benko, and B. Buxton. 2010. “Pen + touch= new tools”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 27–36.
- Hoffswell, J., W. Li, and Z. Liu. 2020. “Techniques for Flexible Responsive Visualization Design”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1–13.
- Hoffswell, J. and Z. Liu. 2019. “Interactive Repair of Tables Extracted from PDF Documents on Mobile Devices”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1–13.
- Holland, S. and D. Oppenheim. 1999. “Direct combination”. In: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 262–269.
- Hoque, E., V. Setlur, M. Tory, and I. Dykeman. 2018. “Applying pragmatics principles for interaction with visual analytics”. *IEEE Transactions on Visualization and Computer Graphics*. 24(1): 309–318.
- Horak, T., S. K. Badam, N. Elmqvist, and R. Dachsel. 2018. “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*. 1–13.
- Huron, S., S. Carpendale, A. Thudt, A. Tang, and M. Mauerer. 2014a. “Constructive visualization”. In: *Proceedings of the ACM Conference on Designing Interactive Systems*. 433–442.
- Huron, S., Y. Jansen, and S. Carpendale. 2014b. “Constructing Visual Representations: Investigating the Use of Tangible Tokens”. *IEEE Transactions on Visualization and Computer Graphics*. 20(12): 2102–2111.
- Hurter, C., N. H. Riche, S. M. Drucker, M. Cordeil, R. Alligier, and R. Vuillemot. 2018. “Fiberclay: Sculpting three dimensional trajectories to reveal structural insights”. *IEEE Transactions on Visualization and Computer Graphics*. 25(1): 704–714.
- IBM. 2018. “IBM Watson Analytics”. Product. URL: <http://www.ibm.com/analytics/watson-analytics>.

- Isenberg, P., N. Elmqvist, J. Scholtz, D. Cernea, K.-L. Ma, and H. Hagen. 2011. “Collaborative Visualization: Definition, Challenges, and Research Agenda”. *Information Visualization Journal (IVS), Special Issue on Information Visualization: State of the Field and New Research Directions*. 10(4): 310–326.
- Isenberg, P. and D. Fisher. 2009. “Collaborative brushing and linking for co-located visual analytics of document collections”. *Computer Graphics Forum*. 28(3): 1031–1038.
- Isenberg, P., D. Fisher, M. R. Morris, K. Inkpen, and M. Czerwinski. 2010. “An exploratory study of co-located collaborative visual analytics around a tabletop display”. In: *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*. 179–186.
- Isenberg, P., D. Fisher, S. A. Paul, M. R. Morris, K. Inkpen, and M. Czerwinski. 2012. “Co-located collaborative visual analytics around a tabletop display”. *IEEE Transactions on Visualization and Computer Graphics*. 18(5): 689–702.
- Isenberg, P., T. Isenberg, T. Hesselmann, B. Lee, U. Von Zadow, and A. Tang. 2013a. “Data visualization on interactive surfaces: A research agenda”. *IEEE Computer Graphics and Applications*. 33(2): 16–24.
- Isenberg, P., B. Lee, A. Joshi, and T. Isenberg. 2015. “ITS Workshop DEXIS 2015: Visual Data Exploration on Interactive Surfaces”. In: *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*. 491–494.
- Isenberg, T., P. Isenberg, J. Chen, M. Sedlmair, and T. Möller. 2013b. “A Systematic Review on the Practice of Evaluating Visualization”. *IEEE Transactions on Visualization and Computer Graphics*. 19(12): 2818–2827.
- Jakobsen, M. R., Y. S. Haile, S. Knudsen, and K. Hornbæk. 2013. “Information visualization and proxemics: Design opportunities and empirical findings”. *IEEE Transactions on Visualization and Computer Graphics*. 19(12): 2386–2395.
- Jansen, Y., P. Dragicevic, P. Isenberg, J. Alexander, A. Karnik, J. Kildal, S. Subramanian, and K. Hornbæk. 2015. “Opportunities and Challenges for Data Physicalization”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 3227–3236.

- Jo, J., S. L'Yi, B. Lee, and J. Seo. 2017. "TouchPivot: blending WIMP & post-WIMP interfaces for data exploration on tablet devices". In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 2660–2671.
- Jo, J., B. Lee, and J. Seo. 2015. "WordlePlus: expanding wordle's use through natural interaction and animation". *IEEE Computer Graphics and Applications*. 35(6): 20–28.
- John, R. J. L., N. Potti, and J. M. Patel. 2017. "Ava: From Data to Insights Through Conversations". In: *CIDR*.
- Ju, W., B. A. Lee, and S. R. Klemmer. 2008. "Range: exploring implicit interaction through electronic whiteboard design". In: *Proceedings of the ACM Conference on Computer Supported Cooperative Work*. 17–26.
- Kandel, S., J. Heer, C. Plaisant, J. Kennedy, F. Van Ham, N. H. Riche, C. Weaver, B. Lee, D. Brodbeck, and P. Buono. 2011. "Research directions in data wrangling: Visualizations and transformations for usable and credible data". *Information Visualization*. 10(4): 271–288.
- Kandel, S., A. Paepcke, J. M. Hellerstein, and J. Heer. 2012. "Enterprise data analysis and visualization: An interview study". *IEEE Transactions on Visualization and Computer Graphics*. 18(12): 2917–2926.
- Kassel, J.-F. and M. Rohs. 2018. "Valletto: A multimodal interface for ubiquitous visual analytics". In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. 1–6.
- Kim, D. H., E. Hoque, and M. Agrawala. 2020. "Answering Questions about Charts and Generating Visual Explanations". In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1–13.
- Kim, N. W., N. Henry Riche, B. Bach, G. Xu, M. Brehmer, K. Hinckley, M. Pahud, H. Xia, M. J. McGuffin, and H. Pfister. 2019a. "DataToon: Drawing dynamic network comics with pen+ touch interaction". In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1–12.

- Kim, N. W., H. Im, N. Henry Riche, A. Wang, K. Gajos, and H. Pfister. 2019b. “DataSelfie: Empowering People to Design Personalized Visuals to Represent Their Data”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1–12.
- Kim, Y.-S., N. Henry Riche, B. Lee, M. Brehmer, M. Pahud, K. Hinckley, and J. Hullman. 2019c. “Inking your insights: Investigating digital externalization behaviors during data analysis”. In: *Proceedings of the ACM Conference on Interactive Surfaces and Spaces*. 255–267.
- Kim, Y.-H., B. Lee, A. Srinivasan, and E. K. Choe. 2021. “Data@Hand: Fostering Visual Exploration of Personal Data on Smartphones Leveraging Speech and Touch Interaction”. In: *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. In press. ACM.
- Kister, U., K. Klamka, C. Tominski, and R. Dachsel. 2017. “GraSp: Combining Spatially-aware Mobile Devices and a Display Wall for Graph Visualization and Interaction”. In: *Computer Graphics Forum*. Vol. 36. No. 3. Wiley Online Library. 503–514.
- Klemmer, S. R., M. W. Newman, R. Farrell, M. Bilezikjian, and J. A. Landay. 2001. “The designers’ outpost: a tangible interface for collaborative web site”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 1–10.
- Klum, S., P. Isenberg, R. Langner, J.-D. Fekete, and R. Dachsel. 2012. “Stackables: combining tangibles for faceted browsing”. In: *Proceedings of the International Working Conference on Advanced Visual Interfaces*. 241–248.
- Kondo, B. and C. Collins. 2014. “DimpVis: Exploring time-varying information visualizations by direct manipulation”. *IEEE Transactions on Visualization and Computer Graphics*. 20(12): 2003–2012.
- Kumar, A., J. Aurisano, B. Di Eugenio, A. Johnson, A. Gonzalez, and J. Leigh. 2016. “Towards a Dialogue System that Supports Rich Visualizations of Data”. In: *Proceedings of the Annual Meeting of the Special Interest Group on Discourse and Dialogue*. 304.
- Kumar, A., T. Paek, and B. Lee. 2012. “Voice typing: a new speech interaction model for dictation on touchscreen devices”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 2277–2286.

- Lai, C., Z. Lin, R. Jiang, Y. Han, C. Liu, and X. Yuan. 2020. “Automatic Annotation Synchronizing with Textual Description for Visualization”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1–13.
- Lam, H., E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. 2012. “Empirical Studies in Information Visualization: Seven Scenarios”. *IEEE Transactions on Visualization and Computer Graphics*. 18(9): 1520–1536.
- Langner, R., A. Augsborg, and R. Dachsel. 2014. “Cubequery: Tangible interface for creating and manipulating database queries”. In: *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*. 423–426.
- Langner, R., T. Horak, and R. Dachsel. 2017. “VisTiles: Coordinating and Combining Co-located Mobile Devices for Visual Data Exploration”. *IEEE transactions on visualization and computer graphics*. 24(1): 626–636.
- Langner, R., U. Kister, and R. Dachsel. 2018. “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.
- Laput, G. 2019. “Context-Driven Implicit Interactions”. *Ph.D. Thesis*.
- Laput, G. P., M. Dontcheva, G. Wilensky, W. Chang, A. Agarwala, J. Linder, and E. Adar. 2013. “Pixeltone: A multimodal interface for image editing”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 2185–2194.
- LaViola Jr, J. J. and R. C. Zeleznik. 2006. “Mathpad2: a system for the creation and exploration of mathematical sketches”. In: *ACM SIGGRAPH 2006 Courses*. 33–es.
- Le Goc, M., L. H. Kim, A. Parsaei, J.-D. Fekete, P. Dragicevic, and S. Follmer. 2016. “Zooids: Building blocks for swarm user interfaces”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 97–109.
- Le Goc, M., C. Perin, S. Follmer, J.-D. Fekete, and P. Dragicevic. 2019. “Dynamic Composite Data Physicalization Using Wheeled Micro-Robots”. *IEEE Transactions on Visualization and Computer Graphics*. 25(1): 737–747.

- Ledo, D., J. Vermeulen, S. Carpendale, S. Greenberg, L. Oehlberg, and S. Boring. 2019. “Astral: Prototyping Mobile and Smart Object Interactive Behaviours Using Familiar Applications”. In: *Proceedings of the ACM Conference on Designing Interactive Systems*. 711–724.
- Lee, B., M. Brehmer, P. Isenberg, E. K. Choe, R. Langner, and R. Dachsel. 2018a. “Data visualization on mobile devices”. In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. 1–8.
- Lee, B., E. K. Choe, P. Isenberg, K. Marriott, and J. Stasko. 2020. “Reaching Broader Audiences With Data Visualization”. *IEEE Computer Graphics and Applications*. 40(2): 82–90.
- Lee, B., K. Isaacs, D. A. Szafir, G. E. Marai, C. Turkay, M. Tory, S. Carpendale, and A. Endert. 2019. “Broadening intellectual diversity in visualization research papers”. *IEEE Computer Graphics and Applications*. 39(4): 78–85.
- Lee, B., P. Isenberg, N. H. Riche, and S. Carpendale. 2012. “Beyond mouse and keyboard: Expanding design considerations for information visualization interactions”. *IEEE Transactions on Visualization and Computer Graphics*. 18(12): 2689–2698.
- Lee, B., R. H. Kazi, and G. Smith. 2013. “SketchStory: Telling more engaging stories with data through freeform sketching”. *IEEE Transactions on Visualization and Computer Graphics*. 19(12): 2416–2425.
- Lee, B., G. Smith, N. H. Riche, A. Karlson, and S. Carpendale. 2015. “SketchInsight: Natural data exploration on interactive whiteboards leveraging pen and touch interaction”. In: *2015 IEEE Pacific Visualization Symposium (PacificVis)*. 199–206.
- Lee, B., A. Srinivasan, J. Stasko, M. Tory, and V. Setlur. 2018b. “Multimodal interaction for data visualization”. In: *Proceedings of the International Conference on Advanced Visual Interfaces*. 1–3.
- Liu, Z., J. Thompson, A. Wilson, M. Dontcheva, J. Delorey, S. Grigg, B. Kerr, and J. Stasko. 2018. “Data Illustrator: Augmenting Vector Design Tools with Lazy Data Binding for Expressive Visualization Authoring”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 123:1–123:13.
- Loper, E. and S. Bird. 2002. “NLTK: the natural language toolkit”. *arXiv preprint cs/0205028*.

- Lundgard, A., C. Lee, and A. Satyanarayan. 2019. “Sociotechnical Considerations for Accessible Visualization Design”. In: *2019 IEEE Visualization Conference (VIS)*. 16–20.
- Martin, J.-C. 1998. “TYCOON: Theoretical framework and software tools for multimodal interfaces”.
- Méndez, G. G., M. A. Nacenta, and S. Vandenheste. 2016. “iVoLVER: Interactive visual language for visualization extraction and reconstruction”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 4073–4085.
- Metoyer, R., B. Lee, N. Henry Riche, and M. Czerwinski. 2012. “Understanding the verbal language and structure of end-user descriptions of data visualizations”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1659–1662.
- Microsoft. 2021. “Microsoft Power BI”. Product. URL: https://en.wikipedia.org/wiki/Microsoft_Power_BI.
- Myers, C., A. Furqan, J. Nebolsky, K. Caro, and J. Zhu. 2018. “Patterns for how users overcome obstacles in voice user interfaces”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1–7.
- Mynatt, E. D., T. Igarashi, W. K. Edwards, and A. LaMarca. 1999. “Flatland: new dimensions in office whiteboards”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 346–353.
- Narechania, A., A. Srinivasan, and J. Stasko. 2021. “NL4DV: A Toolkit for Generating Analytic Specifications for Data Visualization from Natural Language Queries”. *IEEE Transactions on Visualization and Computer Graphics*. In press.
- North, C., T. Dwyer, B. Lee, D. Fisher, P. Isenberg, G. Robertson, and K. Inkpen. 2009. “Understanding multi-touch manipulation for surface computing”. In: *IFIP Conference on Human-Computer Interaction*. Springer. 236–249.
- Okoe, M., S. S. Alam, and R. Jianu. 2014. “A gaze-enabled graph visualization to improve graph reading tasks”. *Computer Graphics Forum*. 33(3): 251–260.

- Olsen Jr, D. R. 2007. “Evaluating user interface systems research”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 251–258.
- Olsen, L., F. F. Samavati, M. C. Sousa, and J. A. Jorge. 2009. “Sketch-based modeling: A survey”. *Computers & Graphics*. 33(1): 85–103.
- Oviatt, S. 1999. “Ten myths of multimodal interaction”. *Communications of the ACM*. 42(11): 74–81.
- Parnami, A., A. Gupta, G. Reyes, R. Sadana, Y. Li, and G. D. Abowd. 2016. “Mogeste: A Mobile Tool for In-Situ Motion Gesture Design”. In: *Proceedings of the 8th Indian Conference on Human Computer Interaction*. 35–43.
- Pausch, R. and J. H. Leatherby. 1991. “An empirical study: Adding voice input to a graphical editor”. *Journal of the American Voice Input/Output Society*. 9: 2–55.
- Pfeuffer, K., J. Alexander, M. K. Chong, Y. Zhang, and H. Gellersen. 2015. “Gaze-shifting: Direct-indirect input with pen and touch modulated by gaze”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 373–383.
- Pfeuffer, K. and H. Gellersen. 2016. “Gaze and touch interaction on tablets”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 301–311.
- Pfeuffer, K., K. Hinckley, M. Pahud, and B. Buxton. 2017a. “Thumb + Pen Interaction on Tablets”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 3254–3266.
- Pfeuffer, K., B. Mayer, D. Mardanbegi, and H. Gellersen. 2017b. “Gaze + pinch interaction in virtual reality”. In: *Proceedings of the ACM Symposium on Spatial User Interaction*. 99–108.
- Plaisant, C. 2004. “The challenge of information visualization evaluation”. In: *Proceedings of the International Conference on Advanced Visual Interfaces*. 109–116.
- Pretorius, J., H. C. Purchase, and J. T. Stasko. 2014. “Tasks for multivariate network analysis”. In: *Multivariate Network Visualization*. Springer. 77–95.

- Reipschläger, P., T. Flemisch, and R. Dachsel. 2021. “Personal Augmented Reality for Information Visualization on Large Interactive Displays”. *IEEE Transactions on Visualization and Computer Graphics*. 27(2): 1182–1192.
- Ren, D., T. Höllerer, and X. Yuan. 2014. “iVisDesigner: Expressive interactive design of information visualizations”. *IEEE Transactions on Visualization and Computer Graphics*. 20(12): 2092–2101.
- Ren, D., B. Lee, and M. Brehmer. 2019. “Charticulator: Interactive Construction of Bespoke Chart Layouts”. *IEEE Transactions on Visualization and Computer Graphics*. 25(1): 789–799.
- Ren, D., B. Lee, M. Brehmer, and N. H. Riche. 2018. “Reflecting on the evaluation of visualization authoring systems: Position paper”. In: *2018 IEEE Evaluation and Beyond-Methodological Approaches for Visualization (BELIV)*. 86–92.
- Ren, D., B. Lee, and T. Höllerer. 2017. “Stardust: Accessible and transparent gpu support for information visualization rendering”. *Computer Graphics Forum*. 36(3): 179–188.
- Riche, N. H., B. Lee, and C. Plaisant. 2010. “Understanding interactive legends: a comparative evaluation with standard widgets”. *Computer Graphics Forum*. 29(3): 1193–1202.
- Roberts, J. C., P. D. Ritsos, S. K. Badam, D. Brodbeck, J. Kennedy, and N. Elmqvist. 2014. “Visualization beyond the desktop—the next big thing”. *IEEE Computer Graphics and Applications*. 34(6): 26–34.
- Romat, H., N. Henry Riche, K. Hinckley, B. Lee, C. Appert, E. Pietriga, and C. Collins. 2019a. “ActiveInk: (Th) Inking with Data”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1–13.
- Romat, H., E. Pietriga, N. Henry-Riche, K. Hinckley, and C. Appert. 2019b. “SpaceInk: Making Space for In-Context Annotations”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 871–882.
- Ryall, K., N. Lesh, T. Lanning, D. Leigh, H. Miyashita, and S. Makino. 2005. “Querylines: approximate query for visual browsing”. In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. 1765–1768.

- Rzeszotarski, J. M. and A. Kittur. 2014. “Kinetic: naturalistic multi-touch data visualization”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 897–906.
- Sadana, R., M. Agnihotri, and J. Stasko. 2018. “Touching data: A discoverability-based evaluation of a visualization interface for tablet computers”. *arXiv preprint arXiv:1806.06084*.
- Sadana, R. and J. Stasko. 2014. “Designing and implementing an interactive scatterplot visualization for a tablet computer”. In: *Proceedings of the International Conference on Advanced Visual Interfaces*. 265–272.
- Sadana, R. and J. Stasko. 2016a. “Designing multiple coordinated visualizations for tablets”. *Computer Graphics Forum*. 35(3): 261–270.
- Sadana, R. and J. Stasko. 2016b. “Expanding selection for information visualization systems on tablet devices”. In: *Proceedings of the ACM Conference on Interactive Surfaces and Spaces*. 149–158.
- Saket, B., A. Endert, and J. Stasko. 2016. “Beyond usability and performance: A review of user experience-focused evaluations in visualization”. In: *Proceedings of the Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization (BELIV)*. 133–142.
- Saktheeswaran, A., A. Srinivasan, and J. Stasko. 2020. “Touch? Speech? or Touch and Speech? Investigating Multimodal Interaction for Visual Network Exploration and Analysis”. *IEEE Transactions on Visualization and Computer Graphics*. 26(6): 2168–2179.
- Satyanarayan, A. and J. Heer. 2014. “Lyra: An interactive visualization design environment”. *Computer Graphics Forum*. 33(3): 351–360.
- Satyanarayan, A., B. Lee, D. Ren, J. Heer, J. Stasko, J. Thompson, M. Brehmer, and Z. Liu. 2019. “Critical reflections on visualization authoring systems”. *IEEE Transactions on Visualization and Computer Graphics*. 26(1): 461–471.
- Satyanarayan, A., D. Moritz, K. Wongsuphasawat, and J. Heer. 2016. “Vega-lite: A grammar of interactive graphics”. *IEEE Transactions on Visualization and Computer Graphics*. 23(1): 341–350.

- Satyanarayan, A., R. Russell, J. Hoffswell, and J. Heer. 2015. “Reactive vega: A streaming dataflow architecture for declarative interactive visualization”. *IEEE Transactions on Visualization and Computer Graphics*. 22(1): 659–668.
- Schmidt, S., M. A. Nacenta, R. Dachsel, and S. Carpendale. 2010. “A set of multi-touch graph interaction techniques”. In: *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*. 113–116.
- Schmitz, A., C. Thoburn, and J. Tangelder. 2020. “Hammer.js: JavaScript library for touch gesture recognition”. Javascript Library. URL: <https://hammerjs.github.io/>.
- Schwab, M., S. Hao, O. Vitek, J. Tompkin, J. Huang, and M. A. Borkin. 2019a. “Evaluating pan and zoom timelines and sliders”. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.
- Schwab, M., J. Tompkin, J. Huang, and M. A. Borkin. 2019b. “EasyPZ.js: Interaction Binding for Pan and Zoom Visualizations”. In: *2019 IEEE Visualization Conference (VIS)*. 31–35.
- Setlur, V., S. E. Battersby, M. Tory, R. Gossweiler, and A. X. Chang. 2016. “Eviza: A Natural Language Interface for Visual Analysis”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 365–377.
- Setlur, V., M. Tory, and A. Djalali. 2019. “Inferencing Underspecified Natural Language Utterances in Visual Analysis”. In: *Proceedings of the ACM Conference on Intelligent User Interfaces*. 40–51.
- Sharma, R., M. Yeasin, N. Krahnstoeber, I. Rauschert, G. Cai, I. Brewer, A. M. MacEachren, and K. Sengupta. 2003. “Speech-gesture driven multimodal interfaces for crisis management”. *Proceedings of the IEEE*. 91(9): 1327–1354.
- Shneiderman, B. 1983. “Direct Manipulation: A Step Beyond Programming Languages”. *Computer*. 16(8): 57–69.
- Shneiderman, B. 1994. “Dynamic queries for visual information seeking”. *IEEE Software*. 11(6): 70–77.
- Shneiderman, B. 2016. *The New ABCs of Research, Achieving Breakthrough Collaborations*. Oxford University Press.

- Sicat, R., J. Li, J. Choi, M. Cordeil, W.-K. Jeong, B. Bach, and H. Pfister. 2018. “DXR: A toolkit for building immersive data visualizations”. *IEEE Transactions on Visualization and Computer Graphics*. 25(1): 715–725.
- Spindler, M., C. Tominski, H. Schumann, and R. Dachsel. 2010. “Tangible views for information visualization”. In: *ACM International Conference on Interactive Tabletops and Surfaces*. 157–166.
- Srinivasan, A., B. Lee, N. Riche, S. Drucker, and K. Hinckley. 2020a. “InChorus: Designing Consistent Multimodal Interactions for Data Visualization on Tablet Devices”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 653:1–653:13.
- Srinivasan, A., B. Lee, and J. Stasko. 2018. “Facilitating spreadsheet manipulation on mobile devices leveraging speech”. In: *CHI 2018 Workshop on Data Visualization on Mobile Devices (MobileVis)*.
- Srinivasan, A., B. Lee, and J. T. Stasko. 2020b. “Interweaving Multimodal Interaction with Flexible Unit Visualizations for Data Exploration”. *IEEE Transactions on Visualization and Computer Graphics*.
- Srinivasan, A. and J. Stasko. 2018. “Orko: Facilitating multimodal interaction for visual exploration and analysis of networks”. *IEEE Transactions on Visualization and Computer Graphics*. 24(1): 511–521.
- Stasko, J. 2014. “Value-Driven Evaluation of Visualizations”. In: *Proceedings of Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization (BELIV)*. 46–53.
- Streit, M., A. Lex, H. Müller, and D. Schmalstieg. 2009. “Gaze-based focus adaption in an information visualization system”. In: *IADIS International Conference Computer Graphics, Visualization, Computer Vision and Image Processing*. 303–307.
- Subramonyam, H. and E. Adar. 2018. “SmartCues: A Multitouch Query Approach for Details-on-Demand through Dynamically Computed Overlays”. *IEEE Transactions on Visualization and Computer Graphics*. 25(1): 597–607.
- Subramonyam, H., W. Li, E. Adar, and M. Dontcheva. 2018. “TakeToons: Script-driven Performance Animation”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 663–674.

- Sun, Y., J. Leigh, A. Johnson, and S. Lee. 2010. "Articulate: A semi-automated model for translating Natural Language queries into meaningful visualizations". In: *International Symposium on Smart Graphics*. Springer. 184–195.
- Swaminathan, S., C. Shi, Y. Jansen, P. Dragicevic, L. Oehlberg, and J.-D. Fekete. 2014. "Supporting the Design and Fabrication of Physical Visualizations". In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 3845–3854.
- Tableau Software. 2021. "Tableau Ask Data". Product. URL: <https://www.tableau.com/products/new-features/ask-data>.
- Taher, F., J. Hardy, A. Karnik, C. Weichel, Y. Jansen, K. Hornbæk, and J. Alexander. 2015. "Exploring Interactions with Physically Dynamic Bar Charts". In: New York, NY, USA: ACM. 3237–3246.
- Thompson, J., A. Srinivasan, and J. Stasko. 2018. "Tangraphe: interactive exploration of network visualizations using single hand, multi-touch gestures". In: *Proceedings of the International Conference on Advanced Visual Interfaces*. 1–5.
- Thudt, A., B. Lee, E. K. Choe, and S. Carpendale. 2017. "Expanding research methods for a realistic understanding of personal visualization". *IEEE Computer Graphics and Applications*. 37(2): 12–18.
- Tory, M. and V. Setlur. 2019. "Do what I mean, not what I say! Design considerations for supporting intent and context in analytical conversation". In: *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*. 93–103.
- Tsandilas, T., A. Bezerianos, and T. Jacob. 2015. "SketchSliders: Sketching widgets for visual exploration on wall displays". In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 3255–3264.
- van Dam, A. 1997. "Post-WIMP user interfaces". *Communications of the ACM*. 40(2): 63–67.
- Voida, S., M. Tobiasz, J. Stromer, P. Isenberg, and S. Carpendale. 2009. "Getting practical with interactive tabletop displays: designing for dense data, "fat fingers," diverse interactions, and face-to-face collaboration". In: *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*. 109–116.

- Walny, J., S. Huron, C. Perin, T. Wun, R. Pusch, and S. Carpendale. 2017. “Active reading of visualizations”. *IEEE Transactions on Visualization and Computer Graphics*. 24(1): 770–780.
- Walny, J., B. Lee, P. Johns, N. H. Riche, and S. Carpendale. 2012. “Understanding pen and touch interaction for data exploration on interactive whiteboards”. *IEEE Transactions on Visualization and Computer Graphics*. 18(12): 2779–2788.
- Wang, Y., A. Segal, R. Klatzky, D. F. Keefe, P. Isenberg, J. Hurtienne, E. Hornecker, T. Dwyer, and S. Barrass. 2019. “An Emotional Response to the Value of Visualization”. *IEEE Computer Graphics and Applications*. 39(5): 8–17.
- Watanabe, T. and H. Mizukami. 2018. “Effectiveness of tactile scatter plots: comparison of non-visual data representations”. In: *International Conference on Computers Helping People with Special Needs*. Springer. 628–635.
- Wattenberg, M. 2001. “Sketching a graph to query a time-series database”. In: *Extended Abstracts of the ACM Conference on Human factors in Computing Systems*. 381–382.
- “Web Speech API”. 2019. Online Draft Community Group Report. URL: <https://wicg.github.io/speech-api/>.
- Weiser, M. 1999. “The computer for the 21st century”. *ACM SIGMOBILE Mobile Computing and Communications Review*. 3(3): 3–11.
- Weizenbaum, J. 1966. “ELIZA—a computer program for the study of natural language communication between man and machine”. *Communications of the ACM*. 9(1): 36–45.
- Wigdor, D. and D. Wixon. 2011. *Brave NUI world: designing natural user interfaces for touch and gesture*. Morgan Kaufmann.
- Winograd, T. 1972. “Understanding natural language”. *Cognitive psychology*. 3(1): 1–191.
- Wobbrock, J. O., A. D. Wilson, and Y. Li. 2007. “Gestures without libraries, toolkits or training: a \$1 recognizer for user interface prototypes”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 159–168.

- Wong, N. and S. Carpendale. 2007. “Supporting interactive graph exploration using edge plucking”. In: *Visualization and Data Analysis 2007*. Vol. 6495. International Society for Optics and Photonics. 649508.
- Wong, N., S. Carpendale, and S. Greenberg. 2003. “Edgelens: An interactive method for managing edge congestion in graphs”. In: *IEEE Symposium on Information Visualization 2003*. 51–58.
- Wu, A., W. Tong, T. Dwyer, B. Lee, P. Isenberg, and H. Qu. 2021. “MobileVisFixer: Tailoring Web Visualizations for Mobile Phones Leveraging an Explainable Reinforcement Learning Framework”. *IEEE Transactions on Visualization and Computer Graphics*. In press.
- Wu, K., S. Tanis, and D. Szafr. 2019. “Designing Communicative Visualization for People with Intellectual Developmental Disabilities”. In: *VIS 2019 Workshop on Visualization for Communication (VisComm)*. OSF Preprints.
- Xia, H., B. Araujo, T. Grossman, and D. Wigdor. 2016. “Object-oriented drawing”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 4610–4621.
- Xia, H., N. Henry Riche, F. Chevalier, B. De Araujo, and D. Wigdor. 2018. “DataInk: Direct and creative data-oriented drawing”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1–13.
- Yang, Y., K. Marriott, M. Butler, C. Goncu, and L. Holloway. 2020. “Tactile presentation of network data: Text, matrix or diagram?” In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1–12.
- Yoon, D., N. Chen, F. Guimbretière, and A. Sellen. 2014. “RichReview: blending ink, speech, and gesture to support collaborative document review”. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*. 481–490.

- Yoon, D., N. Chen, B. Randles, A. Cheatle, C. E. Löckenhoff, S. J. Jackson, A. Sellen, and F. Guimbretière. 2016. “RichReview++: deployment of a collaborative multi-modal annotation system for instructor feedback and peer discussion”. In: *Proceedings of the ACM Conference on Computer-Supported Cooperative Work & Social Computing*. 195–205.
- Yoon, D. and P. Mitros. 2015. “Multi-Modal Peer Discussion with RichReview on edX”. In: *Adjunct Proceedings of the ACM Symposium on User Interface Software and Technology*. 55–56.
- Yu, B. and C. T. Silva. 2019. “FlowSense: A natural language interface for visual data exploration within a dataflow system”. *IEEE Transactions on Visualization and Computer Graphics*. 26(1): 1–11.
- Zraggen, E., R. Zeleznik, and S. M. Drucker. 2014. “PanoramicData: Data analysis through pen & touch”. *IEEE Transactions on Visualization and Computer Graphics*. 20(12): 2112–2121.
- Zraggen, E., R. Zeleznik, and P. Eichmann. 2016. “Tableur: Handwritten spreadsheets”. In: *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems*. 2362–2368.
- Zhang, J. E., N. Sultanum, A. Bezerianos, and F. Chevalier. 2020. “DataQuilt: Extracting Visual Elements from Images to Craft Pictorial Visualizations”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1–13.
- Zhang, Y., M. Pahud, C. Holz, H. Xia, G. Laput, M. McGuffin, X. Tu, A. Mittereder, F. Su, W. Buxton, and K. Hinckley. 2019. “Sensing posture-aware pen+ touch interaction on tablets”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems*. 1–14.