THE UNIVERSITY OF CALGARY

Collaborative Information Visualization in Co-located Environments

by

Petra Isenberg

A DISSERTATION

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF COMPUTER SCIENCE

CALGARY, ALBERTA

DECEMBER, 2009

UNIVERSITY OF CALGARY
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Collaborative Information Visualization in Co-located Environments" submitted by Petra Isenberg in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

Supervisor, Dr. Sheelagh Carpendale
Department of Computer Science

Dr. Saul Greenberg
Department of Computer Science

Dr. Amy Ashurst Gooch
Department of Computer Science, University of Victoria

Dr. Patrick Shiao Tsong Feng
Department of Communications & Culture

Dr. Colin Ware
External, University of New Hampshire

Date

Abstract

Information visualization research has been developing new methods to represent data and interact with graphical displays of information for more than two decades. In many disciplines, however, the size and complexity of datasets are rapidly growing. As a consequence, it is becoming increasingly necessary to join the domain expertise and data analysis skills of several people to inform decisions about the content of a dataset. While the technological possibilities for supporting teamwork are gradually evolving, several obstacles remain for designing information visualizations that can support team members as they collaboratively explore and analyze information. In this dissertation, I examine this problem by identifying and addressing some of the open issues in the design of information visualizations that support small teams of experts in their joint data analysis activities.

Within the general area of collaborative visualization, this research is scoped to focus on a subset of collaborative visualization scenarios that occur in co-located synchronous work environments; where small groups of collaborators share the same physical workspace such as a large digital table or wall display. Specifically, it contributes to a richer understanding of how groups work with each other and with information visualizations in phases of joint and parallel work.

In this dissertation, I show that team members tend to prefer working in parallel on specific types of information analysis tasks and more closely together on others. During phases of parallel work, individual team members take on unique approaches to data analysis. Thus, for the design of collaborative analysis systems, the support of unique analysis approaches and a flexible temporal flow of activities—both in the temporal sequence and co-occurrence of work styles in groups—need to be considered. In addition, the three case studies presented in this dissertation examine possibilities of how this flexibility can be supported. These case studies shed light on issues of parallel and

joint work with multiple views in a collaborative system, parallel and joint work with a single shared visualization, as well as awareness support during parallel work. In summary, this dissertation contributes to the evolving understanding of collaborative work practices around information visualizations and introduces several specific design considerations.

Acknowledgements

This acknowledgements section is the last part of my dissertation left to write and, in a sense, it is one of the most difficult. Completing a thesis is a challenge, thanking all those who contributed to it is an even greater one. Many people influenced me, shaped my thinking and my work and I want to thank all of them, including those not mentioned here by name.

My advisor Sheelagh Carpendale was instrumental throughout it all. Sheelagh was the first to open my eyes to research in information visualization when she gave me the opportunity to work with her when I was "just" some unknown German student emailing her about an internship. As a PhD advisor, Sheelagh has been a great source of inspiration, enthusiasm, critique, and general life and research support. She found the right balance in correcting me when necessary and letting me explore and follow my own research paths. It is mostly from her that I learned how to be a researcher and how to see things in new and unexpected ways.

Thanks also go to my committee members Saul Greenberg, Amy Gooch, Patrick Feng, and Colin Ware for their time, energy, and feedback on my work. In particular, I would like to thank Saul for providing invaluable direction when I needed help and for asking challenging questions about my work.

I was particularly fortunate to be surrounded by amazing graduate colleagues at the Interactions Lab who were always willing to give advice and inspiration and who were not too shy to challenge my research. Thanks especially go to my close friends in the iLab who provided needed diversion and research "counselling" during the many ups and downs of the PhD experience (you know who you are). Special thanks go to: Chris Collins for all the InfoVis advice and for lively discussions about research in general; Mark Hancock for being my immediate cubicle neighbour, enduring my many questions about coding, study design, and stats, and for interesting discussions about almost anything; Uta Hinrichs for sticking around for the whole journey from undergrad to PhD (10 years!) and for giving a lot of advice on my work; Kim Tee for being an enthusiastic sports companion, amazing proof-reader, and for the many tasty meals she provided during midnight writing sessions; Tony Tang for helping me to define my thesis topic, doing my first qualitative study with me, and for asking deep questions that helped me further my research ideas; Matthew Tobiasz for allowing me to work with him on his collaborative infovis project and for teaching me some new C++ skillz. Countless additional thanks also to all other iLab members past and present who have made this research endeavour so much more fun and enjoyable.

During two research visits I met a number of amazing colleagues. In particular, I would like to thank Jean-Daniel Fekete, Anastasia Bezerianos, and Nathalie Henry for their support during my visit to INRIA and who enthusiastically worked with me on a collaborative InfoVis project. Thanks also to Danyel Fisher for challenging me to think about our research project at MSR and my research in general. Many thanks to Kori Inkpen, Bongshin Lee, Tim Dwyer, Chris North, George Robertson, Mary Czerwinski, Merrie Morris, colleagues and fellow interns at MSR for providing valuable feedback and new perspectives on my work. Thanks also to all my co-authors over the years who have greatly shaped my work and research experience. Figure 1 shows who you are.

Generous support from Alberta Ingenuity and iCORE made it possible for me to come to Calgary for my PhD.

I save for last the people who deserve thanks more than others. I would like to specially thank my husband Tobias (Floh) Isenberg for his uncompromising support and love through all stages that a PhD brings with it, for understanding stressful times, and for always managing to encourage me. I would also like to thank my family for their patience and support and for being understanding when I had to work even during visits back home. And I'd like to especially thank my sister who always put everything into perspective.

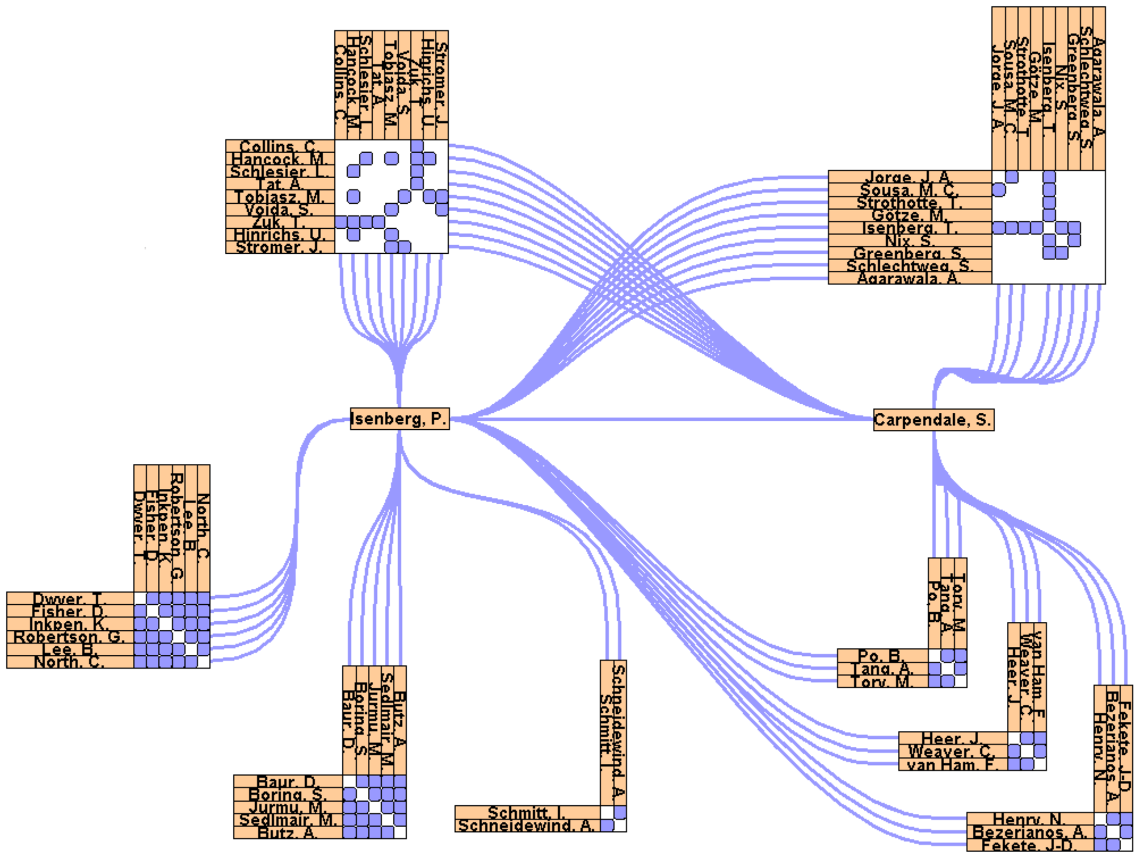


Figure 1: I am indebted to my co-authors for inviting me to collaborate with them, for collaborating with me, for revising and improving my papers, and giving important feedback on my work.

To my friends and family

WITH GOOD REASON

Publications

Materials, ideas, tables, and figures in this thesis have appeared previously in the publications below. After each reference, I note the chapters in which material is used.

Journal Articles

- **Petra Isenberg**, Anastasia Bezerianos, Nathalie Henry, Sheelagh Carpendale, and Jean- Daniel Fekete. *CoCoNutTrix: A Study in Collaborative Retrofitting for Information Visualization*. Computer Graphics and Applications: Special Issue on Collaborative Visualization, 29(5):44–57, September/October 2009.
  Material from this publication appears in Chapter 6.
- **Petra Isenberg** and Danyel Fisher. *Collaborative Brushing and Linking for Co-located Visual Analytics of Document Collections*. Computer Graphics Forum (Proceedings of EuroVis), 28(3):1031–1038, June 2009. Material from this publication appears in Chapter 7.
- **Petra Isenberg** and Sheelagh Carpendale. *Interactive Tree Comparison for Co-located Collaborative Information Visualization*. IEEE Transactions on Visualization and Computer Graphics (Proceedings Visualization / Information Visualization 2007), 12(5):1232–1238, 2007. Material from this publication appears in Chapters 3 and 5. © 2007 IEEE. Portions reprinted with permission from IEEE.

Conference Papers

- **Petra Isenberg**, Anthony Tang, and Sheelagh Carpendale. *An Exploratory Study of Visual Information Analysis*. In Proceedings of the Conference on Human Factors in Computing Systems (CHI), pages 1217–1226, New York, 2008. ACM Press. Material from this publication appears in Chapter 4.

- **Petra Isenberg**. *Information Visualization in Co-located Collaborative Environments*. In Proceedings of the Grace Hopper Celebration of Women in Computing, PhD Forum, 2007. Material from this publication has been used in Chapter 1.

Book Chapter

- **Petra Isenberg**, Uta Hinrichs, Mark Hancock, and Sheelagh Carpendale. *Digital Tables for Collaborative Information Exploration*. In Tabletops - Horizontal Interactive Surfaces. Springer Verlag, Early 2010. To appear. Material from this publication appears in Chapter 3.
- Jeffrey Heer, Frank van Ham, Sheelagh Carpendale, Chris Weaver, and **Petra Isenberg**. *Creation and Collaboration: Engaging New Audiences for Information Visualization*. In Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, and Chris North, editors, Information Visualization|Human-Centered Issues and Perspectives, volume 4950 of LNCS State-of-the-Art Survey, pages 92–133. Springer Verlag, 2008. Material from this publication appears in Chapter 3.

Workshop Papers

- **Petra Isenberg** and Sheelagh Carpendale. *Social Data Analysis in Co-located Environments*. In CHI Workshop on Social Data Analysis, 2008. Material from this publication has been used in Chapter 1.
- **Petra Isenberg**, Uta Hinrichs, Mark Hancock, Matthew Tobiasz, and Sheelagh Carpendale. Information Visualization on Interactive Tabletops in Work vs. Public Settings. In Proceedings of the VisWeek Workshop on Collaborative Visualization on Interactive Surfaces (CoVIS 2009, October 11, 2009, Atlantic City, USA), 2009. Material from this publication appears in Chapter 3.

CONTENTS

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

In recent years, we have seen information visualization tools receive more general adoption and integration into commercial products (Shneiderman and Plaisant, 2009). As a result, information visualizations are increasingly becoming essential tools for information analysis, exploration, and understanding tasks. This is in part the case because visual displays of information have several benefits over exploring information in textual or numerical form. Visualizations can increase memory and processing resources available to a viewer by encoding data visually and making use of our high-bandwidth visual system to perceive these encodings (Card et al., 1999; Ware, 2000). Looking at visual encodings of data has been shown to reduce search time, enhance detection of anticipated or unanticipated patterns, enable perceptual inference operations and hypothesis formulation, help the monitoring of changing data, and help data exploration by providing a manipulable medium (Card et al., 1999; Ware, 2000). Yet, simply finding an effective visual encoding for a given information source is often not enough to aid an individual in completing an information-related task. For example, as datasets become increasing large and complex, several analysts often have to join their knowledge, expertise, and analysis skills in order to be able to make informed decisions about these information-rich datasets (Thomas and Cook, 2005). So far, research in visualization has largely focused on supporting the data analysis activities of a single person. How, in contrast, the collaborative work of teams sharing, discussing, and interpreting information visualizations can be supported, is the general problem area of this dissertation.

In this chapter, I introduce this research context, point out that my scope is within *collaborative information visualization in co-located environments* and define this restricted problem domain of my dissertation work. Finally, I outline my research goals, and end the chapter with a brief overview of the entire dissertation's structure.

1.1 Collaborative Visualization

During the timeframe of this dissertation, a trend towards collaborative data analysis and exploration has emerged in information visualization. From 2007 to 2009, websites such as ManyEyes (Viégas et al., 2007), iCharts (iCharts Inc., 2008), Verifiable.com (Visible Certainty, 2009), or Swivel (Dimov and Mulloy, 2009) have emerged and are among the first to serve as platforms for joint viewing, creation, and discussion of visualization on an Internet scale. Yet, social interaction around data is not a new phenomenon. In everyday practice, data is frequently interpreted, analyzed, and explored not only by individuals but also by teams who work in concert to make decisions, form actions, or learn about information. Such joint and social activities around visualizations are central to *collaborative visualization*. The question of how to best support data analysis as a social process unites several different research approaches and raises interesting new issues for the field of visualization. For example, one research approach (e. g., Viégas et al. (2007)) focuses on the use of online social media techniques such as comment threads, annotations, and bookmarking, in order to bring together large numbers of, possibly unacquainted, people for playful and open-ended exploration or discussion of information. Other researchers have focused on the support of synchronous data analysis meetings in distributed settings (see an overview by Anupam et al. (1994)). Here, another set of questions have to be addressed such as which network and software architectures support synchronous interactions with visualizations, how can joint viewpoints be coordinated, and how can synchronous online discussions be facilitated in reference to the data on the screen. In this thesis, I take a different approach: I focus on small group, task- and work-oriented collaboration in a physically shared work environment. I concentrate on collaboration around large single-display technology such as wall or tabletop displays to facilitate co-located collaborative data analysis activities with computer support. To set the research of this dissertation in a greater context, however, first a more precise definition of collaborative visualization is necessary.

1.2 Context and Definition

Previously, several definitions have been given to describe *specific aspects* of collaborative visualization. None, however, have attempted to give an encompassing definition of the entire scope of group work around visual representations of data. I discuss four of these definitions in the following, note their limitations, and finally provide my own definition for collaborative visualization.

One of the earliest definitions by Raje et al. (1998) emphasizes the goal of collaborative visualization:

> *"Collaborative visualization enhances the traditional visualization by bringing together many experts so that each can contribute toward the common goal of the understanding of the object, phenomenon, or data under investigation."*
>
> *Raje et al. (1998)*

While bringing experts together is an advantage in some collaborative visualization scenarios, collaborators often do not need to be experts. Non-experts can join in collaborative analyses and learn from others' analysis processes and viewpoints on a dataset (Heer et al., 2008). Similar to this restriction by type of collaborators, other definitions may have been too restrictive in terms of the applicable fields:

> *"The term "collaborative visualization" refers to a subset of CSCW applications in which control over parameters or products of the scientific visualization process is shared."*
>
> *Johnson (1998)*

> *"Collaborative visualization [...] allows geographically separated users to access a shared virtual environment to visualize and manipulate datasets for problem solving without physical travel."*
>
> *Li et al. (2006)*

The first definition emphasizes collaboration with interactive, manipulable visualization for the scientific visualization community. The restriction to only the scientific visualization community is overly limiting as the information visualization community can similarly make use of collaborative systems to analyze data. The second definition emphasizes distributed visualization in virtual environments. This is also too limiting because research on groupware systems has a long tradition in both distributed as well as co-located spatial domains. The limitation to virtual environments is another unnecessary restriction. Collaborative visualization also has had numerous applications outside of virtual environments (see Chapter 2). The restriction to only interactive visualizations in both definitions may also be limiting and it is still being debated whether interactivity should be a part of a general definition of visualization (e. g., Pousman et al. (2007)). For example, groups of people may often come together to discuss static visualizations printed on posters, in handouts, or projected as a slideshow. In these scenarios, social interaction around data does occur but the interaction with data may be limited to the selection of which data to look at. In this dissertation, I only consider collaboration with interactive visualizations.

Recently, the term *social data analysis* has been coined to describe the social interaction that is a central part of collaborative visualization:

> *"[Social data analysis is] a version of exploratory data analysis that relies on social interaction as source of inspiration and motivation."*
>
> *Wattenberg (2005)*

The term *social data analysis* emphasizes the possibility of human interactions such as discussions, negotiations, or arguments around visualizations as the driving factors of data exploration. Yet, social interaction around data may occur in more scenarios than just exploratory data analysis. For example, targeted or confirmatory data analysis, teaching, learning, or decision-making scenarios may also frequently involve collaboration. In addition, the term social data analysis has an unfortunate ambiguous connotation in that it could refer to the analysis of social data, such as social networks, email graphs, or instant messaging chats. Even in the form of the definition given by Wattenberg (2005) (see above), it has recently been increasingly used to solely describe web-based social media approaches to collaborative visualization. In order to more broadly describe the entire scope that collaborative visualization can encompass, I use the term *collaborative visualization* as follows to describe my research context in a more general way:

> *Collaborative visualization is the* shared *use of computer-supported, possibly interactive, visual representations of data by more than one person with the common goal of contribution to joint information processing activities.*

This definition is derived from a general definition for visualization as *the use of computer-supported, interactive, visual representations of data to amplify cognition* (Card et al., 1999). It has been augmented by emphasizing the *shared* use of (interactive) visual representations—which could be in the form of joint viewing, interacting with, discussing, or interpreting the representation. Secondly, the word "cognition" has been replaced with the word *"information processing."* This replacement honours the fact that different theories exist for how cognition applies when groups come together to jointly think and reason. Each theory has different terminology, restrictions, and units of analysis. For example, the theory of Group Cognition (Stahl, 2006) describes collaborative knowledge building for small groups by focusing on linguistic analysis, Distributed Cognition (Hutchins, 1996) focuses on social aspects of cognition by analyzing the coordination between individuals and artifacts, and Communities of Practice (Wenger, 1999) describe learning within much larger social communities. In order to avoid favouring any specific theory or unit of analysis, the word *information processing* has been used here as a general term to describe cognitive activities involved in individual or collaborative processing of visual information, such as reading, understanding, applying knowledge, discussing, or interpreting.

Given this broad definition of collaborative visualization, we can look at a number of different scenarios in which it may occur. While much of the early research in collaborative visualization has focused on remote collaborators, several other scenarios fall under our broad definition. Using the space-time matrix (Dix et al., 1998), we can broadly categorize collaborative scenarios according to where they occur in space (distributed vs. co-located) and time (synchronous vs. asynchronous). These distinctions for systems or tools are not strict—systems can cross boundaries and could, for example, be used both synchronously or asynchronously, as pointed out by Dix et al. (1998) for the example of e-mail. E-mail can be used similar to a chat client in synchronous work or asynchronously in conversations that stretch over longer periods of time. Fig-

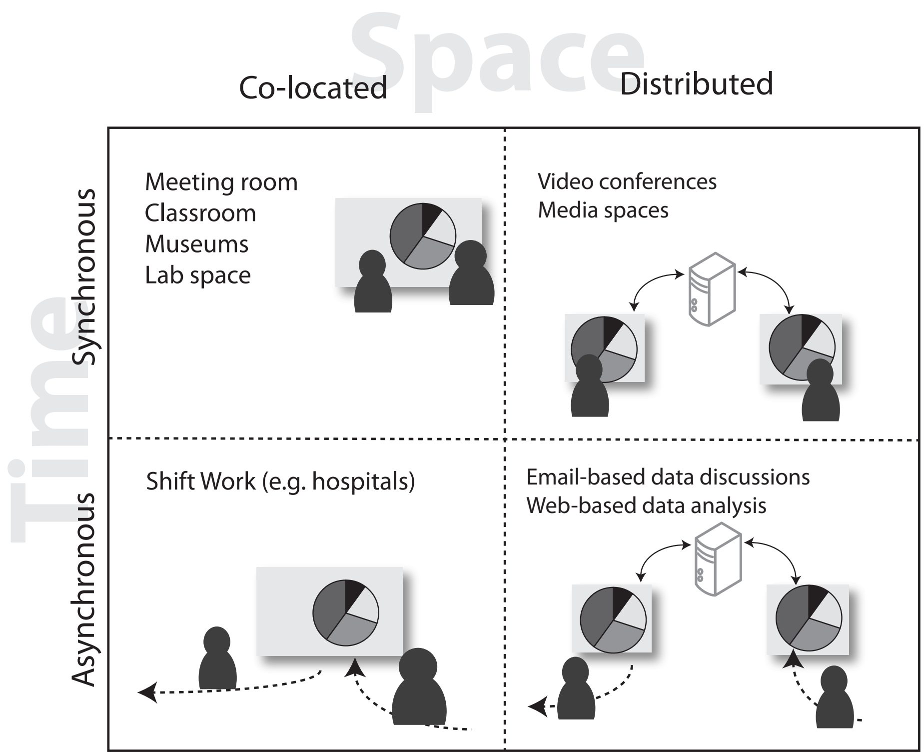


Figure 1.1: Collaborative visualization can occur in many scenarios delineated according to space and time (matrix adapted from Dix et al. (1998)).

ure 1.1 shows several scenarios in which collaborative visualization can occur. I will use this distinction again to more narrowly define the scope of this dissertation work.

At the same time, collaborative information visualization may occur on different levels of engagement with the visualizations. The larger group involved in the social interaction around data can simply view the information, actively interact with and explore it, or even join in creating new visualizations and share those and the underlying datasets with a larger community (Zambrano and Engelhardt, 2008). Different digital systems have been designed to support collaborative visualizations along these different levels of engagement. A few example scenarios are presented next:

Viewing: Presentation systems such as PowerPoint or simple videoconferencing tools can support a group of people viewing static or animated visualizations of data without being able to interact with or annotate the information. Such scenarios could occur, for example, in classrooms or meetings where one presenter explains, teaches, or summarizes information for the larger group. The goal of the group may be to learn, discuss, interpret, or form decisions from a pre-selected set of information and visualizations.

Interacting / Exploring: When groups of people share the same interactive visualization software, either in co-located or distributed settings, they can choose and select alternative views of the data for its exploration, analysis, discussion, and interpretation. In distributed settings, findings can typically be exchanged through chats, e-mail, or a video- or audio-link so that the changing views and alternative representations of the data can be discussed and analyzed. This discussion can also occur face-to-face in co-located settings. The goal of the group is to be able to cover and explore different and more aspects of the data, consider alternative interpretations, and discuss the data in a wider visual context.

Sharing / Creating: Through the emerging trend of user-generated content sites for visualization (e. g., in systems such as ManyEyes (Viégas et al., 2007)), many people are able to create, upload, and share new datasets and visualizations. Often this type of sharing is done with a greater community to raise awareness about a certain issue.

The distinction between these levels of engagement can be blurred. Digital systems may, for example, be intended to mainly support collaborative interaction and exploration of data but may also support the sharing and creation of new visualizations or even the download of new datasets to visualize. However, both time and space dimensions as well as levels of engagement can help to broadly scope a research focus within collaborative visualization.

1.3 Research Scope

Within the general definition for collaborative visualization developed in the previous section, I focus the research scope of this dissertation on a subset of collaborative visualization scenarios that:

- occur in **co-located** environments, where several collaborators share the same physical workspace,
- use **a single shared** workspace such as an interactive digital tabletop or wall display,
- occur in **synchronous** work settings, where collaborative analysis occurs at the same time,
- include **interactive** information visualizations as their main type of visualization,
- include **small groups** of 2–4 individuals,
- support mainly **interaction and exploration** of data with the goal to analyze the encoded information.

Several motivations lie behind this narrower research scope. First of all, *interactive information visualizations* have shown many benefits for individual data analysis as already outlined in the previous section (Card et al., 1999; Ware, 2000). Collaboration around visualizations can also have a number of benefits over individual work. In many disciplines, collaboration allows for multi-disciplinary groups with increased skill sets. Different team members offer different perspectives and expertise that together can improve the quality of solutions and decisions. Also, the analyzed information space may often simply be too complex for an individual to interpret in its entirety. With large data sets, even the task load of *exploring* the data could be shared among several individuals on a team (Thomas and Cook, 2005). The benefits that collaboration offers to this process have motivated my shift from developing information visualization tools for use by just one person toward research on the design of *collaborative* information visualization tools.

Developing collaborative visualization tools is a promising endeavour as several aspects of information work already often involve group work: the acquisition of information, the analysis and interpretation of information, sharing and interpretation of analysis results, and decision making (Chuah and Roth, 2003). Each of these tasks can be supported with digital tools and each has its own design considerations. In this dissertation, I look at a subset of these tasks—collaboration around analysis and interpretation of the visualized data. I assume that data has already been collected and is available for joint analysis. I do not include research on specific functionality to share and disseminate the joint analysis results with outside groups or stakeholders. Both the collaborative collection of data and the dissemination of analysis results are important aspects of collaborative work with information but are complex research endeavours of their own.

Collaboration with the goal to *analyze data* in *small groups* and *co-located synchronous work settings* is relatively common in current workspaces, however, there is currently little software support for this type of work. While it is possible for small teams to work with information visualizations using the standard setup of a small screen, one mouse, and one keyboard—only one person at a time is able to make any changes to the view of the system. Attempting to collaborate under these conditions can be awkward and unnatural. The recent trend toward the use of large interactive displays offers the potential for the development of improved collaborative information visualization systems in which many co-located team members can simultaneously interact and explore data sets. However, it is not yet well understood how interfaces, visualizations, and interaction techniques should be designed to specifically address the needs of small co-located groups.

With this scope, the research topic of this dissertation lies at the intersection of two main fields as seen in Figure 1.2: Information Visualization (InfoVis) and Computer Supported Cooperative Work (CSCW). In the field of information visualization, researchers have been working towards developing new visual representations, presentation, and interaction techniques to amplify human cognition for different types of datasets, tasks, and analysis scenarios (Card et al., 1999; Ware, 2000; Chen, 2006; Spence, 2007a). Research from the field of information visualization informs the topics of this dissertation through its discussion of how individuals work with and perceive visual data representations, how they perform data analysis, and how to design interactive information visualization systems to support these work processes. The field of CSCW (Dix et al., 1998, Chapter 13) is concerned with the challenges of designing software for multiple people to work as a group and how to understand the effect of deployed software on their work processes. Within CSCW in particular, the work on

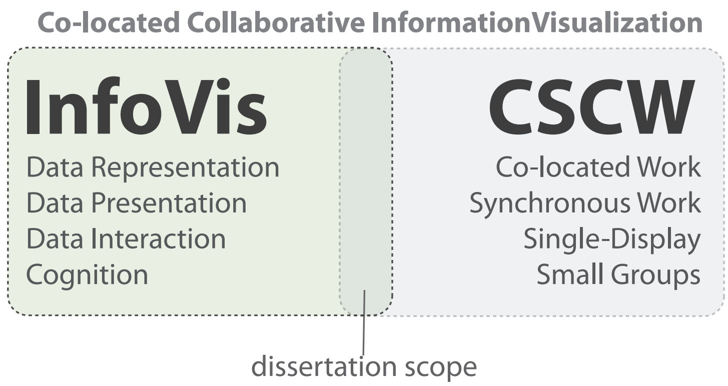
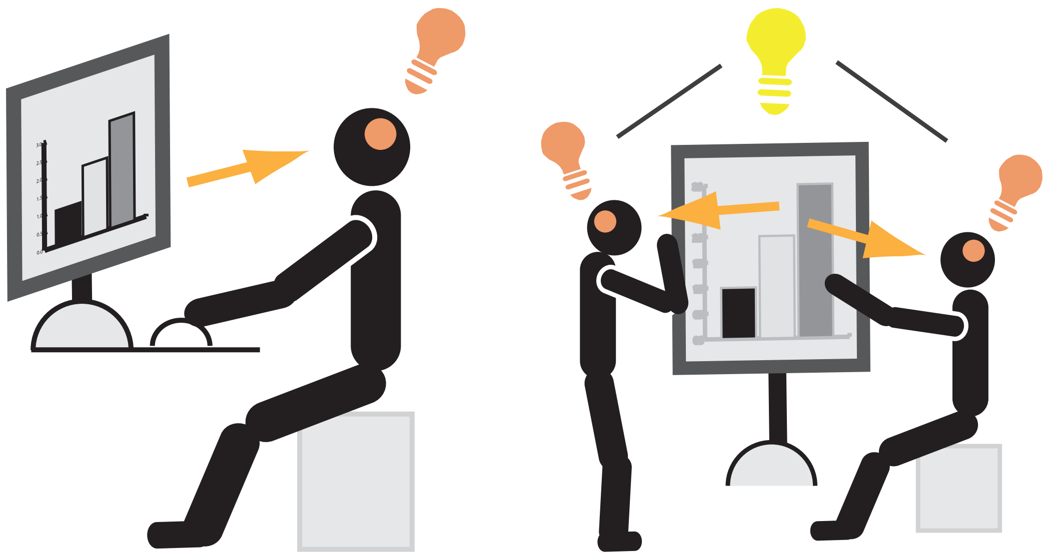


Figure 1.2: Research scope of this dissertation.

single-display synchronous co-located groupware for small teams has a high applicability to the research of this dissertation. I further discuss the applicability of related work in these areas to the specific problem domain of this dissertation in Chapter 3.

1.4 Research Goals

The design of collaborative systems poses challenges in addition to those encountered during the design of information visualization systems that are intended to be used by a single person. In a group setting, the use of co-located collaborative technology needs to support a process of social interaction around the data. Ideally, it should help the group to arrive at a *common* understanding of the data through a process of collaborative interpretation, analysis, discussion, and interaction. For example, in Figure 1.3a a single person works with an information display through a process of looking at and possibly interacting with an information visualization, forming a mental model by interpreting the representation, and ideally gaining an insight and forming a decision (Spence, 2007a). In Figure 1.3b two people join in a collaborative analysis. They both come to individual insights by looking at and interpreting the dataset. However, through social interaction (e. g., discussion and negotiation) they both should reach a common understanding of the dataset in order for both of them to make informed decisions as a group, derive common recommendations, or take joint next step actions together after the analysis. The visionary goal would be that through using collaborative visualization tools, groups are able to gain additional understanding, knowledge,



(a) A single analyst views digital visualizations, engages in cognitive processing about the visual data display, and, ideally, arrives at an insight or discovery about the data set.

(b) Two analysts join in a collaborative analysis scenario. The goal for both is to come to a common understanding of the data through visualization use and possibly derive joint next step actions.

Figure 1.3: Goals of individual and collaborative information visualization.

and insight into the data—different or more encompassing than would have been possible had they explored the data individually.

A challenge in designing information visualization for synchronous and co-located collaborative work is that mechanisms need to be designed that support the ways people work collaboratively, *as a group*, during an analysis. It is still relatively unexplored how to design these systems so that they support the generation of a common understanding through collaborative interaction with and analysis of information visualizations. With this thesis, I do not attempt to describe all the ways that technology does or could impact synchronous co-located collaborative data analysis but instead I have a specific set of research goals.

From information analysis practices of single analysts, we know that people have their own data analysis approaches and styles and hence use varying techniques to solve an analysis problem (Mirel, 2004). When new insights emerge analysts often want to diverge from their current path of inquiry and redirect their investigations or change their analysis approach altogether. For the design of collaborative information visualization systems, this poses an interesting challenge. When we cannot assume a fixed data analysis approach for a single person, we cannot assume that people in a group would agree on one approach when working together. The question arises: how can we build collaborative analysis systems that are sensitive to the needs of the individual, but also the needs of the group as a whole, without polarizing these needs against each other? This has previously been identified as a research challenge in distributed settings (Gutwin and Greenberg, 1998). Exploring ways in which small groups of people can balance their need to analyze data individually and capitalize on the group's shared information processing in a co-located data analysis environment is the main research goal of this dissertation. This research goal was, in part, motivated by an earlier research project (Tang et al., 2006) that I was involved in. This project (led by A. Tang) looked at how people transitioned between several different phases of joint and parallel work while engaged in a route finding task on a digital tabletop display (see Figure 1.4). We found that pairs dropped in and out of different phases of mixed-focus collaboration (shared and individual work) (Gutwin and Greenberg, 1998) and that their different working styles depended on preferred tools, physical arrangement, and the incidence and handling of interference.

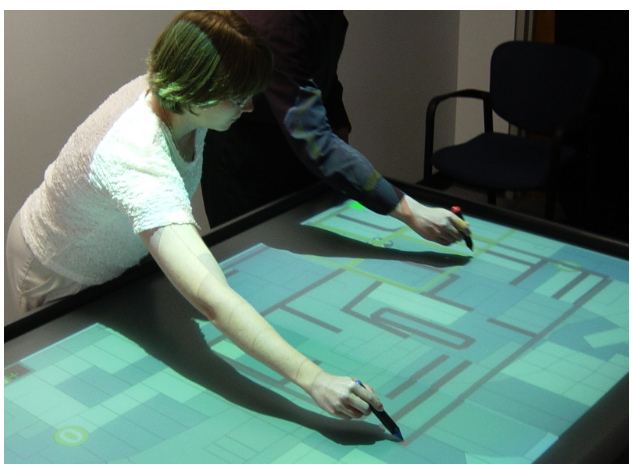


Figure 1.4: Collaborative route finding task over a shared information display as part of a study presented by Tang et al. (2006).

This initial study shed some light on working styles around information analysis tasks but the role of the information representation and the data analysis task remained largely unexplored. Exactly how these phases would manifest in relation to information analysis processes and how visual representations should be designed so that both individual as well as joint work within the same task can be supported, without restricting one another, remained an open question.

The goal of my dissertation research, therefore, is to more specifically explore and to generate a richer description of how information visualizations and work spaces can be designed to support a variety of work styles in collaboration. Towards this goal, I address three specific research challenges:

Challenge 1: We do not have a clear understanding of how related work in CSCW and Information Visualization areas applies to the specific problem of supporting co-located data analysis.

Challenge 2: We do not understand the data analysis practices and processes of small teams. How do they analyze information together and how are information visualizations used in this context?

Challenge 3: We do not know how to design collaborative information visualization systems for co-located work. In particular, we do not know how we can support team members transitioning from parallel to joint work phases.

To address these research challenges, I followed the methodological approach as outlined in the following section.

1.5 Methodological Approach

My research strategy has been to first look at general issues within my research scope and then, in a second phase, to consider specific design challenges through specific implementations of collaborative data analysis systems. The first phase of my research involved summarizing a first set of design considerations from both a literature review and a study of collaborative visualization use in a physical environment. Findings from this phase of my work inspired the design of three specific systems for collaborative information analysis work in the second phase of my research. Two of the implemented systems were subsequently studied in relation to the more general findings and hypotheses from the first phase of my work. During the first phase of my research, a general finding emerged: when working on information analysis tasks, groups show a strong tendency for individual work when searching for, interacting with, reading, and analyzing data but frequently switch to joint work phases to validate, discuss, or interpret information. In the second phase of my research, I focused specifically on the question of how this switch could be supported in information visualization systems so that both types of work are possible, well supported, and that the transition between both is fluid. I designed three systems to shed light on this question in different data and task domains: tree comparison in a biological context, social network analysis, and information foraging in an intelligence analysis task.

Throughout this dissertation, I carefully chose appropriate evaluation methodologies for the evaluation in pre-design (in the first phase) or post-design (in the second phase). During the first phase, I used a mainly qualitative research approach to study people collaboratively analyzing information without existing software support. The goal was to develop a richer understanding of basic work processes that can be used to inform interface design through observations of people's interactions with physical artifacts. Other researchers (e. g., Tang (1991); Scott et al. (2004)) have taken this approach, studying how groups accomplish tasks in *non-digital contexts* in order to understand what activities *digital tools* should support. The reasoning behind this choice is that people's *physical interactions* with these familiar artifacts and tools would closely reflect how they *understand and think* about the problem at hand. For instance, Tang's study of group design activities around shared tabletop workspaces (Tang, 1991) revealed the importance of gestures and the workspace itself in mediating and coordinating collaborative work. Similarly, Scott et al. (2004) studied traditional tabletop game play and collaborative design, focusing on the use of tabletop space and the sharing of items on the table. While these authors studied traditional physical contexts, ultimately their goal was to understand how to design *digital* tabletop tools. Both studies contributed to a better understanding of collaborative work practices involving tables in general. The approach taken in these two studies works well when addressing a design area where the critical issues are poorly understood. For instance, we are *uncertain* how groups will work together with information visualizations if given the ability to do so freely (e. g. prior efforts involved systems where individuals could not work in parallel (Park et al., 2000; Mark and Kobsa, 2005)). Furthermore, we do not know how teams will share and make use of intermediate results, or indeed whether they will even share and work together from the same views or artifacts of the data.

The three case studies developed during the second phase, build both on design challenges uncovered in the first phase of my research as well as on issues uncovered during the study of the case studies themselves. To evaluate these specific examples, I used mixed-methods approaches, combining observational techniques, questionnaires, system logs, and interviews to assess how the interfaces were used and influenced group work. Details about the study methods used for each of these projects are outlined in the respective chapters.

1.6 Contributions

This research builds on previous knowledge from the areas of Information Visualization and Computer-Supported Cooperative Work. It contributes to a richer understanding of how groups make use of shared information visualizations on large interactive displays to gain insight into data and solve problems. Specifically, this work includes three main contributions:

1. This dissertation extends our evolving understanding of collaborative work practices around information visualizations. Different collaborative data analysis processes are identified and described within different phases of joint and parallel group work.

2. Based on three examples, this dissertation demonstrates how co-located collaborative systems can be designed to support group analysis. Experimental findings assess two of the presented designs.

3. This dissertation develops a first set of design considerations for information visualization systems in co-located shared screen settings. These design considerations are derived from a literature review, a study of collaboration practices in a physical setting, as well as two studies of group work with information visualizations in co-located settings.

These contributions are discussed and explored in more detail in the following chapters.

1.7 Organizational Overview

After this introductory chapter, the remainder of the thesis is organized as follows.

Chapter 2:

This chapter forms the first part of a literature review on collaborative information visualization. I give an overview of previous collaborative visualization systems for co-located work, including a systematic review of systems featured in the IEEE Vis/InfoVis/ Vast conferences. I highlight their main features and point to open research questions and extensions of this work.

Chapter 3:

This chapter forms the second part of my literature review and is part of my first research phase. I discuss work from information visualization design, co-located collaboration, and studies that look directly at collaborative visualization. I examine research from these areas in relation to this dissertation work and derive initial design considerations for the design of co-located collaborative information visualizations systems.

Chapter 4:

As part of the first phase of my research, I report on an exploratory study of individuals, pairs, and triples engaged in information analysis tasks using paper-based visualizations. From the study results, eight specific analysis processes are derived that capture the analysis activities of co-located teams and individuals. Comparing these with existing models of the information analysis process suggests that information visualization tools may benefit from providing a flexible temporal flow of analysis actions and that collaborative information visualization systems should support people in fluidly switching between different types of analysis processes. These findings extend the initial design considerations derived in Chapter 3.

Chapter 5:

With this chapter I begin the second phase of my research. This second phase contains three new collaborative information visualization systems. Here in this chapter, I present the first of these three new collaborative systems for co-located data analysis, CoTree. It is based on the considerations derived from work in the two previous chapters. In CoTree, I focused on first providing ways to enable parallel work processes and then included more subtle workspace-based mechanisms for team members to switch to more joint work styles. The system was designed to support hierarchical data comparison tasks for co-located collaborative work. It supports dual-touch input, shared and individual views on the hierarchical data visualization, flexible use of representations, and flexible workspace organization. I discuss this initial design and point to further research questions arising from this prototype.

Chapter 6:

In this chapter, I present a tool and subsequent study in which I explored how a co-located collaborative information visualization and analysis environment can be retrofit from a pre-existing system design for use by a single analyst. This design takes an orthogonal approach to the one used in Chapter 5. I start from a system designed to support only sequential close work and looked at minimal changes necessary to introduce possibilities for parallel work. These changes were based on the results from my previous work and the design considerations developed in Chapter 3. NodeTrix, a social network analysis tool for individual use, was extended to enable parallel interaction in collaborative environments. Details of the retrofitting process and results of a study show the usability of the retrofitted system. The results support the effectiveness of the low-cost collaborative retrofitting for collaborative network analysis and highlight implications for practitioners.

Chapter 7:

In this chapter, I present the design of a tabletop visual analytics tool, Cambiera. Cambiera, supports individual and collaborative information foraging activities in large text document collections. With the design of this system, I take an approach that includes ideas from both Chapters 5 and 6. Similar to CoTree (Chapter 5), I propose a new design specifically tailored towards parallel work but introduce mechanisms to allow people to be more closely aware of each others' activities, a suggestion that came out of the study in Chapter 6. The design of this system focused specifically on the question of how individual and joint analysis activities could be supported with meta-visualizations.

'Collaborative brushing and linking' is defined as an awareness mechanism that enables analysts to follow their own hypotheses during collaborative sessions while still remaining aware of the group's activities. With Cambiera, team members are able to collaboratively search through documents, maintaining awareness of each others' work and building on each others' findings.

Chapter 8:

In the conclusions, I summarize the research objectives and contributions of this thesis and shed light on future issues in co-located collaborative information visualization.

Chapter 2

Research Background

Co-located collaborative information visualization is a research area that is relatively new and still under explored. While there has been considerable research in both CSCW and Information Visualization, comparatively little research has looked at the intersection of both areas and even fewer systems have been developed specifically for co-located data analysis. The research from both CSCW and Information Visualization, however, includes useful information that can be analyzed and used as a basis for my research into co-located collaborative information visualization. I have separated these two types of literature. In this chapter, I discuss the research literature that describes systems specifically related to my research scope: co-located synchronous work with information visualizations. In Chapter 3, I examine relevant literature from other work contexts to generate a set of design considerations upon which I base the research presented in the remaining chapters.

2.1 Introduction

Only a few collaborative analysis systems have emerged thus far for the support of co-located data analysis. However, there has been previous research in collaborative visualization in general. Most of this work has been applied to datasets and techniques from the scientific visualization community (e. g., for volume or flow analysis) and for distributed synchronous collaboration in specific environments such as CAVEs or for head-mounted displays. This focus is, for example, visible in the publication overview of the IEEE Visualization conferences and symposia presented in Figure 2.1. This chart

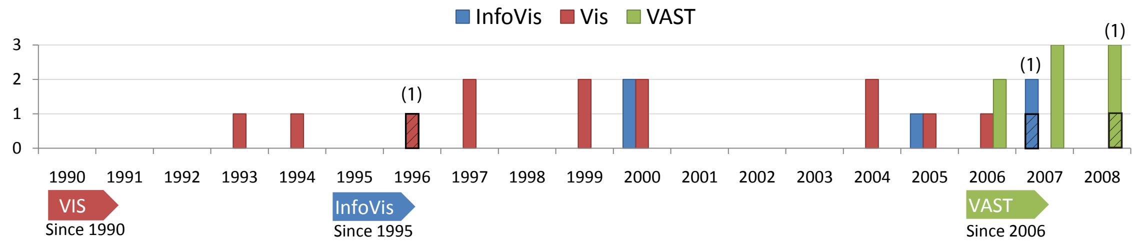


Figure 2.1: History of publications in Collaborative Visualization in the IEEE VIS, InfoVis, and VAST conferences. Out of a total of 1356 published papers, 26 are on Collaborative Visualization (shown here), and only three (hatched and indicated by numbers above the respective bars) covered co-located collaborative visualization.

includes only papers that directly describe collaborative visualization research. For example, I have not included papers that mention collaboration in future work or as work-in-progress, papers of technical solutions that could be used for collaboration but did not include example applications (e. g., large display architectures), spectator interfaces (for viewing but not manipulating visualizations), and ambient displays—unless collaboration was a specific concern in interaction or visual design. The figure shows the number of papers published in three major visualization venues: the IEEE Conference on Visualization (VIS), the IEEE Conference on Information Visualization (InfoVis), and the IEEE Symposium on Visual Analytics Science and Technology (VAST). Out of 1356 papers published in these three venues—VIS since 1990, InfoVis since 1995, VAST since 2006—26 papers focused on collaborative visualization and only three covered co-located collaboration. These three particular conferences were chosen as the top venues representing research interests of the larger visualization community.

Within the area of distributed collaborative visualization, one research focus has been on architectures and synchronization mechanisms for allowing efficient synchronous remote work with large scientific datasets (e. g., Ang et al. (1993); Li et al. (2006); Renambot et al. (2009); Wood et al. (1997)). Much of this research is focused on applications in virtual reality (VR) over the web (e. g., Ang et al. (1993)), in GRID comuting (e. g., Matsukara et al. (2004); Jankun-Kelly et al. (2003)), or for special hardware environments such as CAVEs (see Leigh et al. (1999)). Grimstead et al. (2005) provide an overview and taxonomy of 42 different distributed collaborative visualization approaches.

During the course of my PhD research, distributed web-based *information visualization* applications have emerged with a focus on making information visualization accessible to an internet-sized (mostly lay) audience (e. g., Danis et al. (2008); Dimov and Mulloy (2009); Paper (2009); Smallthought Systems (2009); Viégas et al. (2007)). With these systems, the research focus has shifted from the more technical aspects of network latency, synchronization, and view updates to more social, human-centered questions such as how wide audiences can be engaged to discuss and explore information, how laypeople can effectively share data and visualizations online, or how collaborative contributions can be effectively structured and integrated into a shared visualization to ignite further discussion and common ground formation (Heer et al., 2008).

My research is related to previous approaches in the scientific visualization community in that it involves the use of visualizations for collaborative work in specific hardware environments. In contrast to previous approaches in the scientific visualization, however, my hardware environments do not involve virtual reality technology but consist of single, large, and interactive displays such as tabletop or wall displays. Only a few projects have explored similar work contexts for data analysis and as the most related research, these will be discussed next.

While I have limited the presentation of the related work by the type of hardware and co-located collaborative work environment, my research relates to the collaborative web-based approaches in that I focus on human-centred questions for the support of collaborative data analysis activities. In contrast to much work in the Scientific Visualization community, I do not focus on the development of software and network architectures or synchronization protocols for supporting collaboration.

In the remainder of the chapter, I discuss systems related to co-located collaborative data analysis. The system features will be put in context to the research goal of this dissertation: the study of the tradeoff between private and shared work in collaborative information visualization.

2.2 Collaborative Scientific Visualization on a Shared Display

The 3D Responsive Workbench was an early digital tabletop information exploration system (Krüger et al., 1995; Krüger and Fröhlich, 1994) with the goal to replace computer desktops and provide a work situation more similar to those encountered in an architect's office, in surgery, or scientist research labs. It used a horizontal surface to display stereoscopic 3D information which was viewed through shuttered glasses. Input was facilitated through tracked gloves with gesture recognition. Speech recognition and audio feedback were provided as other possible input modalities.

Later papers focused on visualization applications (Wesche et al., 1997; Durbin et al., 1998) for the Responsive Workbench as well as on extensions to provide two people with independent input and corrected views of the 3D scene (Agrawala et al., 1997). Visualization in fluid dynamics was one of the presented applications (Wesche et al., 1997). Here, the viewer could inject particles and streamlines into a flow, control cutting planes, or select other visualization parameters. In this example, the main view of the scene was shared and input had to be negotiated between all group members. For a collaborative teaching scenario, the possibility of different views for two group members was discussed by Agrawala et al. (1997). The authors presented an example of a student and a teacher exploring a 3D model.Only the teacher could see the labels appear on the model and, hence, had a slightly different view of the scene. Each input in this example, still controlled the underlying view for both group members. In an air traffic visualization example (Agrawala et al., 1997), this problem was solved by partitioning the space into two distinct regions and each viewer was responsible for completing work only on their side. This way, problems of interference with a shared view were specifically avoided which allowed both people to work in parallel.

Around the same time as the Responsive Workbench was being developed in Germany, the Naval Research Laboratory in Washington, DC worked on a similar setup for their Virtual Workbench (Obeysekare et al., 1996). An almost identical setup was used with a stereoscopic display requiring the wearing of shuttered glasses, tracked gloves, and possible speech recognition and audio feedback (Figure 2.2). The Virtual Workbench, however, was never developed into a system that allowed independent viewpoints or mechanisms for several people to simultaneously interact with objects in the virtual

Figure 2.2: The Virtual Workbench. Image reprinted with permission from: *Obeysekare et al., Virtual Workbench–A Non-Immersive Virtual Environment for Visualizing and Interacting with 3D Objects for Scientific Visualization, Proc. of IEEE Visualization*, © 1996 IEEE.

environment. However, a few visualizations were introduced that were meant to be looked at and discussed by several people, including a real-time jet simulator, molecular docking, and flow visualization.

Both the responsive and virtual workbench have inspired more recent work on table-based virtual environments (e. g., Choi et al. (2005)) and multi-user stereoscopic displays (e. g., Kitamura et al. (2001)). The main benefit of these environments lies in their ability to display 3D information in a shared place where collaborators can come together, view the information, and explore, discuss, interpret, and analyze it. The ability to interact with the data, however, may be limited. Often interaction has to be negotiated as the whole view of a 3D model may be influenced when one collaborator decides to change certain parameters (e. g., the rotation, perspective, or view position). This effectively hinders exploration of the data in parallel making these environments difficult to navigate for work in which several team members may want to follow different exploration paths.

More recently a trend towards large horizontal and vertical touch-sensitive displays has emerged and several solutions have been proposed for these environments. With a shift to these types of displays, the question of synchronous, parallel interaction has also received greater attention, as discussed next.

2.3 Collaborative Information Visualization on a Shared Display

Besides my own collaborative information visualization systems (presented in later parts of this dissertation), there are relatively few other examples that have focused on joint data analysis with information visualizations. Next, I discuss systems that had specific data analysis components and involved information visualizations and/or information presentation techniques.

The Personal Digital Historian (PDH) project (Shen et al., 2002) displayed digital collections of photos, video, and text documents to groups of people with the goal to encourage conversation and storytelling about shared past histories. This system was built for casual use by families or friends and used visualization techniques to organize digital documents for exploration by these groups of people. PDH included several simple graphical techniques to organize documents on the table: fisheye distortion, timelines, spatial (map-based) layouts, and hierarchical layouts (Vernier et al., 2002). Fisheye distortion was implemented to enlarge documents in specific regions of the table, either on the inside (called the "central focus effect"), or towards the edge of the table (called the "central black hole effect"). Interaction techniques for opening and closing branches of a tree and their consequences for the hierarchical layout were discussed in more detail by Vernier et al. (2002). In particular, the problem of how to lay out child nodes without overlap when a certain branch is grown outward was discussed. The layout challenge was tackled here by decreasing sizes of nodes and placing them in a fan or circle around the parent node or by letting each group member define where child nodes should be displayed on the table. The first solution is a more standard approach also used within many other graph drawing techniques, while the latter suggests that in tabletop interfaces human-specified approaches may often be appropriate or necessary alternatives even if this results in a less clean data layout compared to the that of an automatic layout algorithm. The system only allowed one input point and people could not work in parallel and actively explore different parts of the data at the same time. Thus, collaborative work was limited to viewing and discussing the data but team members could not divide the work load and follow their own exploration paths.

The PDH system is one example of how simple representational techniques can be used to structure and lay out information on a tabletop display. The DTLens system (Forlines and Shen, 2005) demonstrates how focus+context presentation techniques can be used for the exploration of visualizations in forms of large maps and diagrams on an interactive tabletop. This system addressed an important issue for collaborative information explorations systems: when multiple people want to simultaneously interact with information that is spatially fixed (e. g., as in maps) the parallel exploration of information may be difficult to perform when interests between group members differ. DTLens addressed this problem by letting multiple people geometrically distort, annotate, and explore the visual information simultaneously. The system used a DiamondTouch (Dietz and Leigh, 2001) to provide identifiable input simultaneously, allowing up to four people to control lenses that enabled detailed views of information within a larger context. For example, a lens could be used to zoom into a map in a small portion of the display, while maintaining the context around that zoomed in area.Thus, each person could focus on a portion of interest, without hindering another person's ability to focus on something else within the same dataset. Interaction techniques with the lenses were designed to encourage rapid exploration.

A recent research direction has been the development of gestures for use with information visualizations on touch-sensitive displays. While this research typically does not focus on collaboration, it helps to inform the design of interaction techniques with data on large displays and is, therefore, related to the work of this dissertation. One such project, by North et al. (2009), looked at the problem of selecting and interacting with multiple information items on a multi-touch tabletop display. The researchers studied multi-object selection, grouping, and spreading tasks. In particular multi-object selection is an important interaction for information visualization tasks, as interactions such as filtering, clustering, highlighting, detail-on-demand, or zoom often involve several items. By studying both physical as well as digital contexts, the authors derived a multi-touch gesture vocabulary which can be used to inform future gesture design for direct-touch information visualization system.

While the project by North et al. (2009) looked at low-level interactions that can be common in many different types of information visualization tasks, the research by Frisch et al. (2009) looked at a more specific information visualization task: multi-touch and pen-based diagram or graph editing. By studying participants performing spontaneous gestures for operations such as creation, movement, and deletion of diagram elements, the researchers derived a set of gestures for node-link diagram editing. There is also a rich literature on gesture design for other task contexts that can be useful when considering gesture design for information visualization (e. g., Wu and Balakrishnan (2003); Wobbrock et al. (2009)).

Some projects have discussed the implications of perception on reading information visualizations on large displays. A study by Yost and North (2005) evaluated the scalability of information visualizations in regards to human perceptual capabilities. Their larger research goal was to address questions such as how much information humans can effectively perceive or whether visualization for large displays need to differ from those for desktop displays. Yost and North studied three visualizations across a small and large, high-resolution display and compared three tasks asking for detailed examination of data and four overview tasks. The authors conclude that the visualizations that were studied were perceptually scalable, as participants were—on a per attribute basis—faster in the large screen condition without a decrease in accuracy. While physical navigation along the large display was necessary, participants were able to integrate information from about 2.7m apart. While this study's focus was not on collaboration, it provided several guidelines for designing visualizations for large displays: considering encodings according to viewing angle, choosing visualizations for scalable encoding, providing global and local legends, and strategic label placement.

A study by Wigdor et al. (2007) evaluated the effect of perspective on the perception of graphical variables on a large display. The researchers tilted the display to achieve different viewing angles for participants and studied how well participants were able to perceive graphical variables such as length, position, angle, slope, or area. The study suggests that care should be taken in positioning and choosing the appropriate visual encoding as some graphical elements are more robust to view distortion than others. This study is also important to consider for collaborative work as group members might be positioned on different sides of the display, thus viewing shared visualizations from different directions. However, it still has to be further evaluated how the legibility of information visualizations is affected by different viewing directions and in context of more complex data displays (such as charts, parallel coordinates displays, etc.). So far, it is not known if, for example, turning 2D representations upside-down would lead to inaccurate readings of the data. Wigdor et al. (2007) also evaluated how well participants were able to perceive graphical variables across two adjacent displays. They conclude that visualizations should not be compared across display orientations (e. g., between a tabletop and a wall setup) and again, that certain visual variables are better suited for this comparison than others.

Multi-display setups are also currently being investigated for collaborative work environments, many including data analysis setups. The following section discusses some of these projects.

2.4 Collaborative Visualization in Multidisplay Environments

Tabletops have also been integrated in multi-display environments (MDEs) to support information exploration work. In these environments varying views of the same data can be shown on different displays around a room, effectively increasing the display space available for data exploration and comparison.

Forlines et al. showed two projects in which a tabletop serves a coordinating function in a setting with several vertical displays and a tablet PC. In the first project, Forlines et al. (2006) retrofitted Google Earth to support collaborative exploration of geospatial data.Several instances of Google Earth ran on connected displays, supporting different but coordinated viewpoints of the data on three wall displays. A tabletop display was used as the primary input device to coordinate all instances of Google Earth shown on these separate displays. The viewpoints could be changed in a coordinated manner, but information content could also be individually changed for each display. This made it possible to show different layers of information on each display (e. g., streets or airline routes). The table supported multiple independent input points so that each person could control different viewpoints at the same time. Independent work was facilitated by allowing the unlinking of views so that a single team member could explore parts of the information without affecting other views of the scene. Furthermore, when working alone, a group member could coordinate his or her interactions from a separate tablet PC in order not to disturb others by opening large menu dialogs and, thus, interfering with the global view on the tabletop. The system does not allow team members to synchronously interact with the same data visualizations.

In the second project (Forlines and Lilien, 2008), a single-user visualization application for protein visualization, JMol, was retrofitted to be used collaboratively in a multi-display environment.The environment consisted of a tabletop for controlling the other displays, several wall displays, and a tablet display. The table served as the central coordinating unit for selecting and changing views (e. g., viewpoints and representations) on the wall displays. The tablet was used to allow for fine-grained selection of small protein structures which would otherwise be difficult given the input resolution of the tabletop display. While several people could interact with the table and change views of the data, only one could do so at the same time. Parallel exploration of different parts of the data was not specifically supported.

The project WeSpace (Wigdor et al., 2009) presented a walk-up-and-use environment for collaborative research. The environment consisted of laptops that were brought in by participating group members, a tabletop, and a large wall display. Again, the tabletop served a coordinating function for views sent from the different laptops to the wall and tabletop display. The focus of this project was on allowing researchers to bring their own visualization applications to a joint discussion space rather than presenting them with a custom built visualization tool. The WeSpace tool itself consisted merely of a networking infrastructure to share views from clients installed on the laptops, a layout manager to control the organization of these views on the shared display, and LivOlay (Jiang et al., 2008), a tool to enable overlay and registration of different views of the same data on the shared display. In this setup, individual work is easily possible on each person's individual laptop, while shared viewing, discussion, and interpretation can happen on the shared wall display. However, since visualizations are typically controlled from a single laptop, parallel exploration of the same data is not as easily possible—unless researchers have previously shared the data and tools with each other.

2.5 Summary

Within collaborative visualization research, the challenges of distributed work have so far received greater attention than those of co-located analysis work. However, humans have considerable experience and expertise working together in shared environments, making this form factor a particularly promising one to investigate. In this chapter, I

Name	Max. Group Size	**Data**	**Independent Input**	**Independent Views**
Responsive Workbench	2	Scientific	Yes	Yes
Virtual Workbench	Small group	Scientific	No	No
Personal Digital Historian	4	Document collections	No	No
DTLens	4	Geospatial	Yes	Yes
Forlines et al. (2006)	Small group	Geospatial	Yes	Yes
Forlines and Lilien (2008)	Small group	Scientific	No	Yes
WeSpace	Small group	Group dependent	Yes	Yes

Table 2.1: Collaborative visualization systems discussed in Chapter 2.

introduced several systems that support collaborative data exploration or analysis on shared displays. A summary of the presented systems can be found in Table 2.1. With these systems, researchers have dealt with the problem of providing individual vs. joint work in several ways. Four systems (see Table 2.1) offered independent input and five offered independent views to each team member. With emerging display technologies such as multi-touch tabletop or wall displays, independent input for each group member becomes easier and cheaper to achieve without specific hardware devices. However, the question of how to deal with people's desire to explore parts of the information individually, has so far not received much attention outside of multi-display visualization environments. The DTLens system approached the problem by offering specific visualization tools (lenses) to provide team members individual, customizable views. This approach of allowing individual work with specific view of the same underlying dataset is closest in spirit to many of the research approaches taken in later parts of the thesis.

This chapter forms the first of a two-part literature review. In the next chapter, I review more literature from both research on computer-supported cooperative work as well as information visualization and analysis. The review elicits a first set of design considerations for collaborative information visualization environments for shared large displays.

Chapter 3

A First Set of Design Considerations for Collaborative Information Visualization

This chapter forms the second part of my literature review. The main contribution of this chapter is an analysis of relevant literature which I use to derive a first set of design considerations for co-located collaborative information visualization—drawn from a wide variety of literature sources. [1] The chapter focuses on the problem of supporting both parallel as well as joint collaborative data exploration and analysis in a shared space. This discussion of the literature is conducted with a particular purpose in mind; the intention is that this discussion will form the beginning of design considerations that will be modified and extended through further research in collaborative information visualization. My extensions are presented in the later chapters.

In this chapter, I discuss related literature from the area of Computer Supported Cooperative Work, Information Visualization, and empirical work that looks directly at

[1] Portions of this chapter were previously published in (Isenberg and Carpendale, 2007) © 2007 IEEE. Portions reprinted, with permission, from (1) *IEEE Transactions on Visualization and Computer Graphics, Interactive Tree Comparison for Co-located Collaborative Information Visualization, Petra Isenberg and Sheelagh Carpendale*; (2) (Heer et al., 2008) *Information Visualization-Human-Centered Issues and Perspectives, volume 4950 of LNCS State-of-the-Art Survey, 2008, Creation and collaboration: Engaging new audiences for information visualization, Jeffrey Heer, Frank van Ham, Sheelagh Carpendale, Chris Weaver, and Petra Isenberg. © Springer Verlag, 2008. With kind permission of Springer Science+Business Media.* Any use of "we" in this chapter refers to Petra Isenberg and Sheelagh Carpendale.

collaborative use of information visualization. I begin the overview of relevant literature by discussing research on mixed-focus collaboration or the trade-off between individual and shared collaborative work. Then, I organize the discussion of the related literature based on three aspects related to collaborative data analysis systems. This discussion is structured according to: Setting up a collaborative environment, supporting social interaction around data, and designing information visualizations for co-located collaboration.

3.1 Motivation

As previously discussed, current information visualizations have mostly been designed to support a single person working in a desktop environment. Thus, while most information visualization tools include mechanisms for sophisticated interaction with data, they have only limited facilities to support the collaborative activity of a team (Mark et al., 2003). Attempting to coordinate collaboration around a system designed to support only a single person, however, can be awkward and unnatural (e. g., (Stewart et al., 1999; Amershi and Ringel Morris, 2008)). Previous research on collaborative technology can shed light on questions and issues that need to be considered during the development of co-located collaborative information visualizations. Research within Computer-Supported Cooperative Work (CSCW) is worthwhile to consider since it has developed a number of considerations for systems designed to support co-located collaboration.

However, the requirements of collaborative data analysis tasks and the nature of work with information visualizations may pose particular domain-specific challenges. Work around information visualizations such as discovery and analysis tasks differs from other common collaborative work scenarios like design projects, photo sorting, or document editing in several ways. First of all, the outcome of an information analysis is not typically a product—such as a finished design, organized photo collection, or an edited document—but is often an intangible understanding or insight of the information that was analyzed. This has been cited as a challenge for the design and in particular the evaluation of information visualization systems (Plaisant, 2004). Thus, the support of collaboration around the less concrete work outcomes of data analysis tasks may require particular design considerations. Secondly, information visualizations have both an interaction component and a data representation component. Both of these components may need rethinking and redesigning to be effective in a collaborative work scenario and in non-desktop environments. Research in information visualization draws from the intellectual history of several traditions, including computer graphics, human-computer interaction, cognitive psychology, semiotics, graphic design, statistical graphics, cartography, and art (Munzner, 2002). Therefore, I take a closer look at this literature as well when attempting to generate a richer description of possible design considerations for co-located collaborative information visualization systems. Next, I first discuss literature related to the specific research goal of this dissertation: the study of the trade-off between individual and shared collaborative data analysis activities. Then, the remainder of the chapter discusses related literature on varying topics related to building a collaborative data analysis system for co-located collaboration. It is in this part of the chapter, that the first set of design considerations for this type of work is derived.

3.2 Supporting Mixed-Focus Collaboration

One goal of my research is to study the trade-off between private and shared work in collaborative information visualization. Therefore, I first introduce relevant related research that has discussed or studied this challenge in other contexts.

Many group activities, such as brainstorming or planning, involve phases of *mixed-focus collaboration* (Gutwin and Greenberg, 1998) in which group members transition from loosely coupled, parallel work to closely coupled, group work (e. g., Dourish and Bellotti (1992); Olson and Olson (2000)). It has also been noted, that the spatial partitioning and use of the workspace is influenced by different work phases in a collaborative setting (Scott et al., 2004). For example during parallel work, people make use of specific personal territories on shared displays while they tend to use other specific regions, group territories, when interacting more closely as a group.

For visualization systems the need to support both individual as well as group work has also been previously identified. A study by Park et al. (2000) in distributed CAVE environments, for example, discovered that when the visualization system supported an individual work style, participants preferred to work individually on at least parts of the problem. In collaborative visual analysis, for example, group members may need to be able to work on their own sub-projects, in which tentative hypotheses can be created, followed, and rejected. However, the analysts' desire for private work may be in tension with their desire to capitalize on the group's shared effort. The group may produce pieces of information that could be useful to the analyst; but the analyst, immersed in his or her work, may not want to be distracted (Brennan et al., 2006; Weaver, 2007).

This tension has been previously discussed for other collaborative scenarios. Gutwin and Greenberg (1998) suggest that task-dependent compromises and additional design work are necessary to balance both individual and group needs in distributed collaboration. Several techniques have been proposed in distributed collaborative visual analytics research to address this problem. Brennan et al. (2006), for instance, merge and fuse distinct private views on node-link graph representations, in order to show information overlap and common ground of graph nodes and information items explored and looked at by distributed collaborators. A similar idea, however using computational agents, was implemented in a distributed analysis system by Keel (2006). Here, the computational agents are used to identify when an individual had uncovered potential relationships between information items in his/her workspace; this insight is then automatically relayed to the larger group of collaborators. A more explicit sharing mechanism was implemented in CoMotion (Chuah and Roth, 2003). Here, objects and events can be explicitly shared by placing them in a shared view and implicitly annotated with interaction history information from different collaborators. These projects show the need to balance both group and individual contributions in a data analysis scenario, however, contributions from remote collaborators were the focus. The question arises how this challenge of trading off parallel and joint work in collaborative data analysis, can be addressed in co-located collaboration on a shared display.

Mixed-focus collaboration has recently been shown to apply to the work scenario that is the focus of this dissertation: synchronous co-located collaboration with information visualizations *over shared displays*. In the previously mentioned study by Tang et al. (2006), we[2] studied different types of lenses and filters to understand different types

[2] This paper was published as (Tang et al., 2006). Thus any use of "we" in this paragraph refers to Anthony Tang, Melanie Tory, Barry Po, Petra Isenberg, and Sheelagh Carpendale

of group cohesion. We noted that a trade-off is necessary between providing only a single or multiple independent instances of data views. With a single shared representation individuals' abilities to work independently may be compromised, yet using separate copied views may prevent many group collaborative dynamics from emerging. Our work focused on identifying different coupling styles and influences of tool use and we discussed the influence of the information visualization and the data analysis task in less detail. However, this and the above mentioned distributed projects have one main characteristic in common: they offer—to varying degrees—the possibility for individuals to work with their own views of the data to support parallel work styles. Another, less explored possibility involves making shared visual representations accessible for concurrent input. For example, one could imagine techniques that allow the parallel manipulation of several data characteristics, such as nodes in a tree, colour scales for a set of data items, or changing representations for parts of a dataset in focus. This latter possibility includes many possibilities for view conflicts, for example, when two collaborators want to manipulate the same data items. Design considerations for both of these strategies to support joint and parallel work are discussed next in the context of how to set up a collaborative environment, how to support social interaction around data, and how to design visualizations to support co-located collaborative work.

3.3 Setting Up a Collaborative Environment

The setup of a collaborative environment has an important impact on the possibilities for each individual in a collaborative work scenario. The physical characteristics of an environment influence who can interact with the visualizations, how well the visualizations can be seen, who can best see the visualizations, how the group members are positioned relative to each other, and how well they can discuss the data with each other. Therefore, the main factors that influence these aspects of collaborative work are discussed next with their impact on individual and group work.

3.3.1 Display Size

In collaborative systems, screen space has not only to be large enough to display the required visual representation(s), it also has to be viewed and *shared* by several people. When people would like to work in parallel, independently of one another, they may want to move parts of a visualization off to the side and work on it without disturbing others. In particular in this case, the size of the screen is critical in supporting people's desire for a private work area (Scott et al., 2004). In addition, as the number of people using a shared information display grows, the size of the display and workspace needs to be increased in order to provide a large enough viewing and interaction area that gives adequate access to all group members.

3.3.2 Display Configuration

Several configuration possibilities exist that could increase the amount of available display space, all of which will affect the type of visualization systems that are possible and the type of collaboration work that would be most readily supported. For instance, one could provide team members with interconnected individual displays, as in the ConnecTable system (Tandler et al., 2001), or one could make use of large, interactive, single-display technology, like display walls or interactive tabletop displays (e. g., Stewart et al. (1999); Tang et al. (2006)). An additional possibility is to link wall, table, and personal displays (e. g., Wigdor et al. (2007)), or to consider immersive displays (e. g., Krüger et al. (1995); Obeysekare et al. (1996)). The type of setup most appropriate for an information visualization system will depend on the specific task and group setup. For example, individual interconnected displays allow for private views of at least parts of the data which might be required if data access is restricted or the need for parallel work is particularly high. In addition, private displays can be used for interactive operations that have specific access restrictions. Tabletop displays have been found to encourage group members to work together in more cohesive ways, whereas wall displays are beneficial if information has to be discussed with a larger group of people (Rogers and Lindley, 2004). It has been shown that people are able to coordinate both parallel as well as joint work on tabletops for data analysis but that tools and visual design strongly influence what type of work style people tend to adopt (Tang et al., 2006).

3.3.3 Input Type

In the common desktop setup, input is provided for one person through one keyboard and one mouse. To support collaboration, ideally, each person would have at least one means of input (e. g., Stewart et al. (1999); Amershi and Ringel Morris (2008)). In addition, it would be helpful if this input was identifiable, making it possible to personalize system responses. If a collaborative system supports multiple input points, at least one per person, it has to be coordinated how all team members can access the shared visualization and data sets. For example, synchronous interactions on a single representation may require the design and implementation of new types of multi-focus visualizations. Ryall et al. (2006) have examined the problem of personalization of parameter changes for widget design, allowing widgets to be dynamically adapted for individuals within a group. Similar ideas could be implemented for personalization of information visualizations during collaborative work. This design problem is discussed further in Section 3.5.2.

3.3.4 Display and Input Resolution

Resolution is an issue both for the output (the display) and for the input. The display resolution has a great influence on the legibility of information visualizations and the amount of data that can be displayed. Large display technology currently often suffers from relatively low display resolution so that visualizations might have to be re-designed so that readability of text, colour, and size are not affected by display resolution. Also, large interactive displays are often operated using fingers or pens which have a rather low input resolution. Since information visualizations often display large datasets with many relatively small items, the question of how to select these small items using low input resolution techniques becomes an additional challenge that needs special attention (Isenberg et al., 2006b; Voida et al., 2009). These considerations are important for both parallel and joint work styles.

3.4 Supporting Social Interaction Around Data

Pinelle et al. (2003) provide a set of basic operations that should be supported by groupware systems to help collaborators carry out their tasks as a team. These *mechanics of collaboration* can be grouped into those describing communication and those describing coordination aspects of collaboration. Collaborative information visualization systems, like other groupware systems, require support for both—communication and coordination—to support social interaction among team members, in particular, when they switch between phases of individual and joint work. Further research relating to issues of shared interactions in collaborative information visualization scenarios is discussed next.

3.4.1 Supporting Communication

Communication is an important part of successful collaborations. People need to be able to trigger conversations, communicate their intentions, indicate a need to share a visualization, and to be generally aware of their team members' actions. Co-located synchronous work has a number of characteristics for specific communication support (Olson and Olson, 2000): a shared local context in which participants can interact with work objects, rapid visual and audio feedback, multiple channel information exchange with voice, gestures, etc., and visibility of others' actions. However, the nature of collaborative work with information visualizations impacts team members' ability to communicate. These impacting factors are discussed next, in relation to explicit and implicit communication in a workspace (Pinelle et al., 2003).

Explicit Communication

The possibility for direct exchange of information through many channels such as voice, gestures, and deictic references is one advantage of co-located collaboration (Olson and Olson, 2000). The ease of referencing items in a shared space by simply pointing to them—often combined with verbal alouds—can improve communication about shared items in the workspace (Olson and Olson, 2000). However, this ease of reference to joint objects can be limited in situations where group members are working in parallel.

In these situations, group members may be working either with very different parts of the data, different data altogether, or may be viewing the same data but using different representations (e. g., Figure 3.1). Although being in a shared space, here the reference to an information item may be difficult because—in particular, when visualizations are large and complex—the context may not be immediately understandable and transferable to another view of the data. For example in Figure 3.1, we see two people working with two representations of the same data. Person A is pointing at an information item to initiate a conversation about it with Person B on the other side of the table. B now has the difficult task of finding the data item in his or her working context, or has to switch to the other side of the table and work more closely together with A and, thus, abandon his or her current work context altogether. The problem here lies in the registration of an item from one information space to the next. The team member has to perform a mental navigation task to the new information space, a problem which has been well identified in the literature (e. g., Spence (1999)). The problem of mental navigation in complex information spaces—often represented by information visualizations—lies in the fact that simple spatial referencing is not easily possible. In visualization environments each interacting person has to navigate two types of spaces. First, there is the physical interaction in the workspace. Secondly, there is interaction within the data space. Here the interaction may involve zooming into data, changing spatial variables, or selecting certain information items. The mapping of the physical to the data space may be different for every visualization or every different view in a visualization, limiting the way in which items can be referred to spatially.

While, this problem of reference exists, visual representations can be designed to support explicit communication across views. For example, meta-visual overlays can be designed that identify an information item that is pointed to in one view, in all other views of the same data. Also, it has been shown that the ability to annotate data and share insights in a written way is an essential part of the discovery process in distributed information visualization settings (Heer et al., 2007). Annotation is a form of explicit communication, either with oneself at a later point in time, or with other collaborators looking at the same data. In digital systems, annotations of data items with messages of all types, written, voice, etc. could further support communicative needs of groups, in particular in phases of parallel work.

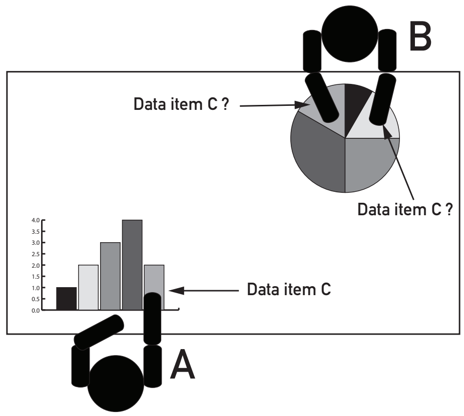


Figure 3.1: Two collaborators working in parallel with different representations of the same data. A is pointing to a data item while Person B on the top is trying to establish which item the other is referring to.

Implicit communication

In co-located non-digital collaboration, people are accustomed to gathering implicit cues about team members' activities through such things as body language, verbal alouds, or by hearing items being moved in the workspace (Olson and Olson, 2000). This is an active research area in distributed collaboration since the co-located evidence does not naturally become distributed. While co-located collaboration benefits from many of the co-present advantages, there are still issues that arise. Some examples include digital actions that are not always readily visible (cursors are hard to see on large screens), menu actions that can affect a remote part of the screen, the difficulty of identifying data items in different views, as well as the general problem of change awareness (Rensink, 2005). Thus, while implicit communications are present and potentially, at least to some extent, noticeable in co-located settings, some system changes made by a collaborator can still remain unnoticed if the collaborative system does not provide appropriate feedthrough (i. e., a reflection of one person's actions on another person's view) (Pinelle et al., 2003).

In collaborative information visualization, for example, it might be important to consider appropriate awareness for operations that make changes to the underlying dataset. Imagine a co-located system in which each collaborator works in parallel on a different view using a different file system representation. If one collaborator discovers an old version of a file and decides to delete it, this change might go unnoticed if the other person is looking at a view of the data that does not include the particular file or it might be completely surprising to the other person to see a file in their representation disappear. Some research has proposed policies to restrict certain members from making unsuspected global changes to a dataset (Ringel Morris et al., 2004). Earlier research on information visualization discussed the differences between view and value operators (e. g., (Chi and Riedl, 1998)). View operations make changes to the view of a dataset only, while value operations make changes to the underlying dataset itself. Most recent research in multiple-view visualization tends to favour view operations (filtering of unwanted data rather than deletion). This seems likely to be most appropriate during collaboration as well. However, when value operations are required during an analysis, appropriate awareness mechanisms have to be implemented. In a system by Tobiasz et al. (2009), we[3] explored this issue in more depth. We proposed a meta-visualization between individual views of the data which allows collaborators working in parallel to remain aware of the scope of their local interactions on other views of the data.

It has also been shown that the location and orientation of artifacts are used to support implicit communication in non-digital settings. They can, for example, suggest *who* is working with an artifact and communicate the intent of one team member to hand-over or pass on an artifact (Kruger et al., 2004). It has also been shown that this translates to digital settings (Kruger et al., 2005). Thus, an important design consideration is the support of artifact mobility and freedom of orientation, in particular when parallel work is supported by giving people different views of the same data.

3.4.2 Supporting Coordination

In group settings, collaborators have to coordinate their actions with each other. Coordination involves activities such as the transfer of resources in the workspace, protecting one's work, or storing items in the workspace (Pinelle et al., 2003).

Typical information visualization systems for individual use, however, impose a fixed layout of windows and controls in the workspace. In contrast, previous research has shown that, on shared workspaces, collaborators tend to divide their work areas into personal, group, and storage territories (Scott et al., 2004). This finding implies that

[3] This paper was published as (Tobiasz et al., 2009). Thus any use of we in this paragraph, refers to Matthew Tobiasz, Petra Isenberg, and Sheelagh Carpendale.

a group interaction and viewing space may be beneficial in collaborative data analysis, in particular when groups need to be supported in transitioning between individual and shared work. In a shared group space, all collaborators can work on a shared representation of the data or they can share tools and representations. In a personal space, individual team members may more commonly explore the data separately from others. Flexible workspace organization can offer the benefit of easy sharing, gathering, and passing of representations to other collaborators. If visualizations can be easily shared, team members with different skill sets can share their opinions about data views, suggest different interpretations, or show different venues for discovery. By offering mechanisms to easily rotate and move objects, aspects such as comprehension, communication, and coordination can be further supported (Kruger et al., 2004). Rotation can, in particular, support coordination by establishing ownership and categorizations. By allowing free repositioning and re-orientation, we can also make use of humans' spatial cognition and spatial memory and possibly better support information selection, extraction, and retrieval tasks. Free arrangements of representations in the workspace, thus, can support changing work styles. Representations can be fluidly dragged into personal work areas for individual or parallel work and into a group space for closer collaboration. Mechanisms for transfer of and access to information visualization in the workspace should be designed in a way that they respect common social work protocols (Kruger et al., 2004; Scott et al., 2004).

3.5 Designing Information Visualizations for Co-located Collaboration

Even though they were proposed with an individual use in mind, many known information visualization guidelines still apply to the design of information visualizations for co-located collaborative use (e. g., Bertin (1983); Tufte (2001); Ware (2000)). This section discusses changes and additions to factors that need to be considered when designing information visualizations for co-located collaborative settings. Thus, much of this discussion simply delineates research questions that may be of specific interest when designing information visualizations to support co-located collaboration.

3.5.1 Representation Issues

Spence (2007a) defines representation as "the manner in which data is encoded," simplifying the definition of representation as a formal system or mapping by which data can be specified (Marr, 1982). The concept of representation is core to information visualization since changes in representations cause changes in which types of tasks are most readily supported. As in Marr's example (Marr, 1982), the concept of thirty-four can be represented in many ways. To look at three of them; Arabic numerals, 34, ease tasks related to powers of ten; Roman numerals, XXXIV, simplify addition and subtraction; and a binary representation, 100010, simplifies tasks related to powers of two. Not surprisingly, Zhang and Norman (1994) found that providing different representations of the same information to individuals provides different task efficiencies, task complexities, and changes decision-making strategies. Questions arise as to what are the most effective representations during collaboration. Will certain representations be better suited to support small group discussions and decision making? Will multiple representations be more important to support different people's interpretation processes? Will new encodings or representations be needed for collaborative work scenarios? Appropriate representations might have to be chosen and adapted depending on the chosen display type; whether completely new designs are required is not yet clear. For example, spatiality or the use of position/location is commonly an important aspect of representation semantics. However, spatiality as manifested in territoriality is a significant factor for communication and coordination of small group collaboration (Scott et al., 2004). It is an open question as to whether there is a trade-off between these two uses of spatiality.

In addition, different representations may have to be accessible in an interface because, in a collaborative situation, group members might have different preferences or conventions that favour different types of representations both in parallel and joint work. Findings suggest that the availability of multiple, interactively accessible representations might be important for information visualization applications since the availability of multiple data representation can change decision making strategies (Kleinmutz and Schkade, 1993). Also differing representations have an influence on validation processes in information analysis (Saraiya et al., 2005), and more easily support people working in parallel on information tasks (Park et al., 2000). Gutwin and Greenberg (1998) have discussed how different representations of the workspace affect group work in a distributed setting. They point out that providing multiple representations can aid the individual but can also restrict how the group can communicate about the objects in the workspace. This extends to co-located settings, in which several representations of a dataset can be personalized according to taste or convention, making it harder to relate individual data items in one representation to a specific data item in another (see Section 3.4.1 and Figure 3.1). Implementing mechanisms to highlight individual data items across representations might aid individuals when switching between joint and parallel data exploration.

3.5.2 Presentation Issues

Presentation has been defined as “something set forth for the attention of the mind” (Webster, 2007) and as “the way in which suitably encoded data is laid out within available display space and time” (Spence, 2007a). From these definitions it is clear that changing from desktop to other display configurations, as is usually the case to support co-located collaboration, will impact the types of presentations techniques that are possible and/or appropriate. In collaborative scenarios, information visualizations might have to cover larger areas as group members might prefer to work in a socially acceptable distance from each other. The display space might also have to be big enough to display several copies of one representation if team members want to work in parallel. These copies should be movable, resizable, and reorientable to allow group members to position them according to their preferences. However, the size and aspect ratio of a visualization can have an important influence on the interpretation of the information (Heer and Agrawala, 2006). This can be both beneficial as it provides new perspectives on the visualized information but also be problematic as pointed out in frequent examples of ‘data lies’ (Huff, 1954).

Assuming a large enough display space and multi-touch input capabilities—if groups are working over a shared presentation of data, presentations might have to be adapted to allow collaborators to drill down and explore different parts of the data in parallel. Collaborative information visualizations will likely have to support multiple simultaneous state changes. This poses additional problems of information context. Team members might want to explore different parts of a dataset and place different foci if the dataset is large and parts of the display have to be filtered out. Information presentations might have to be changed to allow for multi-focus exploration that does not interfere with the needs of more than one collaborator. For example, DOI Trees (Card and Nation, 2002) or hyperbolic trees (Lamping et al., 1995) are examples of tree visualizations in which only one focus on the visualization is currently possible. ArcTrees (Neumann et al., 2005) and TreeJuxtaposer (Munzner et al., 2003), for example, allow for multiple foci over one tree display but these were not designed to take the information needs of multiple collaborators into account and might still occlude valuable information.

An example for visualization presentation changes based on a collaborative circular tabletop environment has been presented by Vernier et al. (2002) (also see Section 2.3). The presentation of the circular node-link tree layout was modified to rotate all nodes towards the boundary and a 'magnet' was implemented to rotate nodes towards just one team member. Nodes were also changed in size; as leaf nodes were placed closer towards team members' personal spaces (Scott et al., 2004) they were decreased in size and the nodes towards the center of the table were enlarged to allow for easier shared analysis of the node contents in the group space (Scott et al., 2004). A possible extension of this work is to think about placing and re-arranging nodes automatically based on the placement and discovery interests of team members or based on the individual or shared discoveries that have been made.

The presentation of visualizations might also have to take available input devices on a shared large display into account. If fingers or pens are used as input devices, the selection might not be accurate enough to select small information items. A common task in information visualization is to re-arrange data items (e. g., by placing points of interest), to request meta-information (Shneiderman, 1996) (e. g., by selecting an item), or to change display parameters by selecting an item. If the displayed dataset is large, it often covers the full screen and reduces individual items to a few pixels. Previous research has attempted to solve the issue of precise input for multi-touch screens (e. g., Isenberg et al. (2006b); Benko et al. (2006)) but they might not be applicable if the whole visual display is covered with items that can possibly be selected. Alternatively, information presentations could be changed to allow for easier re-arrangement and selection of items, for example, with lenses as we[4] explored with the iLoupe and iPod-Loupe (Voida et al., 2009). DTLens (Forlines and Shen, 2005), discussed in Chapter 2, is another example of the use of lenses for co-located collaboration.

The resolution of a large display also has an influence on presentation design as it relates to the legibility of data items. It is known that the reading of certain visual variables is dependent on the size and resolution in which they are displayed (Ware, 2000). Information visualizations also often rely on textual labels to identify data items which may be hard to read on low-resolution displays. The presentation size of individual items and labels may have to be adapted to compensate for display resolution. Both the selection of small data items and the readability of the data on the display are issues for both parallel and joint work phases.

3.5.3 View Issues

The term 'view' is common in information visualization literature and view operations (changing what one currently sees) have been defined as distinct from value operations (changing the underlying data) (Chi and Riedl, 1998). However, this use of the term view also incorporated changes in visual aspects of representation and presentation. Blurring the distinction between view and presentation changes has not been problematic because with a single viewer and a single display these are often concurrent. This distinction rises in importance in co-located collaborative applications. A change of view, for example, can be the result of a person moving physically to another location. This would be quite rare in single person desktop setups but can be quite common in large and/or multi-display environments. It can also be the result of view operations such as pan, zoom, rotation, or re-location, which may be commonly performed when team members have freedom to organize their views in the workspace.

As discussed next, view issues impact both parallel and joint work on visualizations. In a co-located collaborative setting, of necessity, there are as many views of a given presentation as there are people in the group. Also, since collaboration practices often include mobility, a given person's view will change as they move in the physical setup. This factor has recently begun to receive attention in the CSCW community. Nacenta

[4] This paper was published as (Voida et al., 2009). Thus, the use of "we" in this sentence refers to Stephen Voida, Julie Stromer, Matthew Tobiasz, Petra Isenberg, and Sheelagh Carpendale

et al. (2007b) have shown that righting (orienting a piece of 2D information into the proper perspective) by means of motion tracking aids comprehension. Hancock and Carpendale (2007) consider the same problem for horizontal displays looking for non-intrusive interactive solutions. Since a study by Wigdor et al. (2007) has indicated that angle of viewing affects readability of certain visual variables; this issue will be an important one for collaborative information visualizations. This research on how view-angle distortion affects perception in a single and multi-display environment suggests that certain types of representations may need to be modified in order to be used on a digital tabletop display and that information visualizations should not be compared across multiple display orientations. However, as visual variables were tested in isolation (e. g., length and direction only), further evaluations have to be conducted to see whether participants will correct for possible distortion if the variables are presented in conjunction with others or whether view correction (Nacenta et al., 2007b; Hancock and Carpendale, 2007; Hancock et al., 2009) might compensate.

Visualizations that can be read from multiple angles and orientations (e. g., circular tree layouts vs. top-down layouts) might be more appropriate for display on a horizontal surface. However, it is not clear whether participants would try to read oriented visualizations upside down and make wrong conclusions based on these readings or whether they would simply re-orient the visualization to correct the layout.

3.5.4 Interaction Issues

Most interaction issues deal with interaction with representations, presentations and views, thus discussing them here would overlap with points raised under these headings. However, there are some more general interaction issues. These include issues with how collaborative systems may impact interactive response rates and the fact that several inputs are in play.

Interactive Response Rates

Information visualization often deals with extremely large and complex data sets and can have considerable graphics requirements for complex representations. Adding larger screens, more screens, higher pixel counts, and multiple simultaneous inputs will increase computational load adding more requirements to the challenge of maintaining good interactive rates. Thus implementations of collaborative information visualizations have to be carefully designed for efficiency. Individual information displays can already be computationally intensive and require considerable pre-processing. Yet, in collaborative systems several information visualizations might have to be displayed and interacted with at the same time. While powerful hardware can solve the problem to some extent, efficient data processing as well as fast rendering of the graphical representations will be important issues to be addressed. This is particularly crucial when parallel work styles require simultaneous changes to visualization states.

Interaction History

Collaborative information visualization systems should also consider providing access to some form of data analysis history. While this is true for information visualizations in general (Shneiderman, 1996), it might be of even higher importance in collaborative settings. Chuah and Roth (2003) have suggested that capturing and visualizing information about interactions of collaborators with objects in a workspace may enhance collaboration by leading to a better understanding of each others' involvement in solving a task. As group members switch between work on individual and shared views of the data, they might lose track of the interactions of their collaborators (Gutwin and Greenberg, 1998). The access to an exploration history can help in later discussing the data and exploration results with collaborators or informing them about interesting data aspects that have been found during the analysis process.

Information Access

Access to data through information visualizations also needs to be coordinated on a global and local scope. What if, during parallel work, one group member found something in the data that he or she wishes to delete or modify? Who can change the scale, zoom, or rotation settings for a shared view of the data? Policies might have to be put in place to restrict certain members from making unsuspected global changes to the data or representations that might change other group members' view of the same data

(Ringel Morris et al., 2004). Similar issues pertaining to workspace awareness (individual vs. shared views), artifact manipulation (who can make which changes), and view representation have been discussed by Gutwin and Greenberg (1998). Should a system allow for multiple representations or force collaborators to work over a shared representation of the data? Should the exploration on multiple representations of the same dataset be linked or be completely independent? Further studies are required to arrive at answers to these questions.

Fluid Interaction

Collaborative systems should support fluid transitions between activities to improve the coordination of activities (Scott et al., 2003). The fluidity of interactions in a shared workspace influences how much collaborators can focus on their task rather than on the manipulation of interface items (Scott et al., 2003). This implies that in a collaborative information analysis scenario, parameter changes to the presentation or representation of a dataset should require the manipulation of only few interface widgets (menus, slider, etc.) and little or no changes of input modalities (mouse, keyboard, pen, etc.). A study on collaborative information visualization systems has similarly reported that groups worked more effectively with a system in which the required interactions were easier to understand (Mark and Kobsa, 2005). This poses a challenge to information visualization tool designers as typically a high number of parameters are required in visualization systems to adapt to the variability in dataset complexity, size, and tasks. The availability of mechanisms to quickly and easily pass, obtain, and share information visualizations is invaluable for collaborative work around information visualization.

3.6 Summary

This chapter provides a first collection of considerations for the design of collaborative information visualization systems based on a review of the related literature. It shows that one has to consider many factors including: the needs for groups when designing the collaborative environments; the impact that social factors such as communication and coordination, have on data analysis; the importance of thinking about group needs when designing effective representations, presentations, and views of the data; and the needs of group interaction support as well as information interactions that consider impact on the group. I have given examples and shown how related work in several areas can be synthesized to inform the design of co-located collaborative information visualization systems for shared workspaces. The HCI and CSCW literature has high applicability when considering the interaction, coordination, and communication component in information visualizations, whereas research on information visualization and perception has a higher applicability when designing changes in representations, presentations, and views based on group working requirements.

I discussed scenarios in which groups need support for both individual as well as joint work. Parallel work can be supported by providing individuals with their own views of the data to support parallel work styles or by making shared visual representations accessible for concurrent input. Joint work can be supported by providing additional communication and coordination mechanisms in which analysis results, views, and visualizations can be shared and discussed together. This literature review has also shown that there is very little research focused on these issues in collaborative visualization. While Chapter 2 has discussed some previous systems developed for this domain, little work has emerged on how people work together with visualizations and how they collaboratively make use of visualizations during collaborative data analysis. Next, Chapter 4 discusses a first study on collaborative data analysis activities in a co-located setting with a focus on studying teams' joint data analysis processes. Table 3.1 gives a final overview of the design considerations discussed in this chapter.

Consideration	**Aspects to Consider**
	Collaborative Environment
Display size	Socially appropriate work space size per person, establishment of private, group, and storage spaces
Display configuration	Accommodation of group's current work practices, tasks, and goals
Input Type	Impact of input type on possible interactions
Resolution	Input and display resolution
	Supporting Social Interaction
Communication	Explicit data referencing across different representational and viewing contexts, e. g., annotation; implicit awareness cues of changes to the data across different representational and viewing contexts
Coordination	When using individual data views, location and rotation as a coordination and communication tool, sharing of visualizations and views, multiple synchronous interactions with shared representations
	Designing Information Visualizations
Representation	Personal preferences, multiple representation types, awareness support, appropriateness of representation for work environment and social interaction
Presentation	Arrangement of data items for group access, providing copies of the same data, accommodation of input methods, compensations for display resolution
View	Interpretability of data from multiple viewpoints and orientations
Interaction	Interactive response rates despite simultaneous interaction, collaborative interaction histories, conflict reduction arising from global changes to data or view, fluid interaction

Table 3.1: Summary of first design considerations for co-located collaborative data analysis environments.

Chapter 4

Collaborative Visual Information Analysis Processes

The review of the literature, as discussed in Chapter 3, showed that we still know too little about how teams work with information visualizations, what kinds of analysis processes software should support, and which analysis processes require specific support for parallel vs. joint work styles. To improve our understanding of collaborative data analysis, we[1] conducted an exploratory study of groups of individuals, pairs, and triples engaged in information analysis tasks using paper-based visualizations. From the results of our study, we derive information processes that capture the analysis activities of co-located teams and individuals. Comparing this framework with existing models of the information analysis process suggests that information visualization tools may benefit from providing a flexible temporal flow of analysis actions and that certain analysis activities are more commonly performed closely coupled and some more commonly in loosely coupled collaboration. These findings enrich our understanding of the considerations in designing information visualization systems for co-located collaborative data analysis.

[1] Main portions of this chapter were published in Isenberg et al. (2008a). Thus any use of “we” in this chapter refers to Petra Isenberg, Anthony Tang, and Sheelagh Carpendale

4.1 Motivation

Many researchers have explored the information analysis process (e. g. Card et al. (1999); Jankun-Kelly et al. (2007); Spence (1999)) but little has emerged on the nature of this process in a collaborative context (Mark and Kobsa, 2005; Park et al., 2000). How a single doctor would analyze biomedical visualizations, for example, might differ from how a team of doctors might analyze the same data. If teams make use of visual information to solve problems differently than individuals, we need to understand what these differences are so we can redesign or create new information visualization tools to support their activity. To address this problem, we designed an exploratory study to understand the flow and nature of this collaborative process and its relationship to data analysis practices. To derive practical considerations for information visualization tool design, we focused our analysis on how team members engage with the workspace and their collaborators. Teams in the study were given paper-based visualizations to solve tasks, allowing us to view their process independently of the confounds of a specific information visualization system.

4.2 Collaborative Visual Information Processing

With this study, we examine how individuals and teams solve information tasks using simple visual representations of their data. The study results in a description of information analysis processes and, thus, it particularly relates to previous studies that have also resulted in descriptions or information processing frameworks. To provide context for our study, in this section, we outline previous research that articulates an *information visualization process* or the process through which a person extracts insight from a dataset given a problem and visualization tool.

We reserve the detailed comparison of these processing frameworks until after the description of our study and our study results (see Section 4.6). This allows us to compare our results to these existing frameworks. In particular, we will compare both the study by Park et al. (2000) of pairs using distributed CAVE environments, and the study by Mark and Kobsa (2005) of pairs sharing an information visualization software tool that had been designed for single person use. These studies resulted in similar but not identical information processing frameworks. These two studies are most related, but the results of the study presented here differs in that by studying non-digital information processing, the results do not reflect the processing constraints built into existing software but instead reflect how participants would work in non-digital contexts.

Several researchers have modeled an individual's involvement in visual information processing as an iterative sequence of components; however, each model is unique in terms of its focus, and how it abstracts the process. One perspective has been concerned specifically with the design of digital information visualization tools, focusing on how a person manipulates view and visualization transformation parameters (e. g., (Chi and Riedl, 1998; Jankun-Kelly et al., 2007)). Jankun-Kelly et al. (2007) propose a model of visual exploration for analyzing one person's interaction with a digital visualization system. A core proposition of this work is that a fundamental operation in the visual exploration process is the manipulation of visualization parameters. This model is effective in capturing the temporal aspects of visual parameter manipulation; however, it does not capture the higher-level semantics of a person's interaction (i. e., why was a parameter changed?). Chi and Riedl (1998) address this aspect, basing their semantic operator framework on a person's intention of action (i. e., view filtering vs. value filtering), classifying and organizing operators in the analysis process. At the other end of the spectrum, Amar and Stasko (2005) name higher-level analytic activities that a person using a visualization system would typically perform, such as complex decision-making, learning a domain, identifying the nature of trends, and predicting the future. Shneiderman (1996) outlines a process ("overview then detail"), that addresses a task-centric perspective on the analysis process. He suggests seven different operations that information visualization tools should support to facilitate the problem solving process: overview, zoom, filter, details-on-demand, relate, history, and extract.

A model by Russell et al. (1993), derived from studying collaborative information consolidation activities, describes a "Learning Loop Complex," a cyclic process of searching for representations and encoding information. Indirectly, these observations have led to the *Sensemaking Cycle* by Card et al. (1999) and an extension presented in (Thomas and Cook, 2005). We will later revisit the Sensemaking Cycle by Card et al. (1999) as it shares some processes defined in our framework.

In this study, we are interested in the *general processes* that occur during collaborative information analysis (independent of the confines of a computer-based information visualization tool), as well as the interactions with visualizations and those between team members. The focus is on general processes that form the basis of *collaborative* use of information visualization.

4.3 Choosing a Methodology

When developing software tools to augment work practices, at least three fundamentally different approaches exist. One is to study possible improvements for support of the process through studying the current software support or tools in use. Another is to hypothesize about improvements to existing tools, to develop a promising tool and study it in comparison to the existing tools. A third is to work towards an improved understanding of the process in order to develop a better match between the natural human process and its software support.

The study presented in this chapter is part of this third stream of research and as such falls into the tradition of qualitative research (Creswell, 1998). The basic idea is that through observations of participants' interactions with physical artifacts a richer understanding of basic activities can be gained and that this understanding can be used to inform interface design. Researchers (e. g., Scott et al. (2004); Tang (1991)) have previously studied how people accomplish tasks in *non-digital contexts* in order to understand what activities *digital tools* should support. Their approach generally relied on observation of participants, inductive derivation of hypotheses via iterative data collection, analysis, and provisional verification (Creswell, 1998). This style of research has worked well to uncover the basic activities of collaborative work. For instance, Tang's study of group design activities around shared tabletop workspaces revealed the importance of gestures and the workspace itself in mediating and coordinating collaborative work (Tang, 1991). Similarly, (Scott et al., 2004) focused on the use of tabletop space and sharing of items on the table and showed how people established and used different tabletop territories during collaborative work. While these authors studied traditional, physical contexts, ultimately their goal was to understand how to design *digital* tabletop tools. Both of these studies contributed to a better understanding of collaborative work practices involving tables in general. Of particular interest is that in these studies, the researchers chose *not* to use digital tools, and instead to study the participants using traditional artifacts, such as pens, paper, cardboard, and so forth. The reasoning behind this choice is that participants' *physical interactions* with these familiar artifacts and tools would more closely reflect how participants *understand and think* about the problem at hand. Similarly, with the study presented here, we shed light on small team visual analytic processes by observing people working with familiar physical artifacts on a traditional table in order to avoid observing the processes supported by any given piece of visual information software.

4.4 A Study of the Information Analysis Process

In the exploratory study presented next, we focused on examining individuals and small groups working on visual analysis tasks unencumbered by the confounds of any specific digital interface. We developed a set of static charts placed on index cards to represent the visualization tool, and provided participants with traditional tools such as pens and paper. This setup allowed us to observe behaviours such as free arrangement of data, annotation practices, and different ways of working with individual information artifacts—behaviours that we would not otherwise see given most digital visualization tools.

A key drawback of this approach is that it is not possible for us to see how typical interactions in information visualization tools (such as selection, encoding, or presentation parameter manipulations) would be used; however, like Mark and Kobsa (2005), our specific interest was in uncovering the *general processes* involved in collaborative and individual visual analysis, and not on specific interactions with a given visualization tool.

4.4.1 Participants

We recruited 24 paid participants from the University of Calgary population, 14 female, 10 male. The mean age of the participants was 26 years. Participants were assigned to 4 groups each of singles, pairs, and triples. With one exception, all pairs and triples were known to each other beforehand. For further group details refer to Figure 4.1. The sample size for this study was informed by emerging results. After four pilot studies and 12 groups the experimenters were confident that further observations would result in redundant data. 17 of the 24 participants reported to be familiar with all the charts given to them in the study (Figure 4.1). 21 of the participants reported to do data analysis similarly to how it was asked of them in the study at least on a yearly basis (Figure 4.1).

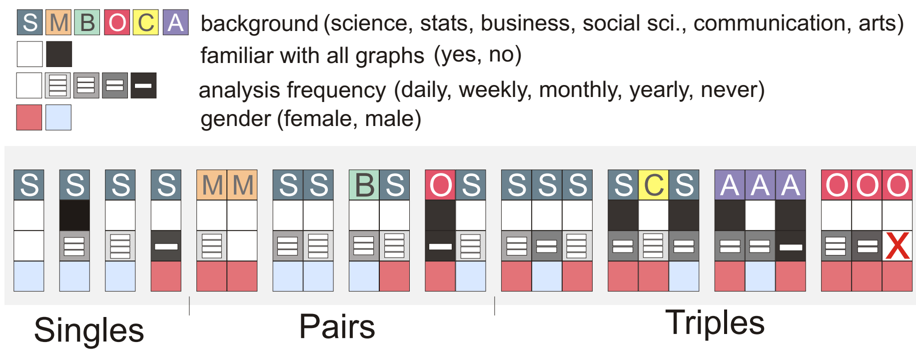


Figure 4.1: Participants' gender, chart familiarity, and data analysis frequency.

4.4.2 Apparatus

Participants worked on a large table (90 × 150 cm) and were given 15 × 10 cm cards, each showing one data chart. The table was covered with a large paper sheet, and several pens, pencils, rulers, erasers, scissors, and sticky notes were provided. Six different types of charts were used. These charts showed different subsets of the data and each data subset was shown in at least two different representations (e. g., line chart and bar chart). Figure 4.3 gives an overview of the charts used and shows how many participants reported themselves to be unfamiliar with a given chart; however, data was always redundantly encoded in familiar charts. All charts can be found in Appendix A.1.

4.4.3 Tasks

Participants worked on two task scenarios, each composed of a different data set with its own representations. The data sets used in the study are part of the sample files provided with the analysis software SPSS 14.0. The behaviour data set (Scenario B,

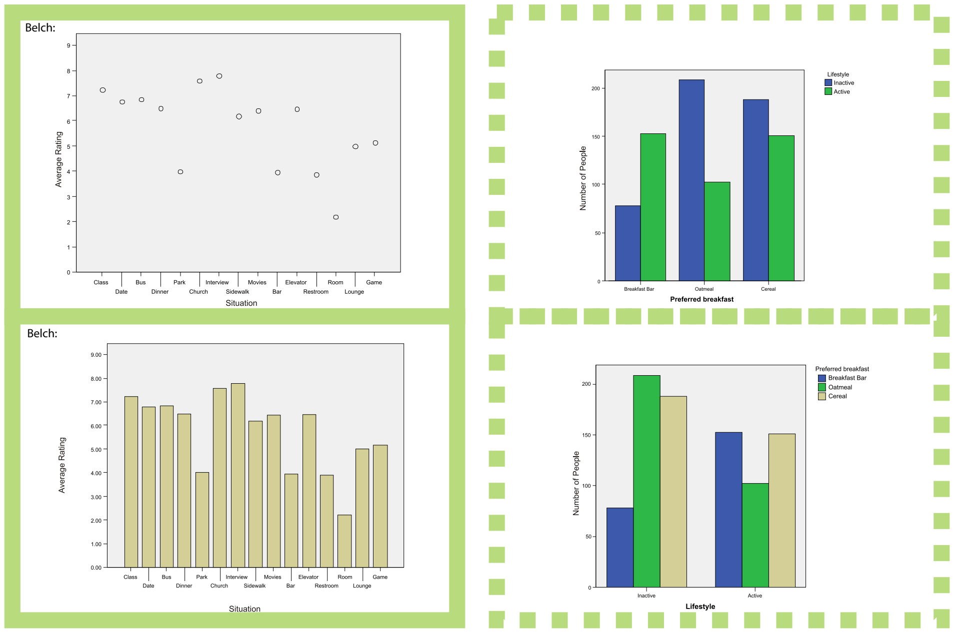


Figure 4.2: Example charts given to participants in the study. Left: Scenario B, Right: Scenario C.

behavior.sav in SPSS) included 32 charts (1 stacked area, 1 line, 15 scatter plots, 15 bar charts). The data shown in these charts was about ratings for the appropriateness of 15 behaviours in 15 different situations (e. g., belch in church, see Figure 4.2). The cereal data set (Scenario C, cereal.sav in SPSS) which included 30 charts (3 pie, 9 bar, 9 stacked bar, 9 line charts, see one example in Figure 4.2) was about an imagined study of preferences for certain breakfast options. No specialized knowledge about the data was required to solve the tasks and high task engagement was evident throughout the observations. The presentation order of these scenarios was counter-balanced between groups. Similar to the design used by Mark and Kobsa (2005), our scenarios each contained an equal number of *open discovery tasks*, where tasks could have several possible solutions, and *focused question tasks* which had only one correct answer. For example, one scenario contained study data on ratings of appropriateness of 15 behaviours in 15 different situations. In this scenario, an example of an open discovery task was, "choose three situations and describe behaviours most appropriate for that situation according to the graphs," and an example of a focused question was, "is it more appropriate to argue or belch in a park?" An overview of all tasks and task charts can be found in Appendix A.1.

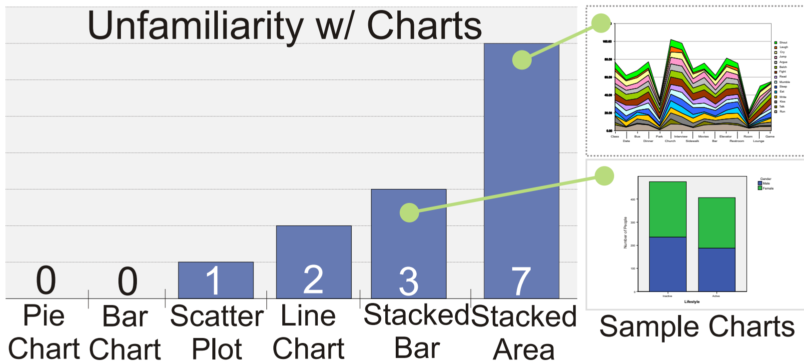


Figure 4.3: Unfamiliarity of participants with study charts.

Scenario	Task	Type
C	1) Give a short description of the participants' characteristics.	open
	2) Who should each breakfast option be advertised to?	open
	3) Do more females prefer oatmeal than active people prefer cereal.	focused
	4) Do more inactive people prefer oatmeal than people over 60? Do you think there might be a relationship between lifestyle and age in terms of preference for oatmeal?	focused
B	1) Find pairs of behaviours that have similar ratings in at least three different situations.	open
	2) Choose three situations and describe behaviours most appropriate for that situation.	open
	3) Find two situations that have at least five behaviours with similar ratings.	open
	4) Is it more appropriate to argue or belch in a park?	focused
	5) Where was it most appropriate to laugh.	focused
	6) What behaviour in which situation was most appropriate and which was most inappropriate.	focused

Table 4.1: Study questions and type for Scenario C (Cereal) and Scenario B (Behaviour).

4.4.4 Procedure

Participants were greeted and then seated themselves around the table. Next, a short tutorial was provided on the types of charts, tasks, and scenarios used in the study. Participants were told that they could use any of the tools on the table (pens, rulers, etc.) to work with the charts, and that they could write on anything as they saw fit (e. g., cards, scrap paper, table, etc.). Participants were then given an example task scenario to clarify the process. Once participants reported to have understood how to proceed, a problem sheet for each task scenario was given to them in turn. Participants were instructed to work on the tasks in any way they felt comfortable. Upon completing both task scenarios, participants filled out a questionnaire asking them about their experiences during the study and to collect demographic information. The groups of two and three participants naturally discussed their tasks and progress and single participants were asked to use a "talk aloud" protocol. The questionnaire given to participants can be found in Appendix A.1.

4.4.5 Data Collection and Analysis

For both our data collection and analysis we followed basic principles of Interaction Analysis as described by Jordan and Henderson (1995). Our study is based on the assumption that information analysis practices are situated within the interactions between collaborators and the artifacts of the workspace. Our goal was to collect evidence from data that would help to ground emerging theories about the collaborative analysis practices. During each study session two observers (one the author of this dissertation) were always present. Both observers collected notes, and each session was video or audio taped. The scene was videotaped from two locations: one above the table clearly showing the information artifacts and another from the side capturing the participants and recording voice. One group did not agree to be video-taped but allowed us to audio-tape the session. 610 minutes of video data was collected ($\sim$50 minutes for each session).

The two observers collected independent notes on the process of participants in solving the tasks. These notes were not compared between sessions as the goal was for each observer to establish an independent picture of groups' working styles. Every attempt was made to refrain from speculation as to participants intentions and instead to support observations with examples of participants' interactions. Next, the two observers went over the field notes and compared and categorized their observations using an informal affinity diagramming approach. The observations were thus grouped into a first set of codes for observed general data analysis processes used by the participants.

I implemented a video coding tool capable of synchronously playing both videos while recording time-stamped codes and notes during the video coding. Next, I engaged in a first video coding pass with the tool in which each participant's activities were assigned to one of the initial codes. When none of the codes fit any of the established codes, I assigned a new code and marked the video sequence for later discussion with the other observer. The video coding tool allowed us to quickly jump back to marked sequences to discuss the new codes in question. During these discussions after the first video coding pass, we refined and extended the initial code set. We changed the descriptions for each code, more rigorously described the activities involved, and refined the terminology. For example, the browse process was originally termed "explore" but the term "browse" was finally chosen as it more clearly points to the observed activities for this process.

Last, I engaged in a final coding pass and used the video coding tool to record timestamps detailing temporal occurrences of the codes for each participant in the study. I also frequently stopped the video to take screen shots of examples of group analysis activities for each code. Some of these screen shots can be found in the remainder of this chapter. The team that did not agree to be video-taped was not included in this coding pass. The final code set is outlined in the following section. Lastly, I recorded and analyzed the qualitative data from the questionnaire in order to combine the results with the observations from the video. Where relevant the data is included in the results section to support our observational data.

4.5 Findings

This section outlines our understanding of the collaborative and individual visual analysis processes uncovered during our analysis. This discussion is followed by illustrating how the processes themselves were *not* temporally organized in a consistent way across groups. Then in the next section, we relate our findings to prior work and discuss implications for the design of information visualization tools.

4.5.1 Processes in Visual Information Analysis

Our analysis revealed eight processes common to how participants completed the tasks in the study (summarized in Table 4.2). Each process is described using real examples drawn from the study, discussing participants' interactions with one another and the

Process	Description	Goal
Browse	scan through the data	get a feel for the available information
Parse	reading and interpretation of the task description	determine required variables for the task
Discuss Collaboration Style	discuss task division strategy	determine how to solve the tasks as a team
Establish Task Strategy	establish how to solve a task with given data & tools	find an efficient way to solve the problem
Clarify	understand a visualization	avoid mis-interpretation of the data
Select	pick out visualizations relevant to a particular task	minimize the number of visualizations to read
Operate	higher-level cognitive work on specific data view	solve task or sub-task
Validate	confirm a partial or complete solution to a task	avoid errors in completing the task

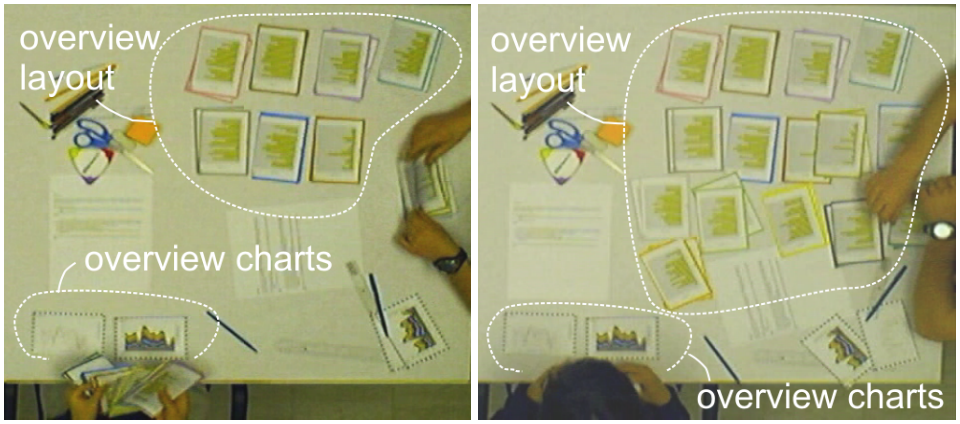
Table 4.2: The eight processes observed in our information analysis study. "Discuss Collaboration Style" only applies to collaborative analysis scenarios.

workspace. Where average process times are reported these are an aggregation of several instances of particular processes during both scenarios.

Browse

The browsing process comprises activities involving *scanning through data to get a feel for the available information*. Browsing activities do not involve a specific search related to a task; instead, the main goal is to gain some understanding of the data set. For example, we observed participants quickly glancing through or scanning the information artifacts—likely to see what types of charts were available and the variables in the charts. Five participants took the complete pile of charts and flipped through them in their hands, while 11 others created an elaborate layout of cards on the table. Figure 4.4 shows an example in which two participants use two very different browsing strategies. One participant (bottom of image) lays the two overview charts out in front of him, flipping through the remaining cards in his hand, while the other participant creates a small-multiples overview of the cards on the table as he browses through them one at a time. A small-multiples layout consists of charts being put in a grid-like fashion next to each other in the workspace.

Participants in groups generally browsed in parallel working independently from one another and with their own set of charts. Groups were more efficient than individuals



(a) Start of a browsing session. **(b)** End of a browsing session.

Figure 4.4: Different browsing strategies: the participant on the right creates an overview layout; the participant on the bottom laid out the overview charts and is flipping through the remaining data charts in his hands.

(average browsing times were ~ 30s for groups, and ~ 60s for individuals), perhaps indicating that, for individuals, having a completely clear sense of the data is more important, whereas groups can rely more on others. In one case, the experimenters observed one participant in a group of three who did not browse through the data himself; instead, he watched as his partners laid their cards out on the table.

Parse

The parsing process captures *the reading or re-reading of the task description* in an attempt to understand how to solve the problem. Participants read the task description either quietly or aloud, and in teams, this choice reflected the collaboration style that teams adopted. For instance, teams working closely together would read task descriptions aloud, facilitating joint awareness of the state of the activity, and discussion of how to interpret the question. On average, pairs and triples spent 2 min reading and re-reading the task description; however, individuals referred to the task sheet more frequently (10 times vs. 9 times in pairs and 7 times for triples in total).

The problem sheet was treated as a special information artifact: it often had a prominent spot in a participant's workspace and was seldom moved. Figure 4.5 shows two examples of typical placements of the problem sheet in participants' workspaces. Even if the problem sheet was covered, as for two of the three participants on the left, the

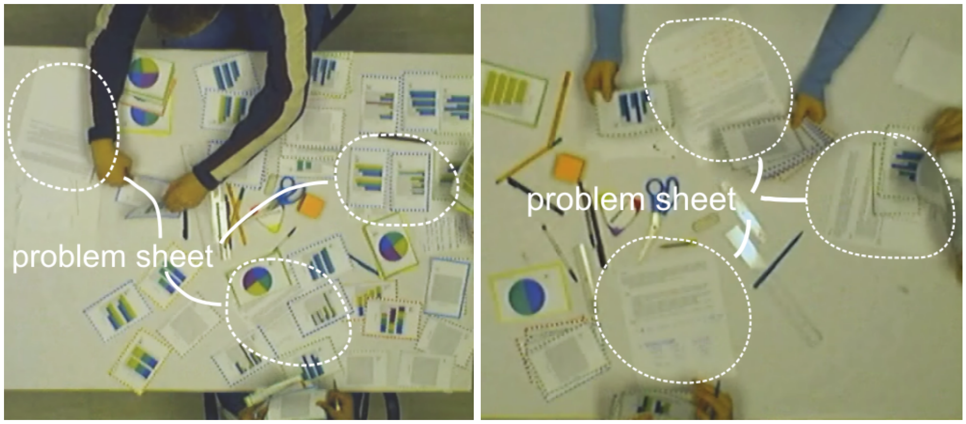


Figure 4.5: Two typical examples illustrating how the problem sheet (outlined) received a prominent spot in participants' workspaces. However, it was often covered by charts that participants were currently working with.

sheet would usually not be moved but accessed by moving artifacts that covered it. The problem sheet was also often used as the primary notepaper to record answers, reinterpretations of the questions, or to retain action lists (e. g., variables to look for in the data).

While many real-world information analysis scenarios may not have a concrete problem description sheet, an assessment of the given problem(s) and the required variables can certainly still occur and would be considered part of this process. The problem sheet can be seen as external textual information that is not part of the current dataset but provides meta-information on the problem, tasks, or data.

Discuss Collaboration Style

Three of the eight teams explicitly discussed their *overall task division strategy*; the five remaining teams seemed to choose their collaboration style on the fly. We observed three main collaboration strategies that teams discussed and/or adopted:

- *Complete task division*. Participants divided tasks between themselves to avoid duplicating work. Each participant worked alone with his or her information artifacts on a separate subset of the problems. Results would be combined at the end without much further group validation.

- *Independent, parallel work*. Participants worked concurrently on the same tasks but independently of each other. When one participant had found an answer, solution and approach were compared and discussed. Other participants might then validate the solution by retracing the approach with their own artifacts, or by carefully examining a partner's information artifacts.
- *Joint work*. Participants talked early about strategies on how to solve the task, and then participants went on to work closely together (in terms of conversation and providing assistance) using primarily their own information artifacts. When one person found a solution, information artifacts were shared and solutions were validated together.

While three teams explicitly discussed a collaboration style, *all 8 teams* changed their collaboration strategy midway through a task scenario or between scenarios. A combination of parallel and joint work strategies was used by six teams and two others used a combination of task division/parallel and task division/joint work. Six of the eight teams started with a loose definition of doing the tasks "together." Strategy discussions were brief: $\sim$1 min on average per scenario. Most of the changes in task strategy were quite seamless, and did not require any formal re-negotiation. This is echoed in the post-session questionnaire in which two participants reported to have chosen their strategy "intuitively" and "by chance." In general, teams showed a strong tendency for parallel work: all eight groups solved at least parts of one scenario in parallel with each team member working with his or her own charts. 14 of 20 participants in teams reported that the main reason they divided tasks this way was for perceived efficiency. Participants in two groups of three reported to have specifically divided the task load by choosing a scribe to record answers.

Establish Task Strategy

In this process, participants *searched for the best way to solve a specific task using the given data and tools*. The goal of establishing such a strategy was to determine the next views or interactions required to extract variables or patterns from the data to solve the problem efficiently. As a team activity, this discussion occurred 22 times with the help of individual information artifacts for all groups and tasks; one participant would present a possible approach to the other participant(s) using examples. For example, Figure 4.6 illustrates an instance where two participants are discussing how to solve a

Figure 4.6: Discussing a strategy on how to solve a task using the chosen chart. Information artifacts are used as aids.

particular task using a specific chart they had chosen. The team frequently flipped between looking at a shared chart and the chart in their own hand. This explicit strategy discussion was more common when teams worked in a *joint work* collaboration style. When participants worked independently or in parallel, the determination of strategy seemed to occur silently (perhaps in parallel to the *parsing* process). For instance, participants might articulate their strategies without discussing the explicit reasoning for it: "I am now going to look for the highest peak." During the video analysis, we only observed on average 1–2 minutes per scenario in which teams specifically discussed their strategy to solve a task. At the end of this process—depending on the chosen strategy—participants often reorganized their information artifacts in the space to create an adequate starting position for solving the task. For example, if the strategy was to find two data charts, then the workspace might be organized to facilitate the finding of these two data charts (as in Figure 4.4).

Clarify

Clarification activities involve efforts to *understand an information artifact*. While the study included common bar, pie, and line charts, also less commonly used stacked bar charts and an area chart were included. The unfamiliar charts required more careful scrutiny by participants. For individual participants, ambiguities in the data display were resolved twice using other charts as aids. Others did not attempt a clarification but chose alternative representations leaving out the one that was unclear. In teams, the need for clarification involved discussion with other participants to decipher and understand the charts and sharing of information artifacts and, thus, led to joint work phases. Overall clarification required less than 1 min for Scenario B and no clarification was required for Scenario C. The clarification times for Scenario B were higher for each group as this scenario contained the most unfamiliar stacked area chart. Only those triples that included participants which were unfamiliar with certain charts required longer than average (1 min, 2 min) for clarification in Scenario B.

Select

Selection activities involved *finding and picking out information artifacts relevant to a particular task*. The experimenters observed several different forms of *selection*, often dependent on the organization of data that was established during *browsing*. We characterized these styles of selection by how artifacts were *spatially separated* from one another:

- *Selection from an overview layout*. Beginning with an overview layout (e. g., small-multiples overview from Figure 4.4), relevant cards are picked out. Selection of cards from this layout involved either a *re-arrangement* of the organization scheme so that relevant cards were placed within close proximity or *marking* by either placing hands or fingers on the cards, or using pens.
- *Selection from a categorization layout*. Starting from a pile-based categorization of information artifacts, piles are scanned and relevant cards are picked out. These cards are then placed in new piles that carry a particular meaning (e. g., relevant, irrelevant). Previously existing piles might change their meaning, location, and structure in the process.

How participants organized these selected data cards was dependent on how they intended to *operate on* (or use) them. The left of Figure 4.7 illustrates an instance where two cards were *relocated* and placed side-by-side for comparison. Figure 4.7 shows an example on the right where a variable was to be measured, so the card was *relocated* closer in the individual person's workspace. The spatial organization of cards relative to piles of data could carry a particular meaning. For example, when an *operation* on a data card was to be brief, a single card was drawn out, operated upon, and then replaced. Similarly, the organization scheme might reflect the perceived importance of

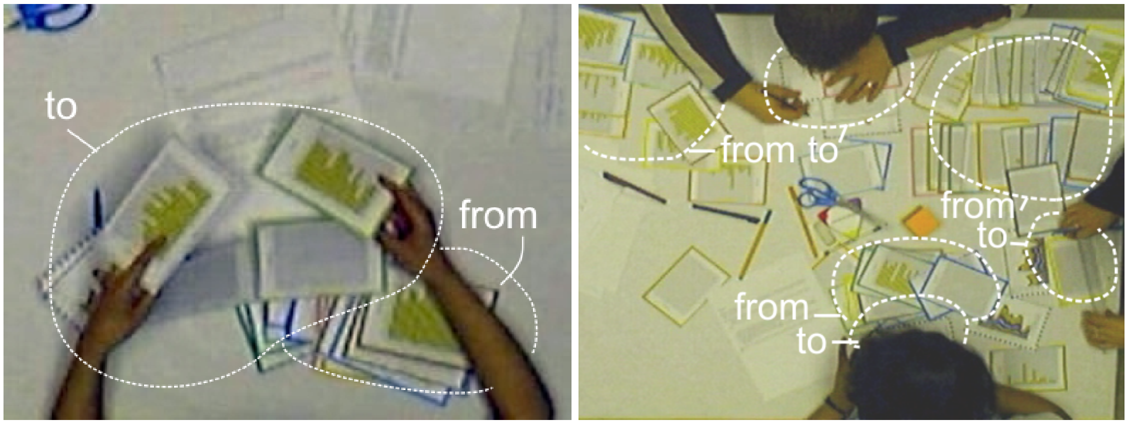


Figure 4.7: Chart organization during selection depending on their intended usage. *Left*: a participant selected four cards for comparison placing them side by side in her hand. *Right*: three participants selected individual charts and placed them in the center of their workspace to measure a specific value.

a set of cards: we observed piles of information artifacts that were clearly discarded (Figure 4.8). Temporally, we also observed different *selection* strategies, which could be loosely classified as "depth-first" or "breadth-first." A "depth-first" approach involved selecting a single card, *operating* on it for a period of time, and then selecting the next card (e. g., Figure 4.8, left). "Breadth-first" strategies selected all cards deemed relevant in a single pass and then *operated* on them afterwards (see Figure 4.8, right). On average, participants spent ~ 4 min selecting data, the second most common process in the study. Selection was an activity that participants in groups performed predominantly in parallel working from their own deck of cards. On few occasions participants in groups jointly selected information from an overview layout when trying to validate an answer.

Operate

Operation activities involved *higher-level cognitive work on a specific view of the data* with the goal of extracting information from the view to solve the task. Figure 4.9 illustrates the two most common types of operation activities: extracting a data value, and comparing data values. To extract a data value from a card, participants often used rulers or some other form of measuring tool (e. g., edge of a piece of paper). To aid recall of these values, participants made annotations: sometimes on the charts themselves, and other times on spare pieces of paper. During the course of both sce-

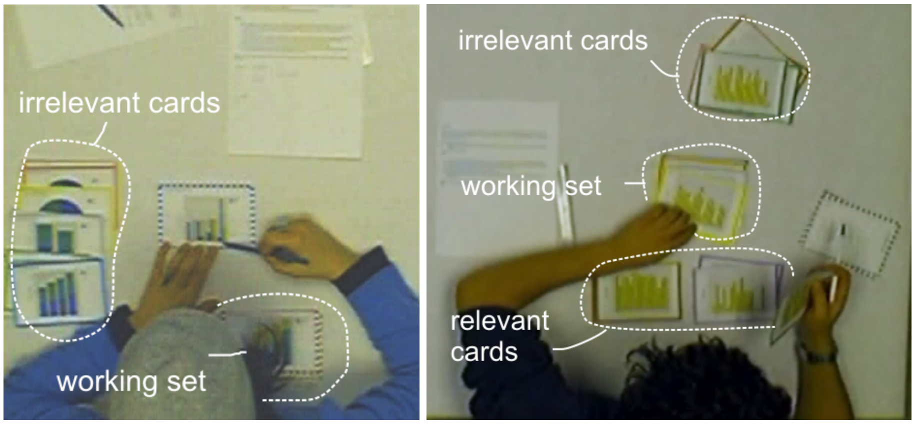


Figure 4.8: Changing categorization during selection. *Left*: a participant placed irrelevant cards to her left and picks single cards to operate on from the working set. *Right*: a participant picked out relevant cards, placed them close to himself, and put irrelevant cards in a pile further away.

narios each participant on average annotated at least three information artifacts (two during Scenario B (Behaviour), one during Scenario C (Cereal)). Comparing values on a specific chart or comparing values across charts was also extremely frequent. Every participant in the study compared charts on at least one occasion. The most frequent comparison involved just two charts but the experimenters also noted 15 occasions of participants comparing three or more charts. In the study, participants arranged the charts for a comparison during *selection*: cards would be placed in close proximity to facilitate easier reading of either individual values or patterns (Figure 4.8). Participants were quite creative in their use of tools to aid comparison: marking individual values, bending or cutting individual charts (to facilitate placing values physically side-by-side), or on seven occasions we noted overlaying of charts atop one another in an attempt to see through the top chart. The operation process typically generated a set of results which were synthesized with previous results and/or written down. Participants in groups predominantly operated on the data in parallel and just reported results to the team if other tasks depended on these results (e. g., when work on the same task had been split up). Operation was the most time-consuming activity in our study. On average participants spent almost half of their time (11 min) on operations per scenario. 64% of operations followed a selection process.

Figure 4.9: Two participants showing two different types of operations on the information. The participant on the right is comparing two cards using a ruler while the participant on the top is measuring a particular value.

Validate

Validation activities involved *confirming a partial or complete solution to a task*. Beyond confirming the correctness of a solution, teams also ensured the correctness of the process or approach that was taken. In teams, the validation process often included discussion coupled with sharing of information artifacts: on 47 occasions participants validated others' solutions by looking carefully at the solution using shared representations, while at other times they searched for the solution by using their own information artifacts (i. e., the process or approach was shared instead of the artifacts themselves). When working more independently, the validation process only involved the presentation of a solution by the group member who had it. In groups where collaborators worked more closely, the collaborators would often ensure that the other participants had understood the process with which a solution was found. For individual participants, the validation process involved looking at other data cards (i. e., different representations) for the same answer. Of interest is that individuals appear to be concerned about the "correctness" of their solution or approach based on other information artifacts, while teams also rely on a collective validation from the social group. On average groups of three spent the longest time validating their answers ($\sim$ 3 min), pairs spent $\sim$ 1–2 min validating, and individuals spent less than one minute validating their answers.

4.5.2 Temporal "Sequence" of Processes

To understand how the processes related to one another in terms of a temporal relationship, the experimenters analyzed the video data from the study, coding the time interval of each individual's activities using these process labels. Looking at the time intervals revealed three aspects of participants' activity: first, while certain processes frequently occurred before others (e. g., *select* most frequently appeared before *operate*), *no common overall pattern appeared*; second, *individuals* varied in how they approached each task, and finally, teams also varied drastically in how they spent their time. A few example charts are shown next. All charts for singles, pairs, and triples exhibit this same extreme variability of approach and can be found in Appendix A.1.4.

Figure 4.10 shows the coded temporal sequence of analytic processes during Scenario B for three pairs. Notice how the sequence of processes was quite different for each pair, even though participants worked on the same tasks using the same tools, representations, and views of the data. Even within teams participants did not show the same temporal occurrences of processes. On average, participants in pairs were concurrently working in the same process for $\sim 70\%$ of the time. For Scenario B (Figure 4.10), P2 has a 65% co-occurrence of the same processes, P3 80%, and P4 69%. This reflects the collaboration strategies participants had chosen. P3 had switched from a complete task division to joint work in this scenario while P2 and P4 were working mostly in parallel. Participants in groups of three only showed a 40% co-occurrence of processes on average. In both charts in Figure 4.10, Tasks 1–3 were open discovery tasks and Tasks 4–6 were focused question tasks. The experimenters noticed that both individuals and teams solved focused question problems quicker than open discovery tasks. In contrast to individuals, teams seemed to have a better understanding of the tasks (established during the task strategy process) and solved them (both focused and open discovery tasks) more correctly. This result echoes findings by Mark and Kobsa (2005) that suggest that groups perform more accurately, albeit slower. Of course, teams also exhibit *establishing a task strategy* more so than individuals, again in order to establish common ground (Clark, 1996), or to ensure a correct or agreed-upon approach.

Figure 4.11 shows a detail view of a specific task, charting individual participants and three of the participant pairs. Notice that even for a single task occurring over a roughly

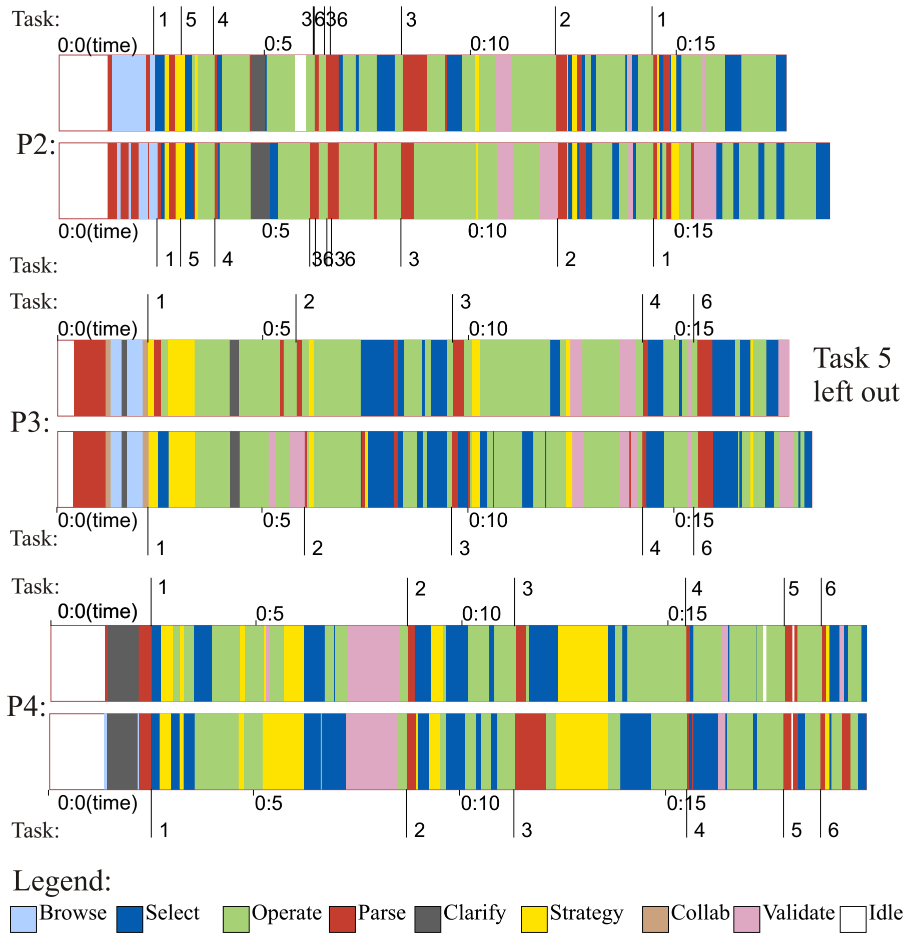


Figure 4.10: Temporal sequence of processes for three pairs during one complete scenario. Time is indicated as hours:minutes.

five minute sequence, *how* the participants engaged in the task, and the temporal distribution of process time varied.

4.6 Discussion

To this point, we introduced a set of processes that occur within the context of collaborative and individual visual information analysis. These processes that are apparent from the study provide an understanding of how teams and individuals use information artifacts in the physical, non-digital workspace to solve visual information analysis tasks and of how team members engage with each other during this process. In this sec-

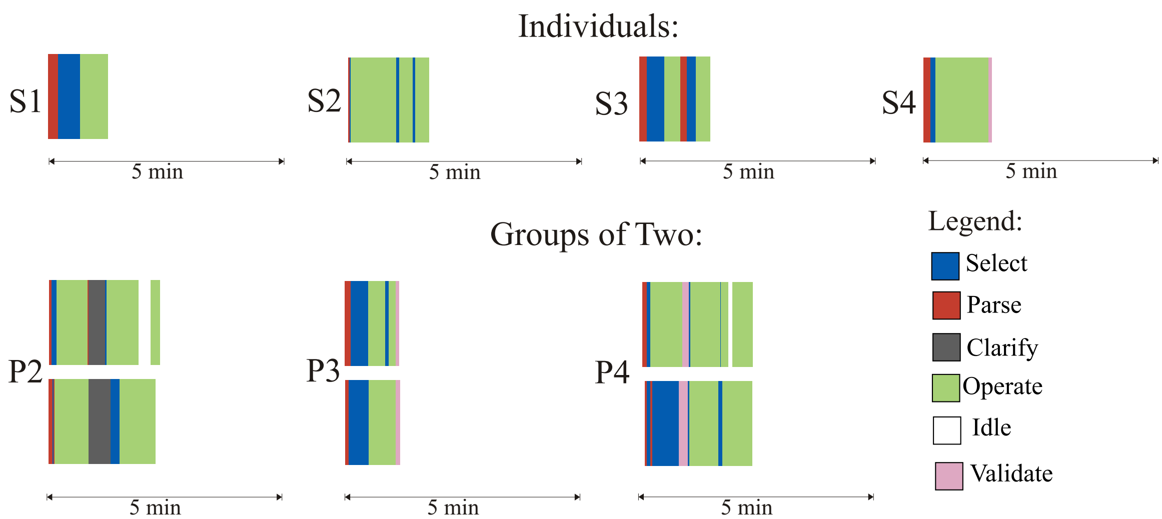


Figure 4.11: Temporal sequence of processes for one open discovery task. The top row shows timelines for individual participants (S1–S4). The bottom row holds timelines for participants in groups of two (P2–P4).

tion, we discuss how the processes we described in the previous section relate to other information analysis models. This discussion reveals that while some processes relate closely to existing models, the temporal analysis from this study suggests that with appropriate tools, both the collaborative and individual information analysis processes may benefit from temporal flexibility.

4.6.1 Comparing Models

Comparison with the Sensemaking Cycle

Card et al. (1999) provide a high-level model of human activity called the "Knowledge Crystallization" or "Sense-Making Cycle" where the goal is to gain insights from data relative to some task. This model, as seen in Figure 4.12, includes five main components: *foraging for data, searching for a schema* (or representational framework), *instantiating a schema, problem solving, and authoring, deciding or acting*. It builds on work by Russell et al. (1993) which involved observations of collaborative work, and an extension can be found in (Thomas and Cook, 2005).

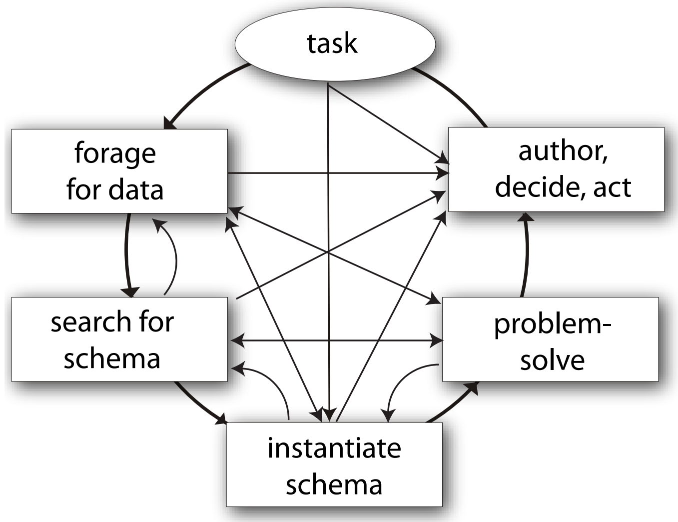


Figure 4.12: Sensemaking model after (Card et al., 1999, pp. 10).

The Sensemaking Cycle has several components related to the processes found in this study. It outlines a process called "foraging for data" that includes our *browse* process. Spence (1999) specifically explores the "foraging for data" component in terms of visual navigation. In particular, he relates visual navigation to cognitive activities (such as internal model formation and information interpretation), thereby arguing that how people can navigate, explore, and visualize a data space will shape how people think about the data. Spence (2007a) distinguishes three different browsing activities: *exploratory browsing* where the goal is to accumulate an internal model of part of the viewable scene; *opportunistic browsing* to see what is there rather than to model what is seen; and *involuntary browsing* which is undirected or unconscious. The experimenters primarily observed exploratory browsing, and saw that, as part of this process, participants established a layout of cards, or put cards in observable categories (e. g., by variables or graph types). It seemed that those participants who created a specific layout of cards in their work area created a type of overview by imposing an organization (even if a loose one) on the information artifacts. Thus, a physical manifestation of the creation of an "internal model of the data" was observed in this study. Furthermore, these physical layouts (a consequence of the browsing phase) clearly relate to Shneiderman's "overview" task (Shneiderman, 1996).

"Search for schema" seems to involve activities that here were characterized as being a part of *parsing*, specifically the identification of attributes on which to operate later.

The activity of identifying attributes to look for in the data described in this model is augmented in our parse component by additional activities of discussion and note taking. "Search for a schema" and "instantiate schema" involve activities that help in the search for the best way to solve the given problem with the provided visualization tool and therefore relate to the *Establish Task Strategy* process, albeit being more tool-centered than our definition. *Clarification* is not an explicit component in the Sensemaking Model but the need for clarification would typically arise during the searching for and instantiating a schema components. The *Selection* process from this study is most closely related to the "foraging for data" component in the Sensemaking Model but can extend into the "searching for and instantiating a schema" components when participants have ended their browsing activities and are ready to select specific information important to solving the task. This may include activities that are part of an *Operation* process: problem-solving, including the three levels of reading by Bertin (1983): read fact, read compare, read pattern. *Validation* is not directly represented in Card et al.'s model; perhaps, as the experimenters have also observed, because validation seemed to be often omitted or quite brief for individual participants and the Sensemaking model focuses on individuals.

The Sensemaking Cycle is the most highly coupled and interactive of the models we compare our found processes to. It makes a strong temporal (cyclical) suggestion but does allow for loops within this cycle over defined forward and backward connections between components. In general, the Sensemaking Cycle does not include the same processes but shares similar activities and predicts some of the findings from this study in terms of temporal flexibility. An adaptation of the Sensemaking Cycle is presented in (Thomas and Cook, 2005) for some types of analysis work. This extension includes two main components: A Sensemaking Loop in which a mental model of the data is iteratively developed and a Data Foraging Loop in which information is searched, read, filtered, and extracted. This extended model tries to cover most aspects of intelligent analysis work and the processes from our study mostly relate to those parts within the Sensemaking Loop as discussed above.

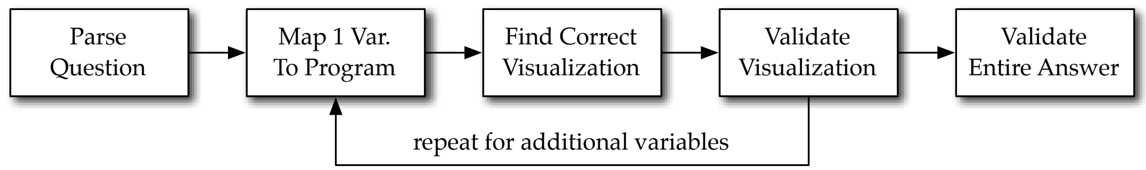


Figure 4.13: Collaborative information visualization model after Mark and Kobsa (2005).

Collaborative Analysis Models

In studying pairs using distributed CAVE environments, Park et al. (2000) articulate a five-stage pattern of behaviour: *problem interpretation, agreement on vis tool to use, search for a trend, discovery reporting, and negotiation of discoveries*. Mark and Kobsa (2005) also provide a five-stage collaborative information visualization model: *parse question, map 1 variable to program, finding correct visualization, validating the visualization, and validation of the entire answer*. A loop is included for additional variables from stages four back to stage two. The temporal sequence of stages in this model was derived from a study of pairs solving both free data discovery and focused question tasks in both distributed and co-located settings. These two models share some similarities, but are clearly not identical. A possible explanation for the disparity is that Mark and Kobsa's model focuses on a context where the pair negotiates exploration through a *shared* tool (i. e., they could not work in a decoupled fashion) whereas Park et al.'s model allows for more loosely coupled work.

Both models share some similarity with the processes discovered in this study. The *parsing* process relates closely to Mark and Kobsa's "parse question" and Park et al.'s "problem interpretation" stages. In our study, additional activities were found that might not have been part of the specific environment under study in these other two models: note taking and frequent discussion about how to interpret a certain task. The *Discuss Collaboration Style* process is not explicitly covered in either model. However, similar to Park et al.'s study the experimenters observed a strong tendency in all group conditions for participants to do at least part of the work using their own views and information artifacts.

According to Mark and Kobsa's model, "map 1 variable to program" is closely related to the *Establish Task Strategy* process in that it would also involve a collaborative agreement on the most appropriate visualizations, parameters, or views to solve the problem, like Park et al.'s "agreement on visualization tools to use."

In contexts where new visualizations are introduced, or individuals are brought in without prior training on particular visualizations, the need for *clarification* would be common. Specifically, beyond providing viewers with aid in developing an understanding of a particular visualization, we would expect individuals to ask for collaborators' interpretations of that visualization or interaction technique or to put their own views and interpretations up for discussion. Considering clarification as a process of analysis is important for designing and evaluating visualization tools but it is not a specific part of the two previous collaborative analysis models.

The articulation of the *selection* process is related to parts of the activities covered by Mark and Kobsa's "find correct visualization" stage and Park et al.'s "search for trend." The description of *selection*, however, more broadly captures the notion of picking out important information beyond operations in a specific visualization system.

"Independent search for a trend including some adjustments to viewing parameters" and "report discovery" include *operations* as defined in our model. Operation is not an individual stage in Mark and Kobsa's model but is integrated in the "find correct visualization" stage. In groups, the validation stage was much more visible and it is also included in these two models as the last stage of information analysis. Mark and Kobsa noticed differences in validation between the free discovery and focused question tasks; a result that was echoed in our study. During more open-ended questions, validation was usually longer and involved more discussion than for focused tasks.

In general, both these models share some of the processes discovered in our study but are quite different in their suggestion of a fixed temporal order.

4.6.2 Temporality and Process-Free Tools

Many of the existing models suggest a *typical* temporal order of components; however, our analysis of the temporal occurrence of the analysis processes suggests that this typical temporal ordering was not evident. This finding of a lack of a common temporal ordering reflects the design of this study; in particular, the stipulation that participants would use a paper-based "information visualization" tool along with traditional tools such as pens, paper and notepaper. Traditional tools have no specific temporal flow in terms of which tools should be used first or for what purpose. Similar observations have been made by Heiser et al. (2004) in a study of non-digital co-located and distributed sketching activities. The flexibility afforded by traditional tools allowed individuals to approach tasks differently. As a consequence, they also allowed groups to transition between multiple stages of independent and closely coupled work rather than regimenting particular work processes. The found processes map to related models, yet our analysis suggests that the temporal ordering of these components is by no means universal. In many digital information visualization systems, the flow of interaction is regimented by structure; in contrast, the use of traditional tools in our study allowed participants to freely choose how to approach and solve problems. The study results also show that in groups certain processes often involved more closely coupled work, in particular: *Discuss Collaboration Style* and *Establish Task Strategy, Clarify,* and *Validate* were often done in close collaboration with shared artifacts while the remaining processes more often involved group participants working more individually with their own artifacts and without much verbal exchange. As close work was more common in some styles than in other styles, this finding leads to new design considerations for co-located collaborative information visualization tools.

4.7 Implications for Design

Here the implications for the design of co-located collaborative information visualization systems are discussed based on the findings of this study.

4.7.1 Support Changing Work Strategies

In group settings, our participants dynamically switched between closely coupled and more independent work. The *Browse, Parse, Operate,* and *Select processes* were most often done on individual views of the data in a more loosely coupled fashion. *Discuss Collaboration Style* and *Establish Task Strategy, Clarify,* and *Validate* often happened in closer cooperation with the other partner(s) and often included shared views of the data. To support these changing work strategies information visualization tools for co-located work need to be designed to *support individual and shared views of and interactions with the data*. Each collaborator should be able to perform individual operations on these views unaffected by his or her team members' actions. However, the tool should also help to share these individual views and, thus, provide awareness of one team member's actions to the other collaborators. To support individual views of the data, interaction with the underlying data structures (deletion of nodes in a tree, change of query parameters, etc.) should be designed so as to not influence others' views of the same data. However, to support shared views of the data, these previous operations should be transferable to group views, for example, to combine highlights, annotations, or other parts of an interaction history. Information visualizations could also be adapted to specifically support *joint* validation, and clarification. For example, highlights of related information across individual team members' data arrangements could support joint validation. For example, by seeing that one team member has moved a piece of information to an "irrelevant" data pile could trigger discussions for validation.

4.7.2 Support Flexible Temporal Sequence of Work Processes

Individuals by themselves as well as those in teams have unique information analysis practices based on their prior experiences, successes, and failures. These well-established work practices should be supported by digital systems. The study showed that all participants worked differently in terms of the order and length of the individual work processes they engaged in, suggesting the need for digital systems to be relatively unrestricting. The temporality of work processes suggested by previous models of the analytic process could imply that common information visualization tools require a specific process-flow. This study, however, suggests that analysts using digital systems may benefit if a flexible order of operations can be performed. Co-located collaborative systems, in which more than one person may work and interact at the same time, should allow team members to be engaged in different types of processes at the same time and also allow them to work together adopting the same processes. For example, one person should be able to select data from or browse a database while another already works on previously selected information.

One very important factor in the support of flexible temporal work processes in our study was participants' abilities to adapt the workspace to their current needs. Information artifacts were re-arranged on the table by all of our participants. We observed that participants had quite distinct individual workspaces on the table in which they laid out their cards. These workspaces were quite flexible and would change depending on tasks as well as, in group settings, on team members' spatial needs. This observation is echoed by the studies of collaborative behaviour reported by Scott et al. (2004) that call for co-located collaborative systems to *provide appropriate functionality in these personal workspaces* (territories). We refer to their paper for further guidelines of how to support personal territories for co-located collaborative work.

Participants also seemed to frequently impose categorizations on data items by organizing them spatially in the workspace. During *browsing*, overview layouts were often created in which the cards were spread across the whole workspace. Mainly during *selection* and at the end of an *operation* process, information artifacts were organized in piles in the workspace. These piles seemed to have inherent categories and varied greatly in size, lifespan, and semantics. Allowing people to *create a spatial organization of the information artifacts* in the workspace should be considered in the design of information visualization systems. These spatial organizations can help to support mental models of the available information. Systems like CoMotion (MayaViz, 2007) are already taking a step in this direction but the typical information visualization system still relies on a fixed set of windows and controls that can rarely be changed, piled, or relocated.

4.8 Chapter Summary

During the analysis of relevant literature in Chapter 3, we noted a lack of dedicated studies on how team members engage with each other and with visualization in the workspace during collaborative data analysis. The purpose of the study presented in this chapter is to help further our understanding of collaborative work with visualization and data analysis tasks and to help establish further design considerations that could be translated to design decisions in the case studies presented in the following chapters.

With the study presented in this chapter, we derived processes involved in *collaborative* and *individual* activities around information visualizations in a non-digital setting. The eight identified processes are: *Browse, Parse, Discuss Collaboration Style, Establish Task Strategy, Clarify, Select, Operate, and Validate*.

For participants in groups, we noted that processes: Clarify, Discuss Collaboration Style, Establish Task Strategy, and Validate were predominantly performed in joint (closely coupled) work phases with frequent discussion and sharing of artifacts, while activities involving the Browse, Select, Operate, and Parse processes were predominantly performed in individual, parallel work. Similar to the study presented by Tang et al. (2006), we observed participants in groups fluidly transition between phases of joint and parallel work. Our finding that specific analysis processes may more commonly be performed in discussion and with joint information artifacts has implication on the design of collaborative information visualization systems.

Furthermore, we have shown how these eight processes relate to other models of information analysis, and provided insights on differences and commonalities between them. Yet, while others have posited a general temporal flow of information analysis, our results suggest this temporal flow may simply reflect an assumption in the design of existing information visualization tools. Thus, we argue that designers should allow for individuals' unique approaches toward analysis, and support a more flexible temporal flow of activity. This flexible temporal flow of activity is particularly important when group members wish to switch between different phases of joint and more parallel work. During parallel work, in particular, group members may follow their own approaches to analysis and want to be able to interact with, explore, and analyze the data without disturbing others. This free switching of work style needs to be supported and a flexible temporal flow of activities—both in the temporal sequence and co-occurrence of work styles in groups—should be integrated in collaborative data analysis tools. The digital systems presented in Chapter 5–7 were designed to provide this type of flexibility.

Chapter 5

CoTree—A System for Collaborative Tree Comparison

In Chapter 3, I discussed a set of design considerations for co-located collaborative information visualization. In Chapter 4, these considerations were extended to include recommendations on analysis processes that team members adopted and the temporal flexibility of their analysis activities. The purpose of this and the following chapters is to show—based on practical examples—how some of these considerations can be translated into a design and to show which design challenges arise in practice.

When designing a collaborative visualization system to support both parallel and joint work activities, several different approaches exist. First, one can design a completely new system that fundamentally supports parallel work and build in mechanisms so that results from parallel work can also be shared with the group in more closely coupled collaboration. Second, one can take a typical desktop based system, which fundamentally only supports joint work through one input pointer and a large shared visualization, and add on capabilities for teams to work more in parallel. Third, one can design a system that fundamentally supports parallel work but includes system features that are meant to encourage and support more closely coupled work. The following three chapters each introduce a system developed under the focus of these respective approaches.

In this chapter I discuss a prototype system, CoTree, which we[1] designed to fundamentally support parallel work with visualizations. The following sections introduce the system and our design decisions related to building the collaborative environment, supporting social interaction, and designing information visualizations. The chapter ends with an example of how the flexible workspace organization principle underlying the system was used to facilitate the tree comparison task that CoTree supports. The experiences from designing this first prototype are summarized in the final section.

5.1 Introduction

In this first case study, I designed a new collaborative analysis system with the goal to fundamentally enable parallel work with data. This is achieved by providing each collaborator their own views and interaction possibilities with the data. Each view of the system is completely decoupled from all others, allowing interactions or annotations that do not affect other views of the same data. However, views can be individually placed, rotated, resized, and, thus, shared in order to facilitate more closely coupled work. The system departs from the design of typical desktop-based information visualization systems and provides a new layout strategy for the visualization workspace, as well as a new interaction design based on touch input. The system supports hierarchical data comparison tasks for collaborative work with dual-touch input, shared and individual views on the hierarchical data visualization, flexible use of representations, and flexible workspace organization to facilitate group work around visualizations.

CoTree supports work with hierarchical data, specifically with two different types of tree representations: a space-filling radial tree layout and a dendrogram. The radial tree layout similar as the one presented by Stasko and Zhang (2000) was implemented with a minor adjustment that places labels in a circular fashion inside the nodes (see Figure 5.1, left). This type of labelling was chosen to facilitate orientation-independent reading from different positions around the tabletop display. Since tree comparison is a task commonly performed on phylogenetic trees (Munzner et al., 2003), a dendrogram

[1] Portions of this chapter were previously published in (Isenberg and Carpendale, 2007). Reprinted, with permission, from *IEEE Transactions on Visualization and Computer Graphics, Interactive Tree Comparison for Co-located Collaborative Information Visualization, Petra Isenberg and Sheelagh Carpendale*, © 2007 IEEE. Thus, any use of “we” in this chapter refers to Petra Isenberg and Sheelagh Carpendale.

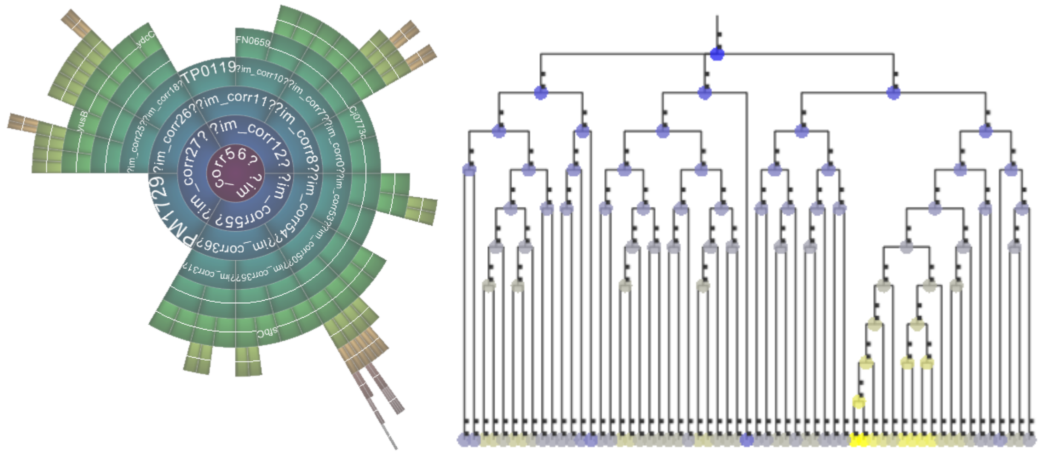


Figure 5.1: The two representations used in CoTree. *Left:* a radial tree layout with radial labelling. *Right:* a dendrogram with additional node colouring to reveal level information.

tree layout (see Figure 5.1, right), the most common representation for this type of hierarchical data, was also implemented. In the dendrogram layout, all leaf nodes are extended to the bottom of the graph. To additionally reveal their place in the hierarchy, nodes are coloured according to their level. CoTree can easily be extended to support other types of representations.

5.2 The Collaborative Environment

CoTree was designed to run on a large digital tabletop display. The tabletop was chosen as the joint working environment as it was shown to have a positive impact on collaboration in joint problem solving tasks (Rogers and Lindley, 2004). The digital table we used was built using a touch-sensitive DViT board from SMART Technologies with two concurrent and independent inputs (see Figure 5.2). The tabletop setup has 2,800 $\times$ 2,100 pixels ($\sim$ 5.9 mega pixels) provided by four rear-mounted projectors (2 $\times$ 2). This setup offers an adequate size, configuration, input, and resolution for small groups of two individuals to work together on data analysis tasks. However, only two simultaneous touches are currently supported by the technology and inputs are not identifiable. The implementation is based on a general framework for tabletop interfaces that provides a method of spatially representing properties of the interface using a

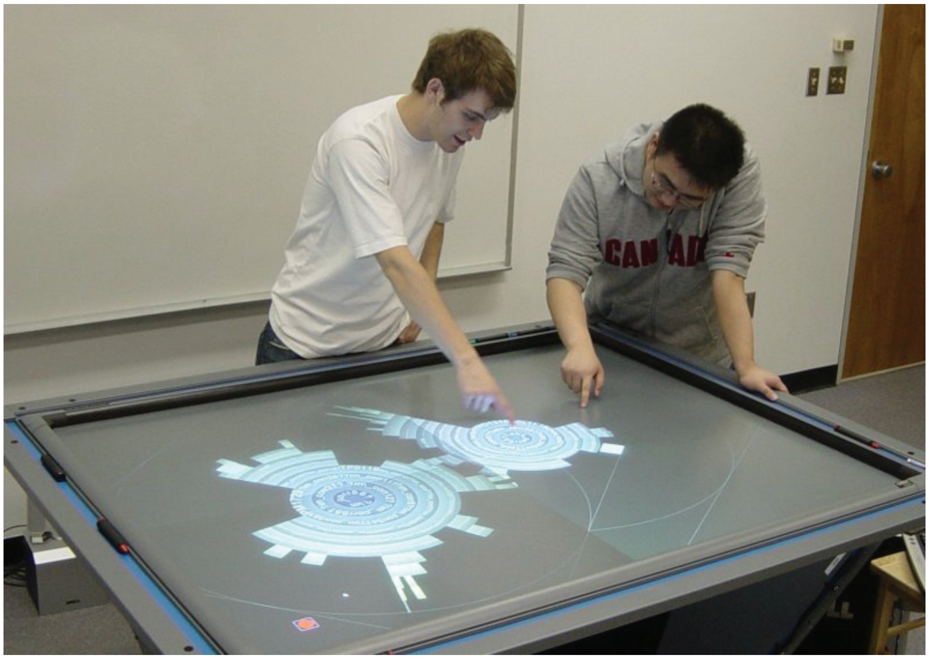


Figure 5.2: The hardware setup for CoTree, the collaborative information visualization application. Two simultaneous pen or finger inputs are possible.

buffer approach (Isenberg et al., 2006a). This framework and the buffer approach help to maintain interactive response on high-resolution tabletop displays. The framework was used, for example, to implement picking and interaction regions for interaction widgets. The framework also provides access to other tabletop interaction metaphors and widgets such as RNT (Kruger et al., 2004), tossing, and Storage Territories (Scott et al., 2005). To facilitate not only an efficient management of memory resources but also to allow people to relate one visual representation of a dataset to a different one of the same data, only one copy of this underlying dataset is maintained. Each view of the data can be customized as discussed next. Appendix A.2 has some code examples discussing how trees were drawn and which methods were used to ensure efficient drawing of multiple views of the same dataset.

5.3 Supporting Social Interaction around the Data

In Chapter 3, I described how the layout of visualizations in the workspace can aid communication and coordination while, in Chapter 4, I discussed the importance of

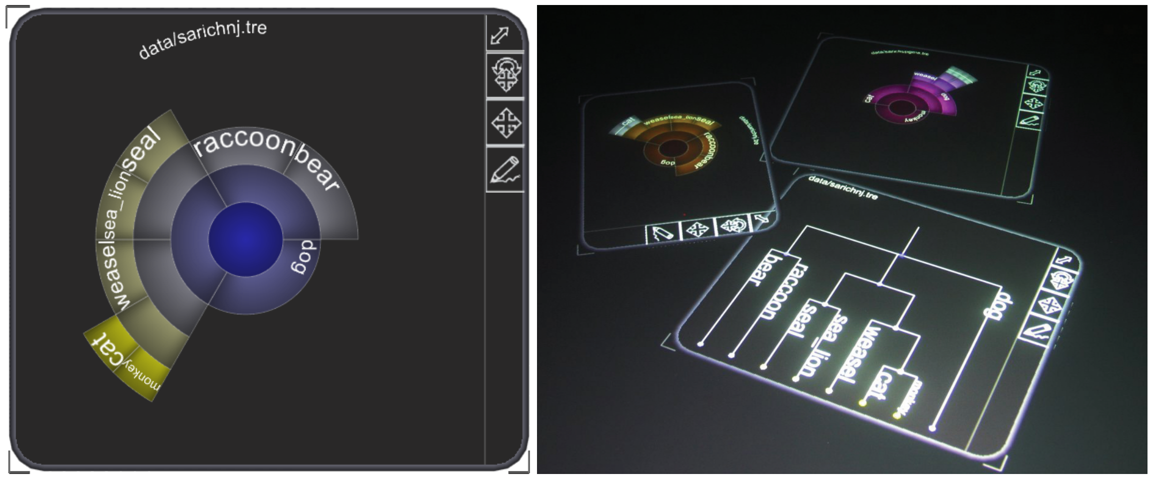


Figure 5.3: A single *visualization plane* showing a radial tree layout can be seen on the left. The right image shows three visualization planes oriented on the tabletop display.

free workspace organization for the support of temporal flexibility of work styles and analysis processes. Therefore, the possibility for individual team members to impose a spatial layout was incorporated in CoTree.

Any information visualization and all control widgets in CoTree can be freely re-oriented and repositioned. Each information visualization is drawn on its own plane with appropriate controls attached to the side. The left of Figure 5.3 shows a single visualization plane showing a radial tree layout and its attached menu buttons. The menu offers common view parameter changes: scaling (zoom), integrated rotation and translation (Kruger et al., 2004), translation only, and annotation. Thus, the plane and attached visualization can be freely moved around the tabletop display. The right of Figure 5.3 shows an arrangement of three visualization planes on the tabletop display. The placement of these menus is further discussed in Section 5.4.

By supporting *free rotation, translation, and scale*, team members working with CoTree can create their own organization of items by putting them in piles, creating a preferred layout (e. g., small multiples or piles as seen during the study presented in Chapter 4). By allowing visualizations to be freely repositioned, *sharing of visualizations* is possible, as the windows can be easily passed to the other collaborators. Representations can also to be passed by tossing them across the table, similar to the implementation for pictures by Scott et al. (2005).

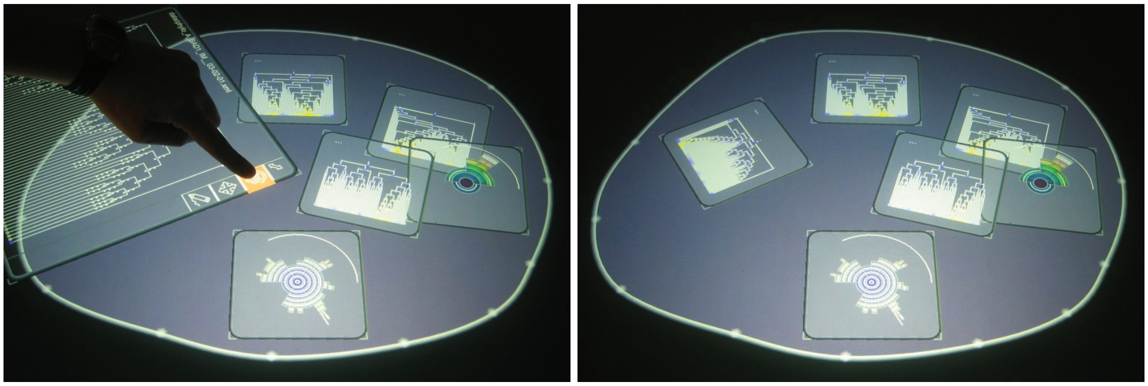


Figure 5.4: A visualization plane is being dropped on a storage container (left) and automatically resized and placed (right).

The possibility for *organizing representations* of data is further supported by providing storage containers that hold visualization planes. In these containers, visualizations can be grouped together, resized, and moved as a unit similar to the implementation by Scott et al. (2005). Figure 5.4 shows an example of a visualization plane being placed in a storage container. First, the plane is dropped on the container (left), and then automatically resized and placed in the storage container (right). Items in the storage container can be placed casually, neatly organized, or piled, and can then be moved as a unit. These containers can provide a means for collaborators to store intermediate exploration results for later reference or comparison.

To support *communication* about the data, CoTree includes annotation directly on the provided visualizations and separately on sticky notes. Interactive sticky notes for low-resolution input (Isenberg et al., 2006b) can be used to take general notes during the exploration process to, for example, write down intermediate results or variables to look for. Using these annotations, collaborators can become explicitly aware of each others' exploration processes even if the individual work takes place in separate areas of the workspace. Figure 5.5 shows how sticky notes and integrated annotations can be used to mark interesting information in a tree layout.

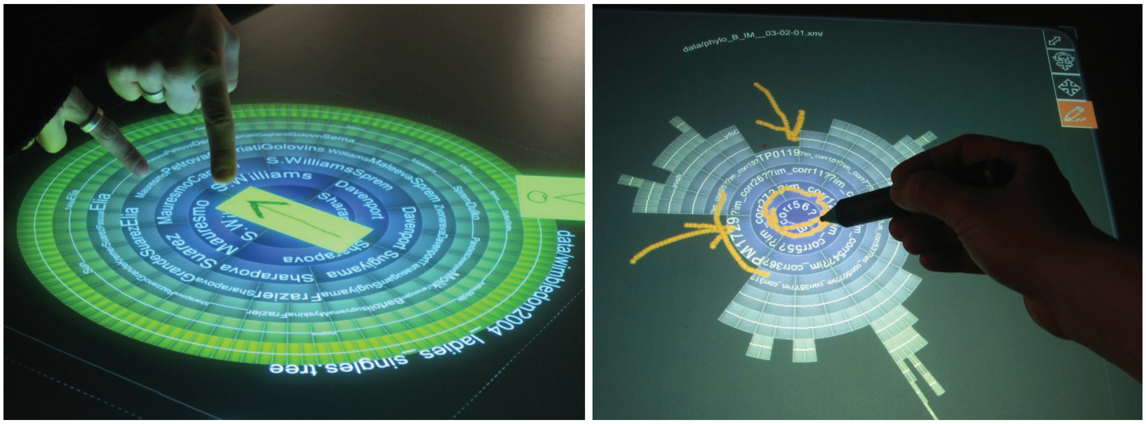


Figure 5.5: Annotation of visualizations. *Left:* Annotation using interactive sticky notes (Isenberg et al., 2006b). *Right:* Annotation integrated directly on the information visualization.

5.4 Designing the Information Visualizations

One of the main challenges encountered during the design of CoTree, was the large number of parameters and interaction possibilities for each view of the data. Since CoTree was designed to support parallel work by providing visualizations on individual view planes, the placement of menus was not as straightforward. On tabletop displays, it is often inappropriate to use standard buttons, taskbars, or menu bars (Scott et al., 2003) since they do not allow for fluid interaction and are not concurrently accessible by several team members. During parallel work in a group setting, for example, several visualizations might have a focus at the same time or a visualization might be interacted upon by more than one person at a time. Furthermore, research on a system for collective co-located annotation of digital photos (Ringel Morris et al., 2006) revealed that team members strongly preferred a replicated set of controls over a centralized shared set of controls because the center of the table was needed for other tasks and because replicated controls avoided accidental touching by other teammates. One possibility to address this problem would, thus, be to redesign menus and embed them on each view plane. For example, for pen-based interfaces pop-up menus (Hancock and Booth, 2004), flow menus (Guimbretiére and Winograd, 2000), or crossing-based techniques (Apitz and Guimbretière, 2004) have been suggested. Due to the technical limit of two unidentifiable inputs, we chose not to implement specific gestures but instead chose to provide basic functionality local to each data view. However, these interactions local to the specific view planes only partially support temporal flexibility of interaction with the data (Chapter 4). When one plane is shared by two team members, interaction still has to be negotiated through a shared menu bar as a view cannot be simultaneously resized and moved, or moved and annotated.

As a design experiment, we only placed common view operations (move, rotate, resize, annotate) on each view plane and chose to provide additional means for parameter changes external to each view. We tried an approach similar to the one by Ringel Morris et al. (2006). We designed parameter menus as separate items in the visualization workspace in order to avoid having to build more complicated menus attached to each visualization plane. Parameters of a specific view can be changed through a drag and drop operation. For example, a visualization can be dropped on ColourChanger widgets in order to initiate a change of its colour scale (see Figure 5.6).

While initial response during informal demonstration sessions has been positive, we acknowledge that further careful studies are required to evaluate the varying effects of this design choice on group work. One obvious disadvantage of this approach is that visualization planes have to be moved in order to change visualization parameters. Such movement can destroy a meaningful layout of the planes that may have been created in the workspace, e. g., to facilitate cross-comparison. Alternatively, these widgets could also be dropped on the visualization plane in order to initiate a parameter change. This alternative would avoid having to reposition visualization planes if a careful layout has been created by the group. However, experimenting and studying this approach in comparison with other input techniques like flow menus (Guimbretiére and Winograd, 2000) would be worthwhile, in order to assess the benefits and drawbacks for group work. The system discussed in Chapter 6, for example made a number of operations available through gestures.

In order to create multiple representations of the same dataset, we designed floating labels for each dataset in the workspace, as can be seen at the left of Figure 5.7. These dataset labels can be freely repositioned and, thus, passed to other collaborators to facilitate shared access to this resource. A new visualization plane with a default visualization, for example, is created by touching the label (see Figure 5.7, right).

To support changing decision-making strategies and personal tastes and conventions, CoTree provides individual access to different types of representations. If an individ-

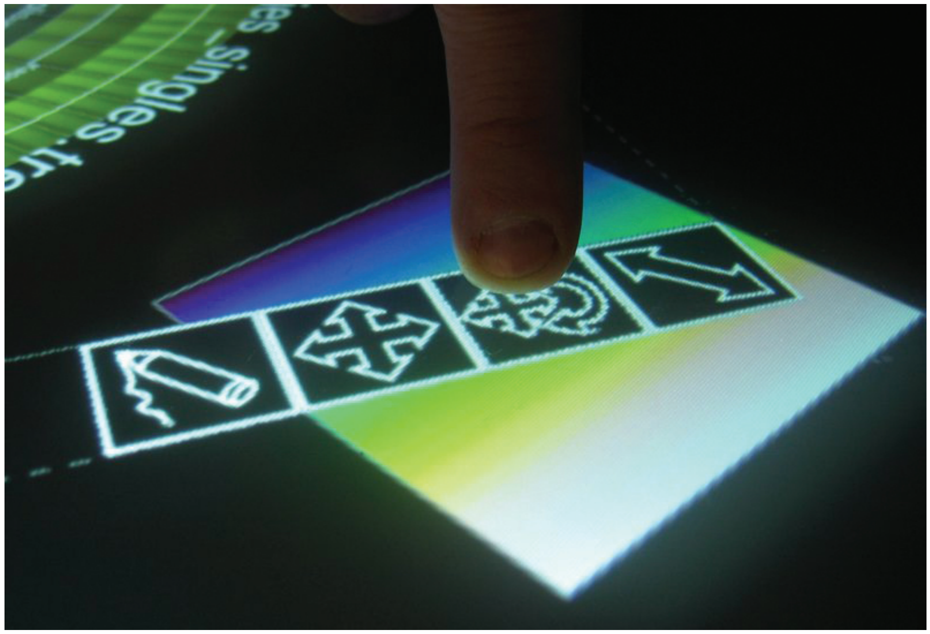


Figure 5.6: A visualization plane is dropped on a ColourChanger widget that changes the colour scale with which the tree is displayed.

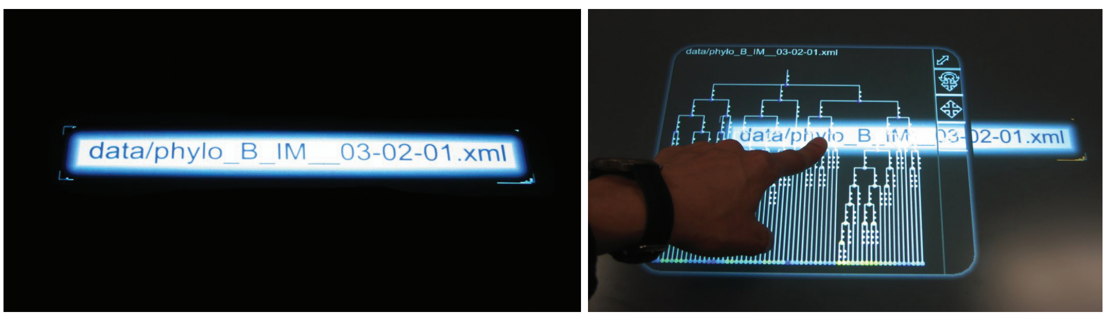


Figure 5.7: Creation of additional representations using dataset labels. *Left:* an example of a floating dataset label on the tabletop display. *Right:* A team member created a new visualization by touching the dataset label.

ual group member wishes to visualize the data using a different representation of the data, e. g., a containment tree layout instead of a node-link diagram, the specific representation can be changed with a drag-and-drop operation without interfering with other group members' operations. Figure 5.8 shows how a representation change is performed. In the left image, the visualization plane is dragged onto the RepresentationChanger widget. As soon as the finger attached to the visualization plane is lifted off the widget, the representation changes to the desired one as can be seen at the right of Figure 5.8.

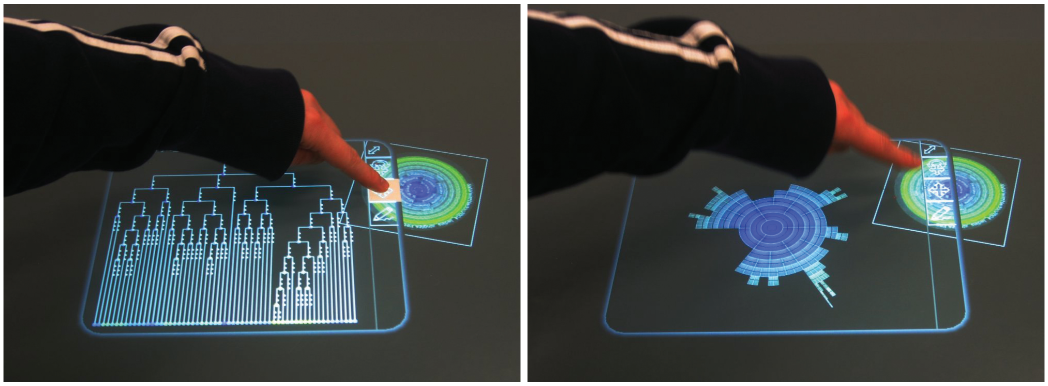


Figure 5.8: Switching a representation type with a drag-and-drop operation.

CoTree currently only includes annotation and note taking capabilities to capture exploration history. Further capabilities will have to be designed for future versions of this or a similar system. The tool presented in Chapter 7, for example, includes persistent information embedded in the information items.

As few evaluations (e. g. Yost and North (2005); Wigdor et al. (2007); Hancock et al. (2009)) have discussed the effects of perspective distortion and orientation on the readability of information visualizations, we did not attempt to correct for possible negative effects.

5.5 Support for Changing Workstyles

Free workspace organization was implemented to be able to support different work styles. When team members transition between more independent work and closer, joint work on information visualizations they can adapt the view parameters of each visual representation to fit the current collaborative needs. Figure 5.9 gives an example in which two collaborators are working individually at first, looking at visualizations in their own area of the workspace (Figure 5.9, left) and then switch to a more closely coupled work style by investigating one visualization together in more detail (Figure 5.9, right). Note that the scaling mechanism has been applied to create a larger visualization to accommodate the concurrent interaction and viewing of both partners and that the plane has been rotated towards both team members. This type of rotation has been previously identified as a strong communicative gesture (Kruger et al., 2004).

Figure 5.9: Visualization planes can be freely arranged in our system. On the left two collaborators are looking at a few representations individually. On the right they are investigating one visualization together.

Any number of windows can be created, moved, and interacted with in the workspace, limited only by the complexity of the graphics and the capabilities of the graphics hardware. By allowing collaborators to each access a copy of a representation, CoTree supports parallel work on the same data. As discussed in Chapter 4, this flexibility in workspace organization as well as the options to personalize representation and presentation of the data can support unique analysis approaches. In particular, by moving representations onto individual view planes team members can select, browse, and operate on the data without disturbing others. However, the system does not contain specific support for team members to specifically coordinate their activities and switch from more parallel to joint work. For example, team members can certainly verbally discuss their collaboration styles, establish task strategies, and together clarify and verify their data (processes discussed in Chapter 4 as being done in more close collaboration). However, there is currently no support built in to support this discussion in close connection to the data. For example, team members may first work in parallel following their own exploration paths in the data. Then, one team member may find a puzzling piece of information in the data and would like to discuss its meaning with another team member. The other team member then has to first either locate this data item on their visualization planes or leave their current work context and jointly view the questioner's visualization. It may be beneficial to further provide assistance to help this switching of work contexts. How this could be done and to what extent this may be necessary is further explored in two other systems presented in Chapter 6 and 7.

5.6 Collaborative Tree Comparison

The development of CoTree was inspired by talking to some biologists at the University of Calgary about their collaborative analysis needs. During discussions with the biologists, they voiced that the comparison of hierarchical datasets would be a beneficial task for them that would also benefit from collaborative analysis. This section shows how CoTree functions as a collaborative data analysis system by stepping through a task of collaborative tree comparison.

5.6.1 Data and Task

As example data for our comparison tasks, we used the InfoVis 2003 phylogenetic data and tasks (`http://www.cs.umd.edu/hcil/iv03contest/`). This dataset contains information on the evolution of two proteins (Protein ABC and Protein IM). It has been suggested that both proteins co-evolve and that such a co-evolution can be detected by comparing the phylogenies of both proteins. The high-level task was to find out whether such a co-evolution was visible. Lower-level comparison tasks included finding where structural changes occurred in the tree. We chose to use the two main files for the ABC and IM proteins and the additional four trees that were provided. Proteins were not paired between the two trees.

5.6.2 Tree Comparison Algorithm and Visualization

We used the similarity measure from the TreeJuxtaposer system (Munzner et al., 2003) in CoTree. The similarity measure is based on comparing the sets of labels of nodes in the subtree under each node. The best corresponding node(s) and nodes with no similarity are highlighted. Figure 5.10 shows a comparison of two trees containing different versions of a carnivore hierarchy. The node “dog” has been interactively selected in the left tree. The best corresponding node “dog” in the right tree is highlighted in yellow, whereas nodes with no similarity are highlighted in red. Nodes in blue are not highlighted in the right tree as they contain the node “dog” (yellow) in their subtree and are therefore “somewhat similar.”

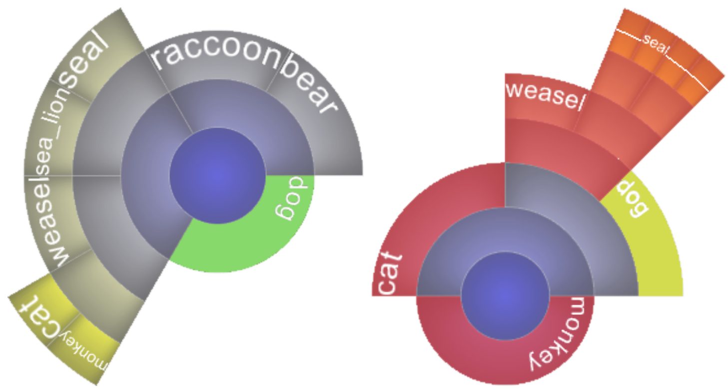


Figure 5.10: Tree comparison of two different versions of a carnivore data set. *Left:* The node “dog” has been selected for comparison. *Right:* The node “dog” is highlighted in yellow as the best corresponding node. Nodes in red have no correspondence with the node “dog.”

Trees in CoTree can be compared by moving their visualization planes close to one another. When planes are close enough for comparison, the borders are highlighted and nodes can be selected to start a similarity calculation. In Figure 5.11, we show two planes on the left in comparison mode (orange border) and a smaller tree to the side that is not currently compared. Any number of trees can be compared by moving them close to others that are already being compared.

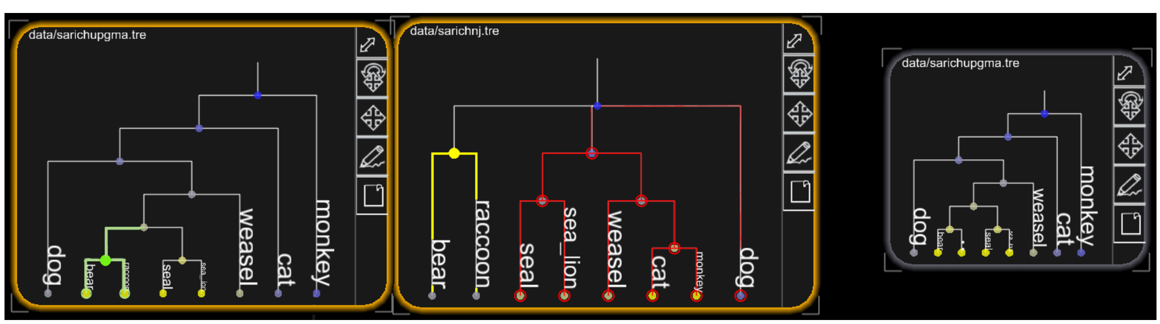


Figure 5.11: Trees can be compared when their planes are in close proximity. Here the two planes on the left are in comparison mode as can be seen by the highlighted (orange) border. The tree on the right is not currently compared with the others.

5.6.3 Solving Collaborative Tree Comparison Tasks

To gain an overview of the available information, each visualization plane can be arranged to facilitate a comparison between all available datasets. In Figure 5.12, two

Figure 5.12: All six datasets have been moved together to facilitate a comparison across all representations.

group members created a comparison overview by organizing their planes to facilitate cross-comparison. Figure 5.13 shows a close-up screenshot of such a comparison. The middle two planes show the main IM and ABC protein representation. The root node of the ABC protein (top row) has been highlighted (green). The two trees on the left, the alternative versions of the IM protein, and the IM protein tree show only dissimilar nodes to the ABC protein (in red). However, the alternative versions of the ABC proteins both show a few dissimilar nodes that need to be inspected further.

This more detailed investigation within the versions of the ABC and IM protein was performed in parallel. The left of Figure 5.14 shows two collaborators who have decided to each investigate one of the proteins. To inspect which nodes have dissimilar values, they have chosen to annotate the dissimilar nodes first and to then examine the nodes and their structure in the hierarchy in more detail. However, closer examination of nodes can also be performed in joint work as shown in Figure 5.14 (right).

A contest task required the examination of the hierarchical structure in terms of whether subtrees moved in the hierarchies or nodes changed position. To facilitate a structural comparison of nodes in this sense, trees in CoTree can be overlaid and then examined. All visualization planes are semi-transparent in order to support this type of tree comparison. Figure 5.15 gives two examples of structural comparison through overlay. The top image shows an overlay of Protein ABC (blue) and Protein IM (magenta). It can be seen that Protein ABC is generally more shallow than Protein IM but has one main

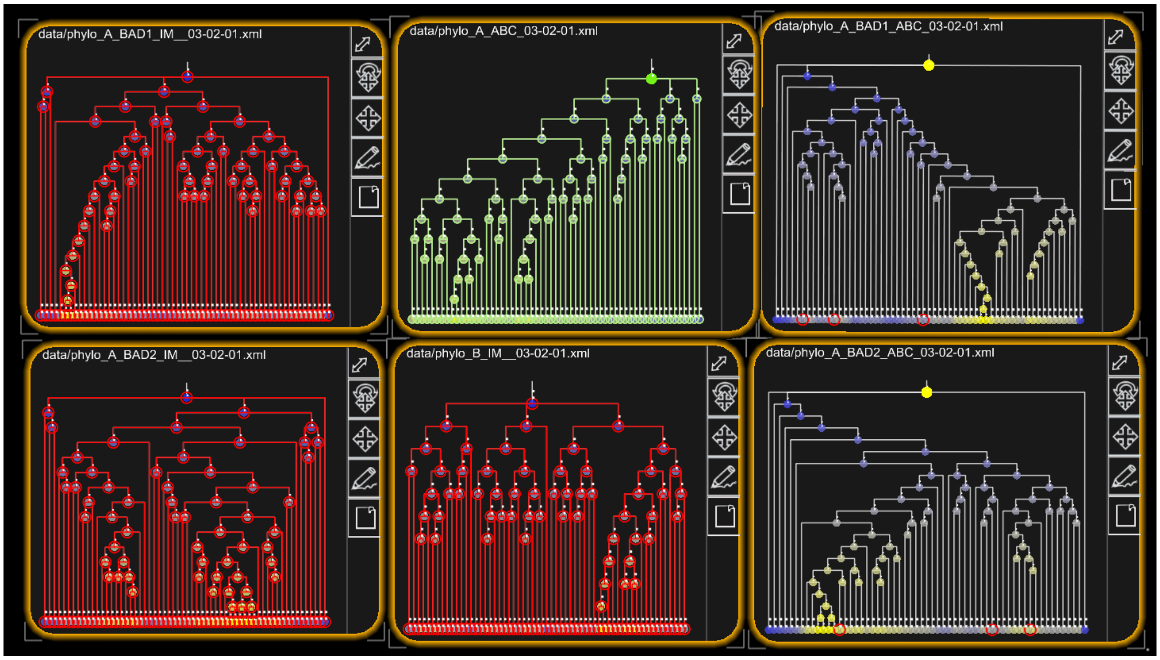


Figure 5.13: Screenshot of the system showing all six trees. The root node of the ABC protein in the top center plane has been highlighted.

Figure 5.14: Closer examination of a few trees. *Left:* Parallel work with each person comparing three trees each. *Right:* Joint work comparing four trees together.

subtree that is wider and deeper than can be found in the other tree. In the bottom image, two collaborators overlaid their exploration history including annotations of similar trees. Similar and dissimilar nodes are highlighted. We are considering options to auto-rotate planes to show the best possible match.

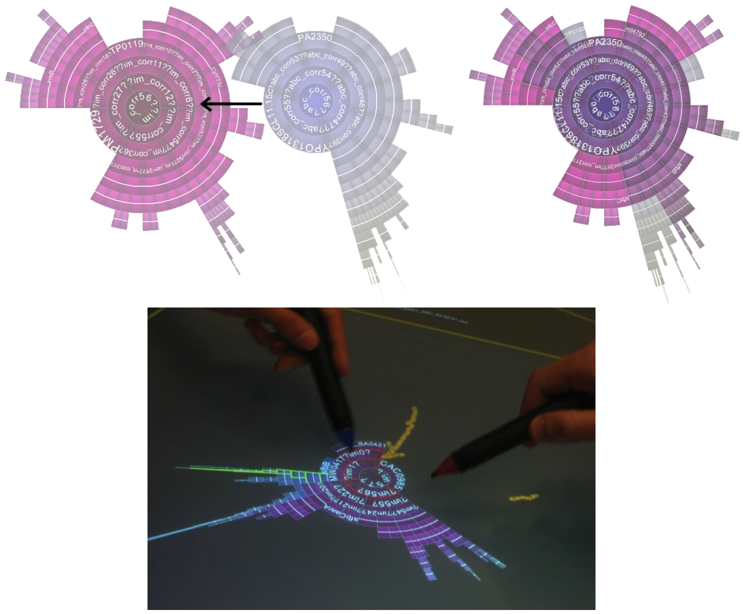


Figure 5.15: Structural comparison through overlay.

5.7 Informal Evaluation

The tree comparison system was shown to a professor in the microbiology department at the University of Calgary to assess its usefulness for supporting collaborative data analysis activities. The response was very positive leading to further demonstration sessions to his students. The students responded positively to the idea of using the tabletop for their work in exploratory analysis of their experimental data. In particular, the benefit of being able to explore different parts of data side-by-side and to be able to cross-compare it, was seen as an advantage. Subsequently, ideas of this system have been used and integrated into a Master's thesis project—Lark (Tobiasz et al., 2009)—that contains improved capabilities to coordinate interactions from several team members as they work on several views of the data in parallel.

5.8 Chapter Summary

The main contribution of this chapter is the implementation of CoTree based on many of the design considerations discussed in Chapter 3 and 4. With the development of this system I have shown that decisions on many levels are necessary to design a collaborative information visualization system for a tabletop display. These decisions are not always straightforward as they depend on a number of different factors. These include not only the data, task, and analysis goals but also the collaborative environment with hardware, input capabilities, group setup, and preferred group work strategies. Some of the design decisions taken here, were also based on information derived from previous literature but would require further study to validate within this work context. For example, the design of menus and parameter changing widgets needs to be validated and compared to other designs that have been suggested for tabletop or pen-based interfaces. In particular, the high number and type of parameters that are typically required in an information visualization system make this an important but difficult part of the design.

CoTree incorporates support for working in parallel by allowing people to create their own copies of views of the data, to flexibly arrange these in their workspace, and to choose personal visualization parameters. This design of the workspace is novel for information visualization as it breaks with the traditional design of visualization systems based on multiple windows, menubars, and dialog boxes. It does not, however, directly support the synchronous interaction on one shared view of the data. Also, the system does not specifically support coordination of activities between team members that would help to transition from parallel to joint work phases. In the following chapter, I introduce a project in which a simple multi-input visualization system was created and subsequently studied to assess the need for such additional coordination support. In contrast to CoTree, in the next chapter I start from a system designed to support only sequential close work and look at minimal changes necessary to introduce possibilities for parallel work.

Chapter 6

CoCoNutTrix: A Study in Collaborative Retrofitting for Information Visualization

In Chapter 5, I presented a new system for collaborative data analysis, CoTree. CoTree focuses on providing individual views of the data on individual customizable view-planes and, hence, fundamentally supports parallel work styles. Beyond customizable view parameters, it did not include specific mechanisms to support changing work styles as people transition from parallel to joint work styles. In this chapter, I discuss the collaborative retrofit of an information visualization tool originally designed to support only a single analyst. The new tool, CoCoNutTrix, differs from the system seen in Chapter 5 in that it involves one large shared visualization. With the development of CoCoNutTrix, we,[1] thus, started from a system that supported joint styles through shared input and a shared visual representation. We attempted a minimal retrofit of the original single-user tool to introduce mechanisms for parallel work. A study assesses the retrofit and shows that parallel and joint work styles were supported by the retrofit. Further, we saw that the nature of the visualization helped team members to remain

[1] Portions of this chapter have been published in (Isenberg et al., 2009). Reprinted, with permission, from *CoCoNutTrix: Collaborative Retrofitting for Information Visualization; Petra Isenberg, Anastasia Bezerianos, Nathalie Henry, Sheelagh Carpendale, and Jean-Daniel Fekete; Computer Graphics and Applications: Special Issue on Collaborative Visualization, 29(5):44–57* © 2009 IEEE. Thus any use of "we" in this chapter refers to Petra Isenberg, Anastasia Bezerianos, Nathalie Henry, Sheelagh Carpendale, and Jean-Daniel Fekete.

aware of each others' activities and to transition between different work styles but that additional awareness information is necessary to better support these transitions.

6.1 Introduction

In the information visualization field, a number of sophisticated tools have been developed for individual analysts. The question arises whether it is possible to simply retrofit such systems for collaborative work by introducing several concurrent inputs that would allow team members to work in parallel or whether new challenges would arise in practice.

In this chapter, I report on a project in which we retrofitted a single-user information visualization system for a low-cost collaboration environment. We evaluated our retrofit in order to assess which factors impacted team members' collaboration styles during data analysis tasks. To create a low-cost collaborative environment we used multiple off-the-shelf projectors that can be simply pointed at a blank wall to create a large display, coupled with technical solutions that replace single mouse or keyboard input streams with multiple input devices (e. g., Jinput (2008)) as can be seen in Figure 6.1. In addition, the visualization tool that was used for the retrofit consisted of one large network visualization filling the entire screen. In contrast to the approach taken with CoTree in Chapter 5, we decided not to duplicate the network visualization in individual planes but to study how teams would coordinate their interactions and work with one shared visualization.

We retrofitted a version of NodeTrix (Henry et al., 2007), a single-user graph visualization environment, to support multiple independent mice. Then, we conducted an observational study to assess how analysts viewed our low-cost environment (e. g., Figure 6.1), and whether it effectively supported collaborative data analysis among data experts using real datasets in the context of social network analysis. To ensure that our low-cost collaboration setup was effective under different realistic settings, the observational study was conducted in three research organizations, using technical facilities present in each organizations. With this research we assess one example for transitioning from single-user to multiple-user information visualization support for co-located

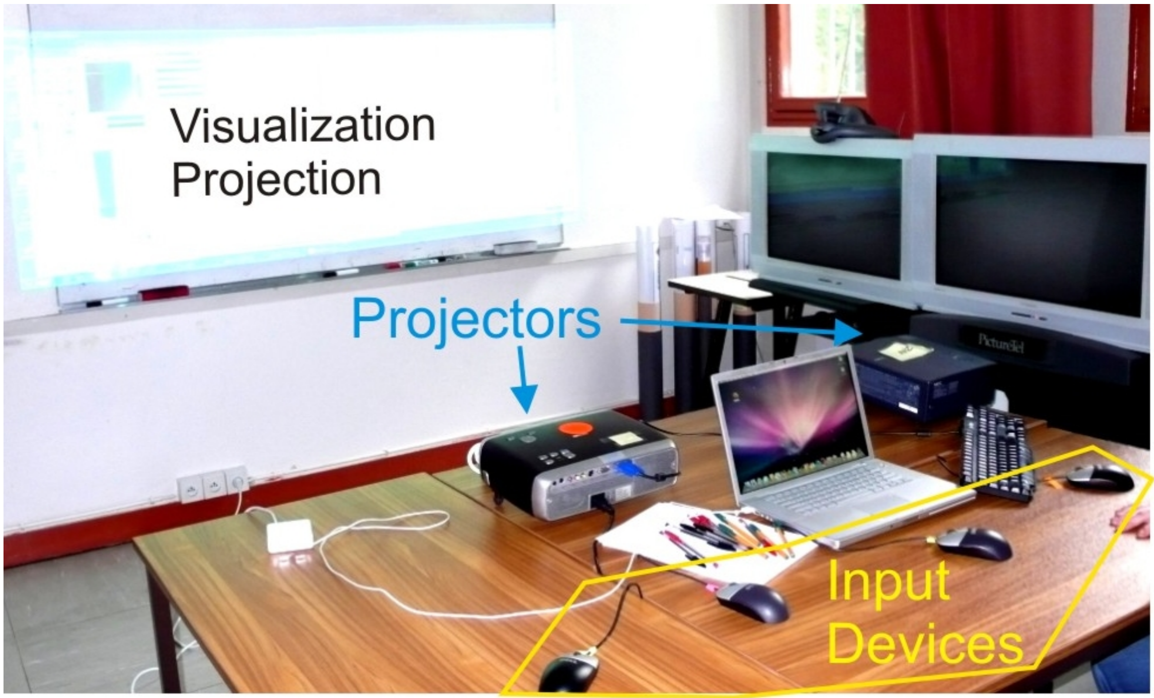


Figure 6.1: An example of a low-cost setup for co-located collaborative data analysis using four mice, two projectors, and a wall for projection.

collaboration. The intention of this work was that from our results and with further research, our knowledge about retrofitting and hence designing co-located collaboration visualization systems will adjust and expand and that this will lead to further design considerations for future collaborative data analysis systems.

6.2 Related Work

The system presented in this chapter was designed to support collaborative social network analysis by retrofitting an existing visualization tool. Thus, this section gives a brief introduction to social network analysis in order to provide context for the tasks supported by our retrofitted tool. Also, to provide background to our retrofitting approach, we briefly outline other efforts in retrofitting visualization tools for collaborative work.

6.2.1 Social Network Analysis

Any collection of people or organizations connected by relations is a social network. In the last decade, the popularity of social networking applications has dramatically increased. Social network analysis tools are used by intelligence agencies to monitor terrorists networks (Yang et al., 2006), by epidemiologists to study transmission networks and to detect and contain disease outbreaks (Krebs, 2008), or company managers and research institutes to examine the flow of communication between their employees or the strength of their employees' collaboration (IBM, 2007). In our work, we focus on visual analysis of social networks that is more exploratory in nature (information on statistical and structural analysis methods can be found in an overview paper by Wassermann and Faust (1994)).

With the increasing popularity of social networking and the progress of Internet technologies, many systems have emerged to visualize and analyze social networks (Henry et al., 2007). Visualizations are used in the social network analysis field to analyze how people communicate and collaborate, what information they exchange, or what role they play in the social group. The two most common types of representations are node-link diagrams and matrix-based representations (Freeman, 2000). Node-link diagrams are commonly used to understand the global structure of the network while matrices have been shown to improve readability for detailed community analysis (Henry et al., 2007).

From trial demonstrations of social network analysis software, we have empirical evidence of spontaneous analysis sessions of co-located colleagues that came together over a small shared display to make sense of, discuss, and explore their data. Similar observations were reported by Heer and boyd (2005) in their study of Viszter, a visualization tool for online social networks in a public setting. Social network analysis can benefit highly from collaborative analysis through the combination of knowledge, expertise, and skills as well as the combined cognitive power of several analysts that can tackle larger networks together. These observations and benefits motivated us to retrofit a tool for this type of collaborative analysis work and data.

6.2.2 Collaborative Retrofitting

The possibilities to connect several mice, keyboards or other input devices to one desktop computer is limited due to support issues at four levels:

1. Operating systems: some systems such as Windows explicitly limit the support for multiple-mice and keyboards due to security issues. Others (including Linux, most flavours of Unix and MacOS) allow the management of extraneous input devices but with a different level of support than the standard input devices. For example, these systems do not provide any cursor feedback for extraneous positional devices so this capability has to be done by applications or window-manager extensions.

2. Low-level libraries for access to USB devices or game devices allow the reading of input devices in system-dependent ways. In the recent years, there has been some progress in trying to standardize access to these libraries with projects such as JInput for Java (Jinput, 2008). There are issues raised by these libraries because the window manager applies many hidden operations to the standard input devices (acceleration management for relative positional devices, key mappings for keyboard devices). These are difficult or impossible to emulate through external libraries, except when integrated with the window systems (e. g., the X Input Extension (`http://en.wikipedia.org/wiki/DirectInput`)).

3. Graphical Toolkits such as Swing for Java or Qt for C++ provide support for GUI components (Widgets) and input managements. Like most of the toolkits, they only manage a limited set of input devices through typed events. Even for well supported devices, like the mouse, they usually do not support more than one reliably. Only recently have there been attempts at supporting multiple input devices at this level (Hourcade et al., 2004; Dragicevic and Fekete, 2004; Huot et al., 2004; Tse and Greenberg, 2004).

4. Applications like MMM (Bier et al., 1992), supporting co-located collaboration, have been built from scratch due to the lack of toolkit and library support. However, newer generations of co-located applications have been trying to build toolkits or rely on special toolkits to simplify the design of these types of applications.

Some researchers have described their process of retrofitting single-user applications for collaborative use; however, only few have specifically studied this in the co-located information visualization context and considered the implications of offering multiple independent inputs.

Forelines describes collaborative retrofitting for Jmol for molecular visualization (Forlines and Lilien, 2008) and Google Earth (Forlines et al., 2006). Both tools were adapted for a multi-user and multi-display environment. Their research focuses on describing how the visualization was adapted to be shown and interacted with in a co-located scenario using different views on different display configurations. Both projects have been previously described in Chapter 2.

Comparing distributed and co-located information visualization work, Mark and Kobsa (2005) studied collaborative use of pre-existing information visualization tools and found that group performance increased with the transparency of the system. Collaborative retrofitting for this study was minimal. While a large shared display was used in the co-located setting, participants also shared a single input.

Some graphical toolkits built for managing scene graphs (e. g. Jazz by Bederson et al. (2000)) or information visualization (Fekete, 2004; Heer et al., 2005), use the Interactor abstraction to implement modular interaction techniques. They decouple display management and interaction, simplifying the retrofitting for multiple inputs. Moreover, they provide support for a layering mechanism on which to draw additional cursors and highlights without interfering with the standard display management. We implemented our extensions in Java with the Infovis Toolkit (Fekete, 2004) in which NodeTrix is implemented.

6.3 Collaborative Retrofitting of NodeTrix

This section includes a short introduction to NodeTrix (Henry et al., 2007) and why it was chosen as a potentially good candidate for a retrofit to a collaborative work environment.

6.3.1 A Short Introduction to NodeTrix

NodeTrix (Henry et al., 2007) is a hybrid visualization in that it combines a node-link representation and an adjacency matrix-based representation of a social network in a single view. This makes it possible to view all data entities represented as nodes and all inter-node relationships as links. Alternatively, one can view all data entities as labels in rows and columns in a matrix and their relationships as the matrix cells. Most importantly, the two representations can be used in combination, with part of the data presented in either node-link or matrix form (Figure 6.2). Whether a particular entity in the data is shown in either of these representations is interactively controllable. For instance, one can group node-link data entities to form a matrix, or select a data entity and drag it into or out of any given matrix. This interactive dual representation combines in a single view the benefits of node-link diagrams and adjacency matrix-based representations, and is conducive to visual data exploration. Figure 6.2 gives an overview of the visualization in which communities within a computer-science department are grouped together in matrices and connected by links representing co-authorship relations.

6.3.2 Choice of NodeTrix for Collaborative Work

To explore collaborative retrofitting of existing information visualizations, we wanted to begin with a tool that seemed to be a promising candidate in its existing state. Thus, we first looked at the considerations discussed in Chapter 3 and 4. According to these considerations we found a promising candidate in NodeTrix (Henry et al., 2007). Specifically, it supports:

- *Free categorization of items*: Nodes can be grouped into matrices with a lasso gesture. Single matrices can be dissolved with a single click. Nodes can be added to or removed from matrices with drag-and-drop. Hence, work on a given item can be done independently from work on others. This could support concurrent, parallel work.

- *Free workspace organization*: Data items can be freely repositioned. This allows individuals to work on the task in different areas of the display and independently

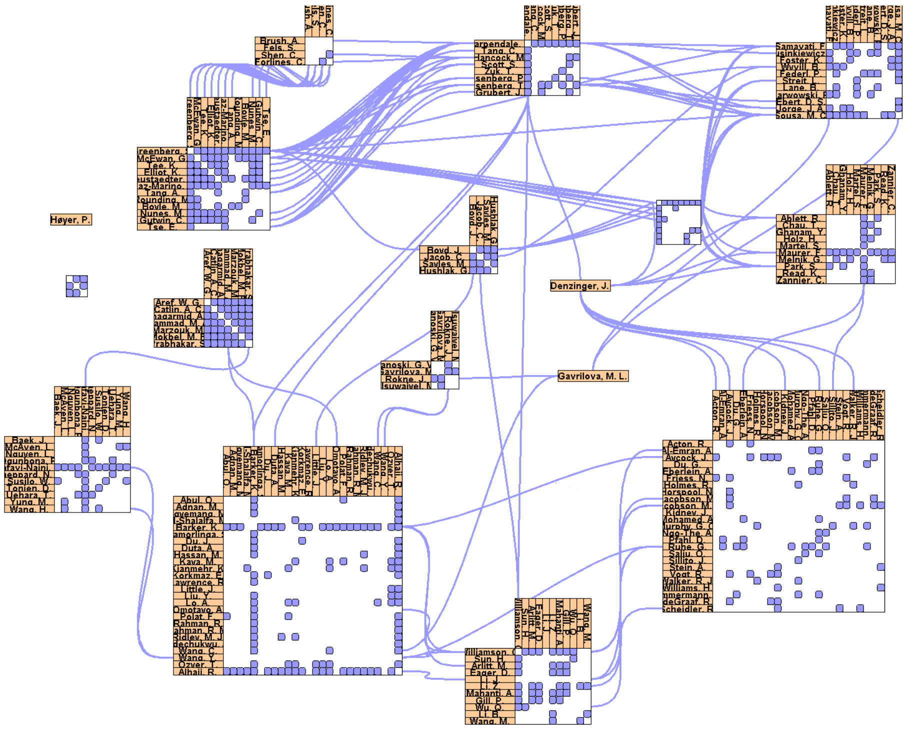


Figure 6.2: NodeTrix Visualization integrating node-link and matrix visualizations. This image shows the co-authorship network of a university department in which research labs have been grouped into matrices.

of one another. In addition, it allows group members to organize information so that they can work more closely together.

- *Individual viewing preferences*: Through a number of local changes in the representation, individuals can adapt parts of the representation to their own preferences.

- *Fluid interaction*: The number of changes of input modality, the manipulation of interface widgets and dialogs is kept to a minimum and can improve the coordination of activities within a group.

- *Focus on mouse interaction*: Almost all actions are mouse-interactive, which makes the tool accessible to retrofitting for multiple inputs. The keyboard is only required for three tasks: to type labels, to trigger undo & redo, and to trigger a graph re-layout.

- *Minimal global changes*: NodeTrix includes only two possibilities for global changes, limiting the possibilities for accidental changes that affect all group members. This may lead to less interruption of the group work.

In addition, several practical aspects of NodeTrix made it a good candidate for our work. It has previously been used successfully with experts in the context of social network analysis and has been shown to be useful in single-user work (Henry et al., 2007). We also had access to the underlying source code and could make necessary adjustments to introduce concurrent inputs. We nick-named our retrofitted co-located collaborative NodeTrix—*CoCoNutTrix*.

However, some of the considerations as outlined in Chapter 3 and 4 are not specifically supported. There is no specific support for communicating findings or discoveries, solving conflicts of interaction, graphical history, or maintaining individuals' awareness of each other's efforts. Thus, while NodeTrix presents a promising starting point, it is not clear whether a retrofitted version will help group members to collaborate effectively. Through an observational study and interviews we explored how participants utilized our retrofitted collaborative software and whether such minimal retrofitting sufficiently supported collaborative data analysis. We were specifically interested in seeing how the following questions applied in the context of parallel and joint work phases:

- Is communication between analysts enabled?
- Do interaction conflicts occur that hinder the collaboration?
- Can group members stay aware of each others' work?
- Are group insights achieved?
- What is the qualitative analysis experience with the system?

6.3.3 Implementation Details

To implement CoCoNutTrix, we made adjustments to the underlying source code. We kept all our re-implementation choices to a minimum. Wherever possible we opted to leave things as they were, as our goal was to study whether a minimal retrofit would accrue collaboration benefits.

General Collaboration Support

One of the challenges in re-designing software for collaboration is that global changes should be kept to a minimum to avoid interrupting group work. Yet, many information visualization systems, NodeTrix among them, offer a high number of parameters to change the visualization output. In our retrofitted tool we turned off menu bars and control panels and chose appropriate default values for all visual features such as link width, colour, or label size appropriate for our task and dataset. The defaults and available functionality were chosen to work well with the type and size of dataset that we were working with. Other tasks and datasets may require different defaults.

Since the main current operating systems do not support multiple windows to be in focus we chose to provide a fullscreen visualization environment, in which no accidental resizing, repositioning, or a change of focus of the application windows could occur. Since all control panels were turned off already this was achieved by giving all available screen space to the rendered visualization.

Adding Multiple Inputs

In NodeTrix, since mouse interaction is the most common type of input, we decided to give each collaborator their own. On the other hand, keyboards were only used for three relatively rare interactions (labelling, triggering a global re-layout, and undo & redo) and take up a lot of physical space on the table, we decided to provide one shared keyboard.

To capture independent input from any attached mouse, we used the JInput library (Jinput, 2008) and added a GlassPane, a transparent panel, on top of the application to render the additional mouse cursors and dispatch modified mouse events to the application. We derived a new mouse event class that carried individual mouse ids in addition to the traditional mouse event data. These ids were necessary to be able to react to person-specific input. In addition, user-specific data structures were put in place to keep track of which items were being drawn or dragged by which mouse. For example, the lasso gesture was used to select multiple nodes. To capture this gesture it was necessary to save a mouse path per person. Appendix A.3.1 contains some sample code detailing modifications to mouse events and selection that were necessary to accommodate multiple synchronous inputs.

In keeping with the spirit of making as few changes as possible in our retrofitting and because it has been suggested that social protocols are often an effective conflict resolution method (Tse et al., 2004), we chose to leave the resolution of conflicts to these social protocols.

Changing Representation and Interaction

We made three changes to visual representation and interaction: (i) We provided additional visual feedback. To differentiate the available mice, each cursor was enlarged to 50×50 pixels and received an individual colour. Click or drag interaction from these mice created a similarly coloured glow effect on each clicked node or matrix. We extended the rendering code for both objects and rendered a coloured semi-transparent rectangle on top of them to achieve this effect. (ii) We changed keyboard input for matrix labels. Previously, labels were created by selecting a matrix and typing the desired text. When several mice are attached to the application, several matrices can be clicked on and in focus at the same time and, thus, it is unclear to which one a label should be added once a group member starts typing. To circumvent this problem, we created a new label object, representing the label text. This object was added to the visualization after a team member finished entering text. It could then be dropped on a matrix to create a label. (iii) We mapped several operations that were previously accessible in control panels, to a mouse gesture. To allow zooming in and out of rendered matrices, for example, we mapped the resizing action to the mouse wheel, a simple fix to address the previously mentioned problem of several matrices being possibly focused on concurrently. All the interactions were implemented using Interactor objects. This allowed decoupling the interaction from the visualization rendering and from the logic of the application. This feature of the InfoVis toolkit made the retrofitting easier.

Retrofitting Cost

Estimating the retrofitting cost is difficult as it relies on the developers's knowledge of the underlying code and the number of places to edit. As an indication of the amount of code, we created ten classes and wrote less than a thousand lines of code to retrofit NodeTrix. NodeTrix is based on the Infovis Toolkit ($\sim$ 750 classes, $\sim$ 65 000 lines of code) and contains around 50 classes and 10,000 lines of code. We only extended the classes in charge of the interaction and created a number of classes to detect and draw the multiple mice. The retrofitting was mainly conducted by the author of this dissertation, an expert in Java but new to both the application and the toolkit it used. It took one week full-time to retrofit NodeTrix with help from the main developers of NodeTrix and the InfoVis toolkit.

6.4 Study

The goal of our study was to determine whether our retrofitted version of NodeTrix could support collaborative social network analysis in realistic settings and to examine how groups viewed our cost-effective design decisions. We strove to provide a study environment as close as possible to (a) real environments, (b) using real data, and (c) with data experts who are (d) performing real social network analysis tasks.

We studied groups of four experts performing social network analysis using data from their own organization. Our participants were experts in the data, not social network analysis experts. To ensure that our collaboration setup was effective in different realistic settings, the study was conducted in three organizations (Org A, an educational institution, Org B and Org C, research organizations) using existing technical facilities.

6.4.1 Social Network Data

Our three organizations have an interest in determining how their internal research groups collaborate and how effective these collaborations are. We, therefore, decided to use research collaboration social networks as data for our study. Given that research publications are a good indication of collaboration, the co-authorship network of each organization was used as a dataset. Authors in the dataset became nodes of the network, and co-authorship relationships became links. Each institution had a high number of authors (exceeding 800 in all three), making the analysis difficult to complete in less than one hour. To ensure a whole experimental session could be concluded in

Org.	Screen Size	Resolution	Projectors	Distance	Figure
A	1.5 m × 1.1 m	2048 × 1536	2 × 2	1.0 m	6.3, left
B	4.0 m × 1.5 m	2560 × 1024	2 × 1	1.5 m	6.3, right
C	2.0 m × 0.8 m	2560 × 960	2 × 1	2.0 m	6.1

Table 6.1: The physical study setup in the three organizations.

approx. 1.5 hours, thus making it easier to recruit knowledgeable experts with limited available time, we filtered out authors with a low number of publications. This resulted in 423 authors for Org A, 327 for Org B, and 430 for Org C.

6.4.2 Participants

44 participants (14 female) took part in our study. All had been with their organization for at least 6 months and were experts in either parts or the entire social network they were asked to analyze. Their positions included senior professors/researchers, group and project leaders, administration personnel, human resources personnel, technical personnel, and few graduate students. We recruited four groups (16 participants) in Org A and Org C, three groups (12 participants) in Org B. To ensure a realistic and comfortable collaborative setting, participants were either work collaborators or friends. With the exception of only one person, all participants reported to be familiar with their group.

6.4.3 Apparatus

Resources in the organizations differed slightly, but an effort was made to keep the settings as similar as possible. The same visualization software ran on a dual core 3GHz CPU, with 2GB RAM, running Windows Vista. In each setting, the four physical mice were colour-coded to match their respective cursors on the screen. The details of our physical setup can be found in Table 6.1.

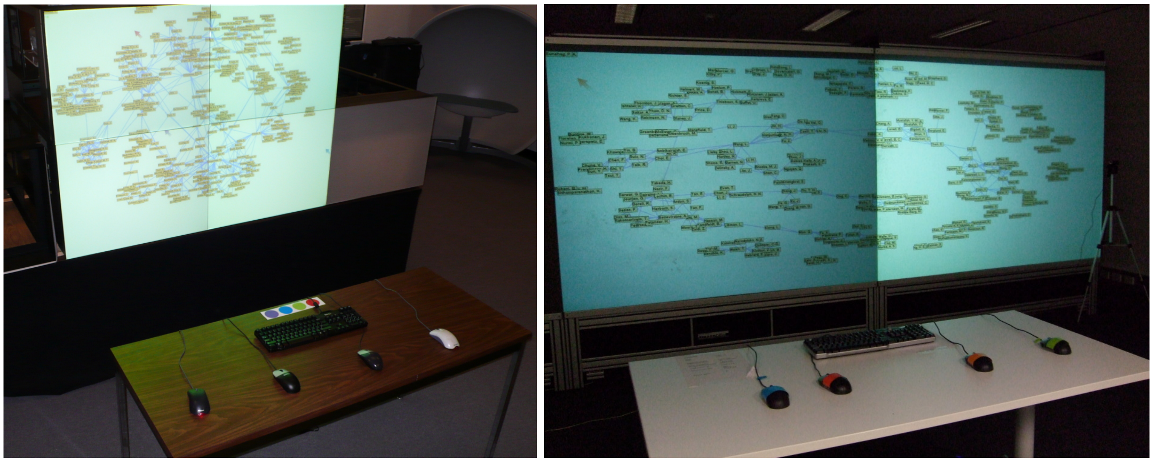


Figure 6.3: Study setup in Org A (left) and Org B (right) using display and computer resources available at each organization.

6.4.4 Task

Participants were presented with a visual representation of a social network that they had intimate knowledge of in terms of: actors (i. e., researchers), their roles and positions in the organizations, and their working relations. Participants were asked to create a representative view of the researchers in the organization that could later be printed in poster form. They were provided with a single shared network representation using a force-directed layout (LinLog as presented by Noack (2005)). For this task they were asked to identify and name the different communities, defining their own criteria. This type of open-ended task of identifying communities and examining their connections is commonly performed in social network analysis (Wassermann and Faust, 1994).

6.4.5 Procedure

Each study session lasted approx. 1.5 hours. Participants were asked to complete a brief demographic questionnaire eliciting their background, their familiarity with the rest of the group, the dataset, and their experience in using social network software (for the full questionnaire see Appendix A.3.2). They were then introduced to the NodeTrix collaborative system and were allowed to experiment with it for 15–20min on a training dataset. After reporting to feel comfortable using the system, they proceeded into the main task of organizing and labelling the co-authorship social network of their organization. The task ended when they completed their labelling and grouping of the network, or when they reached the 40min mark. After a short break, the entire group took part in a semi-structured group interview eliciting their opinions on the task and the system. We asked interview questions from different categories: information content, collaboration experience, awareness and history, work process, and the display space. For the full list of possible interview questions see Appendix A.3.2) An experimenter was present for the duration of the study to answer any questions.

6.4.6 Data Collection and Analysis

Apart from the pre-trial questionnaire, observations, and interview, a number of other data were collected for later analysis. All sessions were video-captured from two distinct locations, one focusing on the participants and one on the screen. Moreover, detailed system logs were stored for each session capturing individual mouse movements and mouse events with the data items that were interacted with. Finally, an observer was present taking detailed notes on the use of the system and interaction of the participants. There was a different observer for each organization, due to the geographical distance between the three locations (Paris, France; Calgary, Canada; and Sydney, Australia).

To analyze our data we first combined the observations and notes from the three observers in an online document and met via video conference to discuss our joint observations. Next, each observer created transcripts of the interviews conducted at the end of each study and included these in another online document. Each observer used their own preferred transcription tool; I used the same as for the study discussed in Chapter 4. We then combined our information from transcribed interview data, notes, and observations in affinity diagrams to reveal patterns in the data. We created these affinity diagrams using an online spreadsheet for remote collaboration. We call this form of remote affinity diagram an aSynchronous Affinity Union List. Affinity diagram categories included topics such as group characteristics, work processes, qualitative work experiences, conflicts, or awareness. For example, an affinity diagram category containing mostly interview data was entitled “Would have preferred to do task alone? ” and would include descriptions such as “No, saw biggest benefit in shared knowledge and consensus,” “3 say no, due to sharing of knowledge, being more fun - doing it alone is work in the group is fun. One part does network analysis as part of his work. He would have liked to do a different strategy that the group didn’t adopt and he would have preferred to do the task alone.” Information such as this could then further be coded by identifying different reasons for the preferences such as “fun, knowledge sharing, task strategy.” The coded results for this category, for example, are reported on in the next section. The affinity diagram’s categories as well as the interview questions can be found in Appendix A.3.2. Using the final affinity diagram categories, we compared the results from the three organizations and searched for overlaps in work patterns. I then took system logs from all the study sessions, analyzed interaction patterns and conflicts, and compared them with our observational data. We then combined all of our coded information to derive and describe our study results as discussed next.

6.5 Results

This section presents how our retrofitted collaboration environment provided collaboration support and assesses whether this support was effective. We group our results according to our study questions from Section 6.3.2 and take inspiration from the “mechanics of collaboration” (Gutwin and Greenberg, 2000) or low-level actions and interactions that a collaborative system must support in order for group members to be able to complete a task in a shared manner. Similar to Gutwin and Greenberg (2000), we consider the collaboration to have been effective when activities could be completed successfully, and if no major errors or conflicts arose.

6.5.1 Explicit Communication

In face-to-face settings like ours, the majority of explicit communication is verbal and is the main means to establish a common understanding of the task at hand.

Observations:
We observed frequent verbal communication: in 9 of the 11 groups, lively communication arose around the content of the data, often in phases of joint work between at least

two team members. We observed two types of explicit communication: running commentary and direct discussions. Running commentary was common in parallel work when participants wanted to quickly inform others of an action performed or planned without an intent to start a conversation. Direct discussions were used to directly contribute to social knowledge building during joint work: groups exchanged rational and argumentation regarding actor placement or grouping choices, group members would agree, disagree, and negotiate, building a shared understanding of the network they analyzed.

Since participants were not directly interacting with the display, our system needed to facilitate deictic references and gesturing for communication in and with a group. Participants performed deictic references not only by pointing with their hands at the display and making verbal references, but also by gesturing and pointing *indirectly* with their uniquely coloured mouse cursors. Moreover, they repurposed the system to their communication needs, for example by enlarging an object to attract attention. During phases of joint visual attention, mice were commonly moved to the joint focus area to show that attention was given to a specific information item that was under discussion.

Requests for Improvement:
Participants only requested additional features to support deictic references. This seems to suggest that transitioning from parallel to joint work was hindered by the inability to easily find the object under discussion or question. Three groups asked for a visual feature, such as an individually controlled glow or animation, that could explicitly draw the visual attention of the group to a particular mouse cursor.

Summary:
We observed that our system provided adequate support for intentional verbal communication, facilitated mostly through the face-to-face setting. Participants made creative use of the visual representation to perform deictic referencing, with a few participants asking for better support. One of the goals of collaborative information visualization tools is to allow groups to come to a common understanding of the data through the use of the visualization. Through our observations of instances of explicit communication we are quite confident that this goal was reached but simply highlighting of data items under discussion could further improve the transition to joint work.

6.5.2 Consequential communication, monitoring and group awareness

Information in physical collaborative settings is unintentionally given off by collaborators and by artifacts as they are being manipulated, for example seeing hands move in the space or hearing paper being dragged by others. This consequential form of communication is very important in digital collaborative tasks as well, as it is an important mechanism for gathering awareness information about what is going on, who is working on what, and where others are in the workspace. Having this awareness can help collaborators transition between different types of work styles.

Observations:
We observed four main visual features with which the representation mediated consequential communication and enhanced awareness within the group:

Colour Coding: Our environment provided a single explicit awareness mechanism in the form of uniquely coloured cursors and matching colouring of selected artifacts. This colour coding indirectly indicated to participants areas of the display and specific artifacts that others were focusing on.

Labelling: Participants labelled communities to indicate that they had been analyzed or needed further work, implicitly informing the group of the work to be done. For example, in 9 of the 11 groups, participants would only give a community a name once they felt it was reasonable finalized, while in 2 of the 11 groups, unknown or not finalized communities would be given a predefined default name (e. g., 'unknown 1').

Location: Participants implicitly communicated their decisions regarding communities by placing them at predefined areas of the display. Some groups (2 of 11) used the periphery of the display to place finalized communities, while others used a predefined area of the screen for 'unknown' or 'draft' communities (2 of 11). Although in most cases this placement started out unintentionally, it often became an explicit work practice (e. g., "I am putting unknowns to the right").

Scale: In 6 of 11 groups, matrices representing finished groups were scaled down in size which communicated that they should not be edited further.

Participants generally reported to have been aware of group work processes on the visualization. Yet, we observed several participants stop their interaction for moments at a time and focused their attention on the representation. When asked about this behaviour in the interview, they reported to have done so to gain an overview of what had changed in the dataset, what the group strategy was, and what areas they could work on next.

Requests for Improvement:
One known issue that pertains to awareness is that people easily lose their mouse cursors on large displays (Baudisch et al., 2003). Participants in 6 groups reported to have lost their cursors occasionally, even though we had increased the mouse cursor icon width and height to four times that of the standard Windows desktop and given each cursor a distinct bright colour.

During the interview, some groups (5 of 11) also asked for more explicit ways of labelling and annotating their work to assure that decisions would not get lost in the work process (e. g., changing colours of communities to indicate they are completed, giving matrices specific descriptions like "Do not merge!", etc).

Only participants in 4 of the 11 groups requested a feature for viewing the interaction history of the group, to see each other's actions and the history of a specific area of the network.

Summary:
Although our participants were able to collaborate on the retrofitted setup, half of them felt the coloured cursors did not provide enough awareness of other peoples' actions. Annotation functionality was also requested to mark the state of communities indicating that additional annotation features may be required. However, most felt that although detailed actions were missed, they were globally aware of the group process and progress. Interaction history was not frequently requested maybe due to the task and length of our study. We generally saw the visualization itself being used as the medium to indirectly capture, represent, and communicate the group understanding and knowledge of the communities in the dataset.

6.5.3 Action coordination, assistance, and protection

An important part of effective and fluid collaboration is how collaborators mediate their actions and share common workspace resources. We observed how team members organized their actions to avoid conflict with others and strategies they adopted to efficiently complete their task.

Observations:
Our participants clearly organized their actions in order not to conflict with others. This was achieved by either explicitly dividing the task and working areas through verbal communication, or by observing where others were working. Collaborators worked predominantly individually or in pairs in different areas of the workspace, moving fluidly between closely and loosely coupled work styles. When questions arose or global changes had to be negotiated, all teams came together and evaluated a solution, performing coordinated actions on the workspace. Coordinated actions were also common when participants helped each other out. Such peer aid would either be requested (e. g., "Could you remove X from that community while I . . . " or would be voluntarily offered by observing the actions of others (e. g., "Let me do that").

In groupware systems accidental conflicts of concurrent input can be disruptive and special control mechanisms have been suggested (e. g., Gutwin and Greenberg (2000); Ringel Morris et al. (2004)). Since we chose not to provide any conflict control mechanism, we logged potential sources of interaction conflicts to validate our choice. These included two or more participants grabbing the same node or matrix, or trying to lasso select an item that was currently worked on by another person. These conflicts occurred rarely. In 10 of the 11 groups, a maximum of two conflicts were logged with concurrent dragging actions being the most common one (4×). One group had 7 such conflicts, mostly caused by two people interacting with the same matrix concurrently. When discussed in the interview, none of participants perceived the logged conflicts as problems. Outside of the logging, we observed conflicts dealing with inadvertent dropping of elements in matrices or a participant editing matrices after others considered it finished. All these conflicts were solved socially, and some groups even established rules (e. g., "ask before editing a reduced size matrix" or "if you see labels don't touch it, that's the rule"). When interviewed, participants felt these conflicts were easily solved and did not interfere much with the task.

Requests for Improvement:
Participants perceived little conflicts of interaction. When asked if they would have wished for a mechanism to lock control or indicate ownership of items, all but one group responded negatively.

Summary:
As in previous studies of collaborative work (Elwart-Keys et al., 1990; Mandviwalla and Olfman, 1994) and even collaborative data analysis tasks (e. g., (Tang et al., 2006) or Chapter 4) we observed participants moving between joint and parallel work styles. Our participants coordinated their actions very fluidly. We feel that our choice of not to include specific protection mechanisms was further justified as conflicts were resolved socially and mistakes could be easily reverted through local or global undo.

6.5.4 Analysis Strategy and Group Insight

One of our original goals, was to determine if our discount environment supported successful collaboration with the visualization. An indicator for successful collaborative visualization use is the establishment of an effective strategy leading to group insight. Group insight is difficult to measure (Stahl, 2006), but can be visible in interactions between participants and with the visual representation, or interview comments like "we found out that . . . "

Observations:
Although no explicit planning support was given in our environment, most of our participants verbally negotiated their strategies. Almost all groups (9 of 11) started the task with a short group exploration phase in which initial obvious clusters were identified. The establishment of an analysis strategy seemed to evolve naturally from conversation and participants observing each others' actions.

When asked, all 11 groups reported to have gained new insight from working with the dataset and reported several surprising or confirmatory findings, such as close collaboration patterns between research groups previously thought unconnected, and even findings about their close working environment "I had no idea that many people collaborating in our lab, I even learned things about my own team!" Peer-learning and teaching of these insights occurred often in groups that had an imbalance of shared knowledge. In one group, for example, a participant helped to identify the initial communities and taught others about parts of the dataset they were unfamiliar with, so the work could then commence in parallel.

Summary:
We observed participants smoothly establishing an analysis strategy and they did not request any additional features for activity planning. Observations and comments showed that our tool helped the group gain insight, teach each other facts about the data, and support knowledge building in the group. We see this as an important part of a successful collaborative data analysis environment.

6.5.5 Work preference

As an indication of successful collaboration, we asked participants whether they preferred conducting this analysis task as a group rather than individually.

Observations: The majority (40 of 44) of participants preferred group work and 4 preferred to do the task alone. Three of the latter were among the most knowledgeable members of their group and felt that they could have done a reasonable job on their own, although they all admitted it to be potentially slower. The fourth had a completely different opinion than the rest of her group about what criteria to use in forming communities. The participants who preferred group work named as reasons for their preference: shared knowledge (27 of 44), fun of collaboration (25 of 44), shared process of forming consensus (6 of 44), brainstorming (4 of 44), efficiency (4 of 44), and shared working styles (1 of 44). One participant commented that "doing it with 3 people was fun, doing it by myself would be work." In addition, 9 of the 11 groups reported feeling happy with the result of their analysis and the communities they had created.

Requests for Improvement:
Most participants stated that additional time and meta-information would have helped to resolve questions about unknown people and improve the visual presentation of the analysis.

Summary:
Groups were generally very happy with their collaboration and result of their work. We take this as an indication that the retrofitting was successful for this setting and task and could effectively support collaborative data analysis as perceived by these participants.

6.5.6 Reaction to low-cost environment choices

While observations on collaboration and group insight can establish whether collaboration in our low-cost setup was effective, observations on the usability of the environment can further inform the effectiveness of the retrofitted tool in use.

Observations:
One observed strength of the CoCoNutrix visualization was its intuitiveness of interaction. All participants were at some point interacting with the information items and, over longer periods of time, all mice in a group were in movement concurrently. Participants were comfortable interacting anywhere on the screen. Even though the screen sizes were slightly different, this observation was unaffected. The keyboard as a shared device was typically used by one dedicated scribe who would type in the labels for communities as they were requested. Groups rarely used features that would have created global view changes (undo, redo and a re-layout of the graph), and when they did, it was generally after negotiating and obtaining group approval. Five groups never made use of these functions, two groups used them 6×, and the remaining groups used it 2–3×. Participants commented that our low-cost setup of mouse input and large screens supported their group work well.

Suggestions for Improvement:
Three groups expressed the need for a second keyboard to avoid interrupting others' work process by asking for a label, or handover of the keyboard. There were 15 requests (from all 44 participants) for functionality that was originally part of NodeTrix and was removed during the retrofitting. These requests were mostly for visual features mentioned earlier, such as highlights, more meta-data, or for additional interactions (such as sorting) on matrices. Participants reported they did not feel the sitting configuration influenced their collaboration, but to further improve communication some would have preferred a slightly curved seating arrangement to be able to talk to each other better. In Org C, dealing with a larger network on a slightly smaller display, participants would have preferred a larger screen display or functionality to "push nodes to get more space." Thus, the ratio between the display and network size used in Org C was perceived as a threshold condition for comfortable use.

Summary:
While participants requested additional functionality for the system and physical setup,

they generally reported to have been well supported in their global task. Lack of interaction capability and the lack of meta-data affected their work efficiency, but the work quality was not generally compromised. We see this fact as proof that our discounted interface was a good compromise for this task.

The requested additional visual and interaction features are difficult problems to solve when multiple people interact with the system. Selection actions can induce input conflicts and parameterizing actions requires consensus as they affect the entire representation. This is the reason we removed them originally in our retrofitting, but further research is necessary to reduce global changes in visualizations or make them less disruptive. While the actual sitting position did not seem to interfere with the collaboration, we found that the display size was very important. Finding the optimal screen size for visualization tasks requires further research attention.

6.6 Discussion

To summarize our findings we return to our initial questions in regards to the utilization of our retrofitted collaborative software.

6.6.1 Assessment of the Results

Communication

We observed frequent interaction between analysts, with the data and with the visualization. Analysts slipped in and out of interaction by themselves, with the full group, or with varying subgroups as work progressed. This confirms previous CSCW studies on information visualization in other settings where frequent switching between loosely and closely coupled work was observed (e. g., Tang et al. (2006) and also see Chapter 4). Active data interpretation, discussion, and negotiation occurred throughout the collaboration while participants interacted on all areas of the display. This finding is important as information visualization analysis requires seeing and interacting with all parts of the representation to explore all available data and avoid misleading or incomplete data analysis.

Conflicts

Control mechanisms to avoid interaction conflicts have been studied and suggested (e. g., Gutwin and Greenberg (2000); Ringel Morris et al. (2004)) for co-located collaboration. Even though we included no specific control mechanisms, we observed and logged few interaction conflicts between participants, echoing previous findings (Tse et al., 2004) that people naturally avoid interfering with each other by spatially separating their actions in the workspace. Moreover, participants did not request any additional control mechanism features, so our decision to leave them out was further justified.

Awareness

The visualization mediated the awareness of decisions made about the data and helped group members to build on each others' work. Factors like labelling were used to help the group coordinate which data aspects were decided upon and which were still in flux. Yet, several additional awareness features were asked for and this is a promising direction for further work in collaborative visualization.

Group Insight

The hybrid nature of the visualization helped in facilitating, and hence observing group insight, as it captured the evolving construction of knowledge within the group. We noticed that participants did not simply view a matrix as a different representation of a group of researchers in the dataset—a matrix expressed a particular research group and together with a label became the result or artifact of choices made by one or several participants during the collaboration. This artifact was then visible to others and facilitated the emergence of a common understanding of the data within the group. Thus, the visualization evolved and became an archive of the participants' process, what work was completed or needed discussion, and of the participants' insight, the interpretations and meaning that they had given together to specific information in the dataset. Similar observations have been made by Stahl (2006) for collaborative communication and learning in online communities.

Qualitative Feedback

Both the chosen physical environments (use of a large back projected display and sitting arrangements) and the use of multiple mice for interaction was positively received by our participants. Together with other positive responses and feedback regarding the usability of the system, we feel confident that NodeTrix was sufficiently retrofitted to enable effective collaboration.

6.6.2 Impact for Other InfoVis Systems

The study results have implications for other information visualization researchers or designers considering how to adapt their own single-user applications to co-located collaborative work settings. The remainder of the chapter presents a number of different general considerations for retrofitting based on our study. As our study used a qualitative observational methodology, the presented considerations should be seen as grounded hypotheses that have to be further evaluated.

In Section 6.3.3 we have described necessary changes to allow for multiple inputs. We believe these changes to be generally possible in other information visualization systems. Getting differentiable user IDs may be difficult on most multi-touch technologies (a notable exception being the DiamondTouch display) and some retrofitting techniques for these technologies may have to work around the lack of user IDs. The most difficult changes, however, pertain to the effects of multiple people concurrently interacting with a system. The next sections highlight the sets of common information visualization features that may be most impacted by the introduction of synchronous inputs and, hence, may be most important to consider when retrofitting information visualization tools. We see synchronous inputs as an important part of collaboration support as it allows collaborators to work in parallel.

Global Controls

Global controls can raise problems when multiple people collaborate as they may change all team members' view of the data. When they are working closely coupled this may not be an issue, but during parallel work where each team member may be concentrated on different aspects of the data, sudden changes to the display can be surprising and disruptive. For example, if a shared visualization is too large to fit on the screen, a single team member's panning modifies the view for all of the collaborators which may be attempting to work in parallel. There are a number of information visualization features that are commonly implemented as global controls and their design or use may have to be reconsidered when multiple concurrent inputs are introduced. Such features include filtering, navigation (e. g., pan, zoom), transformations (e. g., changes between different representations, changes of data encodings), or view changes (e. g., projection, rotation). The amount and importance of these features will influence the difficulty of a retrofit.

Several strategies can be followed when retrofitting global changes. The simplest solution is to either remove interface features that permit global changes, implement control policies, or simply let collaborators deal with potential conflicts themselves. We have made good experiences with a combination of the first and last option. Another solution requires the development of novel interaction techniques in the system, replacing the one causing problems. For example, zooming, panning, or filtering could also be restricted to only influence a local scope. For example, using a node-link diagram, we can imagine taking advantage of the topological information of the graph to replace panning by bringing the neighbours of a node into the view (e. g., as in bring and go by Moscovich et al. (2009)). Similarly, to avoid panning as much as possible, lens techniques can be used to shrink areas of less interest and introduce new visual features.

Introducing these more elaborate interaction features may be more important for a number of different visualization systems. For example, in visualizations that cover the whole display, are space-filling, or include 3D views, global changes are often central to the tasks. A space-filling treemap (Johnson and Shneiderman, 1991), for example, relies on zoom and filter operation and a 3D visualization such as ConeTrees (Robertson et al., 1991) relies on global view changes for information to be understood and read in its entirety. In this case, the visualizations may need a more intensive re-design to introduce multiple foci exploration or coordinated views.

Undo and Reversible Actions

Being able to undo an action is important in almost any computer program and also in collaborative information visualization systems. An undo feature has to be associated with a thread of actions. This makes undo for synchronous collaboration a difficult issue to retrofit since several people could affect a single information item successively and the effects of an undo may be harder to coordinate (Prakash and Knister, 1992).

One could simply keep a system's current undo capabilities (like we did in CoCoNutTrix) but then an undo is global and will undo the last actions, no matter which collaborator issued it. A user-specific undo may seem semantically more meaningful but then conflicts may arise when one team member tries to undo an action on an object which has already been subsequently modified by a collaborator. Based on our study, we believe that participants would prefer this personal undo, however, more meaningful ways to deal with undo conflicts would have to be incorporated. In CoCoNutTrix, participants dealt with the problem by reversing their actions manually in the interface (by taking a node out of a matrix or back in). Adding local undo functionality to a system may be a viable alternative approach.

Windows and Dialogs

When multiple people synchronously interact with a visualization system, the connections between interface and information display items may have to be coordinated differently. Reimplementing multiple windows, widgets, or dialogs to allow more than one concurrent interaction in the system will likely have to be performed. This retrofit can be difficult but is crucial to allow for synchronous data exploration and manipulation. The difficulty of this retrofit depends mostly on the underlying windowing toolkit. It applies, in particular, to multiple-coordinated-view systems or other information visualization tools that depend on a high number of dialogs and widgets. The difficulty in retrofitting dialog boxes lies in making the right connection from an interface item to the data that is to be affected. For example, on which data item should a colour change apply if multiple items are currently highlighted by different people? One possibility is to track an item+widget selection through input IDs but this would effectively hinder collaborators to team up and apply changes together—one highlighting the items, the other operating a dialog box.

In CoCoNutTrix, we made a more fundamental decision. We removed all dialogs from the interface and made only a subset of features available through direct mouse control (left click, right click, scroll wheel, gestures). This effectively made all interaction local to the information items. We made a fundamental choice between two options: either to provide all possible features of the original tool which would potentially result in more interaction conflicts or to reduce the number of features and minimize them for the task at hand. This is a choice that similarly has to be made by other people who are thinking about retrofitting an information visualization tool. One has to consider what task the retrofitted tool should be used for. Is it meant to be as powerful as the original tool? Or, alternatively, is it going to be used by a group to answer more specific questions with specific tasks that are better solved by a group than an individual? In CoCoNutTrix, participants solved the task well with our minimal set, so our choice was justified. However, some of the original features were missed, so in a second redesign phase we would reconsider this design choice. Also, for a different task or dataset we might have to do another retrofit and consider remapping interface gestures and inputs to other parameter changes leading to more retrofitting effort on our part.

Awareness Features

One very important aspect that a retrofitted tool needs to support is the awareness of what has been looked at, analyzed, and about which data items decisions have been made. In our case, this was mostly facilitated through the hybrid nature of the visualization. We, therefore, hypothesize that information visualizations in which group members can give the data meaning by either transforming data items into different representations (as in our case), or by annotating and marking them (e. g., through spatial positioning or graphical markers), will not require much additional functionality to be added. It is likely that a large number of other 2D network and graph visualizations can be easily retrofit in regards to awareness features. Most such tools already allow for free spatial repositioning, which could be used to annotate or mark data by changing their position. Coupled with user-specified visual clustering, group insight could be captured and group coordination and communication supported. Other systems should consider adding capabilities for data annotations. These could simply include adding colourful labels, highlights, or the ability to reposition or reorganize information items, if information is not already otherwise encoded in this manner.

6.7 Chapter Summary

CoTree from Chapter 5 supported parallel work with multiple personalized views of the same data. In this chapter, I contribute insight on a different approach. Here, a system that initially supported more joint work through a single input and one single data representation, was retrofit to introduce mechanisms for parallel work.

We retrofit a social network system, NodeTrix, for synchronous collaborative work which represents the social network across the full screen. The question arises how team members coordinate their interactions across one large shared representation and whether parallel work would still occur, similar to what we observed in Chapter 4. Our study results showed that our retrofit environment was reasonably functional and that it allowed participants to work well in parallel and more closely together. Even though we had not designed any specific conflict resolution mechanisms, participants working in parallel hardly ever issued conflicting commands and rarely attempted to work on one specific matrix at the same time. Due to the hybrid nature of NodeTrix and the task given, team members were able to track team members' interactions with the data by observing how the visualization itself adapted and evolved during the task. In this sense, the visualization itself captured the groups' progress and was used as a medium to communicate decisions and group insight among the individual team members. Nevertheless, participants asked for additional awareness information to help them more easily be aware of where each others' work overlapped. This finding led us to further investigate this issue of joint awareness in co-located collaboration. The resulting work is presented in the following chapter.

In regards to the results from the study presented in this chapter, we caution that the overall success of retrofitted collaborative software is very dependent on an identified set of interaction capabilities of the existing software. To refine our results and to be able to make further recommendations for retrofitting, in particular in relation to the need to support parallel and joint work styles, it needs to be studied how other types of visualizations fair in a retrofitted scenario, and how they are used in real-life situations where the outcome of the analysis has a big impact on participants' everyday work. This work and our previous considerations for collaborative information visualization can be a useful starting point.

Chapter 7

Cambiera: Collaborative Visual Analytics in Document Collections

In this chapter, I introduce Cambiera, a collaborative tabletop system for the analysis of text document collections (see Figure 7.1). With this last case study, I look more closely at the issue of providing awareness to team members who are working in parallel in order to encourage group discussion, negotiation, and shared knowledge building. The work is based in particular on findings derived in Chapters 4 and 6. The systems presented in these two chapters did not include specific visualization based mechanisms to encourage the transition between different work styles. The possibility to easily divide the work and to freely organize, move, and transform information items encouraged group members to analyze data in parallel. Particularly, the fact that CoCoNutTrix included implicit mechanisms that helped team members to remain aware of what had been worked on helped teams to transfer to more joint activities of discussion or validation of their findings. Cambiera is an attempt to more specifically incorporate awareness support to encourage transitions between parallel and joint work during data analysis.

With Cambiera, we[1] designed a tabletop system that allows teams of up to four people to collaboratively search through document collections. It incorporates the possibility to divide the work, supports free arrangement of documents and search results lists, and displays indicators of search and read activities in each person's view of the data.

[1] Portions of this chapter have been published in (Isenberg and Fisher, 2009). Thus, any use of "we" in this chapter refers to Petra Isenberg and Danyel Fisher.

Figure 7.1: Two team members collaborate around Cambiera, implemented on a Microsoft Surface.

An exploratory study at the end of the chapter shows that this design did indeed support a range of collaboration styles and that the awareness features were seen as a valuable feature by the pairs participating in the study.

7.1 Collaborative Brushing and Linking

With Cambiera, we explore the idea of trading off individual and group work in a context of collaborative visual analytics. Our work concentrates on situations in which small groups of people come together in face-to-face meetings to make sense of textual documents. Our goal is to support both the individual and the group in collaborative foraging activities including searching for, reading, and extracting information from visual representations. By providing visual awareness of individual activities we hope to encourage group discussion, negotiation, and shared knowledge building. We use an interactive multi-touch table (Figure 7.1) as the single shared display on which foraging activities are performed. The particular task we discuss in this chapter is that

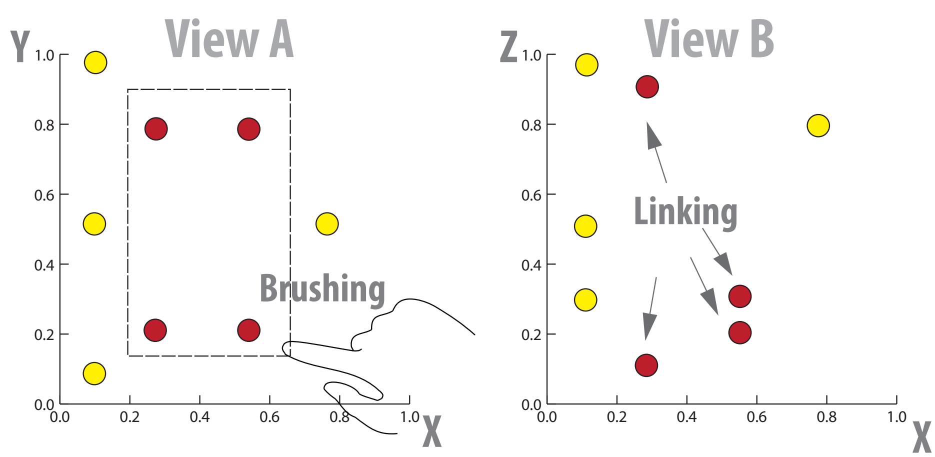


Figure 7.2: Example of a classig brushing and linking scenario. Two scatterplots of the same three-dimensional dataset are shown. View A shows the *x/y* dimension and View B shows the *x/z* dimension. Four data items in A are brushed and highlighted and this interaction is reflected on the corresponding items in View B.

of analysts attempting to find information relevant to reconstruct a story from a large set of textual documents.

We introduce the concept of *collaborative brushing and linking* as the basis of an implicit visual awareness technique across multiple team members for information visualization settings. The idea of "brushing and linking" (Buja et al., 1991) is now seen as a standard technique in many information visualization contexts. "Brushing" refers to the selection (and often subsequent highlighting) of data items in one view; "linking" means that changes in state to a datapoint are reflected in all places that the datapoint is shown. An example is given in Figure 7.2.

Similar to this need for integration and linking of individual views in applications for single analysts, we define *collaborative brushing and linking* as:

> *An awareness technique, in which the interactions of one collaborator on a visualization are visible to other collaborators viewing the data items in their own visualizations or views of the data.*

This concept augments traditional brushing and linking with additional information about the social data analysis process. Collaborative brushing and linking allows team members to communicate implicitly, by sharing activities and progress between visualizations and, thus, may help them to transition more easily between individual and joint activities. In asynchronous situations, collaborative brushing and linking essentially becomes interaction-based annotation: it is a way of labelling data with interaction information for other team members to consider. Here, we consider synchronous work, where collaborative brushing and linking is an awareness tool, communicating current work. During synchronous work, a collaboration must maintain a balance between individual and collaborative activities. The collaborators must be able to *share knowledge* with each other, establish *common ground*, and *reduce redundant work*. Cues that help to support these tasks can include metadata on what information has been looked at and by whom, when and how long information was read, or it could capture additional information, such as importance and reliability indicators, that collaborators wish to add to the data.

7.2 Background: Visualizing Awareness

When collaborating in shared workspaces, people naturally make use of a variety of cues to learn about what their teammates may be working on (Olson and Olson, 2000). By explicitly using verbal information and deictic references, items and actions in the workspace can be specifically referenced while facial expressions, gestures, body postures, or other auditory feedback may implicitly reveal cues about what a particular person may be working on at a given moment. How to support collaboration through design of awareness features is an active research topic in Computer-Supported Collaborative Work (Carroll et al., 2006). This section discusses techniques developed by CSCW researchers to support workspace awareness as well as techniques that have been developed to display interaction history in information visualization applications. These are related to the concept of “collaborative brushing and linking” as is used in the Cambiera project.

7.2.1 Workspace Awareness in Collaborative Work

The topic of awareness has received considerable past research attention by human factors researchers and, in particular, in the area of distributed collaboration. Several different types of awareness have been discussed, including, for example, informal awareness (e. g., Dourish and Bly (1992)), conversational awareness (e. g., Clark and Brennan (1993)), structural awareness (e. g., Leland et al. (1988)), or workspace awareness (Gutwin and Greenberg, 2002). Closest to our goal of providing awareness of interaction with information items in the workspace, is the concept of workspace awareness.

Workspace awareness has been defined as the "up-to-the minute knowledge a person uses to capture another's interaction with the workspace" (Gutwin and Greenberg, 2002). It includes information about the workspace itself, team members' locations within in, as well as their activities and intentions relative to the tasks. Gutwin and Greenberg (1998) discuss workspace awareness for distributed collaboration and present several techniques designed to increase workspace awareness for people working synchronously from remote workstations. Their techniques include split workspace views, feedthrough and action indicators, as well as multiple vs. shared consistent representations for distributed work scenarios.

In co-located collaborative scenarios, workspace awareness is supported by *territoriality*, *artifact feedthrough*, and *consequential communication* (Nacenta et al., 2007a). As our focus is on supporting awareness for this type of collaborative work scenario, we briefly discuss how these three forms of awareness relate to our work. Research on collaboration in tabletop workspaces has shown that group members divide their work areas into personal, group, and storage *territories* (Scott et al., 2004). A team member's increased use of personal spaces to perform individual work may hinder workspace awareness. This may particularly be the case when these spaces are spatially separated across a large display, and interactions require detailed work and the manipulation of small data items. As we have frequently observed detailed work in personal, separated parts of the workspaces (see Chapter 4, 6, and Tang et al. (2006))—support for workspace awareness may be particularly useful for data analysis. *Artifact feedthrough* includes feedback sent to others in the workspace through actions on workspace artifacts. For example, hearing a piece of paper being dragged across a physical table is considered auditory artifact feedthrough, as others receive feedback and awareness information of the dragging action (Baker et al., 2001). In virtual workspaces, actions on artifacts may go unnoticed as they rarely produce auditory feedback themselves, and because actions may be quite sudden and ephemeral. Collaborative brushing and linking addresses this problem and offers a type of artifact feedthrough by making interactions with an artifact directly visible on other copies of this artifact. This information is also made persistent so that it can be used at a later point when needed. *Consequential communication* is a type of feedthrough that communicates information about interactions in the space through the movements of arms and bodies. In co-located virtual workspaces with direct-touch capabilities, feedthrough about who is manipulating a workspace artifact is naturally available. However, one can design interaction mechanisms that do not require one to reach and directly touch workspace items. For example, Nacenta et al. (2007a) discuss several direct and indirect techniques and their influence on awareness in tabletop workspaces. They conclude that direct-touch drag-and-drop interaction had the best all-around performance in their study and ranked highest in preference. They also argue that the use of arms and hands over the workspace offers the most awareness cues to the group in terms of consequential communication. The Cambiera project makes use of direct-touch interaction and does not use indirect object manipulation techniques such as laser pointers or radar views. Therefore, consequential communication about general interactions in the workspace should be supported.

7.2.2 Visualizing Interaction

A general set of questions that one may have about interactions in the digital workspace has been previously proposed (Gutwin and Greenberg, 2002; Tam and Greenberg, 2006): *where, who, what, how, when*, and *why*. Five of these categories were initially used by Gutwin and Greenberg (2002) to describe awareness for real-time groupware and extended by Tam and Greenberg (2006) to cover asynchronous change awareness. A subset of these questions has been previously considered for the visualization of people's interaction with an information visualization system. However, visualization of interaction history is often done retroactively, based on log files created while a person was interacting with a graphical user interface (GUI) on a client or corresponding with a server. HotMap (Fisher, 2007) is an example of such a visualization of Microsoft's Live Search Maps usage, showing—in form of a heatmap—which map tiles had been loaded by viewers of Live Maps. Willett et al. (2007) discuss the visualization of social activities in terms of *information scent* (Pirolli and Card, 1999) for social navigation (Dourish and Chalmers, 1994). Their work on Scented Widgets embeds small visualizations into standard user interface widgets to aid navigation and show possible avenues for discovery. While Scented Widgets focus on supporting navigation cues, our work uses the dynamic and changing information created during synchronous interaction in a shared workspace to offer real-time awareness of team members' actions. Despite this difference, our techniques may similarly serve as social navigation cues in co-located workspaces. Similar to the goals discussed by Chuah and Roth (2003), it may also serve to create common ground during collaboration by providing artifact feedthrough.

7.3 Cambiera: A Tool for Co-located Collaborative Information Foraging

Our tool, Cambiera, is designed for information foraging in the domain of visual analytics (Thomas and Cook, 2005). In the following sections, I describe the data and task used when developing Cambiera, its general system components, and then discuss the awareness features built into the tool in detail.

7.3.1 Data and Tasks

The task we are interested in exploring is that of intelligence analysts attempting to reconstruct a story from a large set of textual data. This scenario is not an original one; rather, we have adapted it from a series of ongoing challenges within the Visual Analytics (VAST) community (e. g., Grinstein et al. (2006)). In the standard VAST scenario, intelligence analysts attempt to decode a large set of documents. Buried within the documents is a single story, spanning multiple documents, that must be reconstructed. For example, a set of newspaper articles may all provide indirect evidence of a crime being planned.

Intelligence analysts often use an exploratory, cyclic process of foraging, evidence gathering, and hypothesis generation. Our system allows analysts to search through a document collection; it visualizes the resulting documents, and allows direct access to document texts in order for the analysts to get detailed information and to form hypotheses. Evidence can be gathered and arranged on the surface to represent information relevant for a specific hypothesis. Our main dataset comes from the VAST

2006 contest (Grinstein et al., 2006), representing some 1200 fictitious newspaper articles. We will describe two fictional analysts working together on an analysis reasoning task, in order to outline features of our tool. The analysts, Ana and Ben are trying to understand an outbreak of BSE, or Mad Cow Disease, in a farming town; they fear this outbreak may be linked to corruption in city hall. Ben is investigating newspaper articles that mention BSE, while Ana focuses on political events.

7.3.2 Implementation

The design of Cambiera is intended to take advantage of the special properties of multi-touch, co-located surface computing. Cambiera runs on a Microsoft Surface interactive tabletop (Figure 7.1) which supports multiple, independent, direct-touch inputs on a horizontal display with a resolution of 1024×768. Cambiera currently supports a small group of up to four analysts, working together. Cambiera was implemented using WPF and the Surface SDK and did not make use of other external libraries. All visuals were created from scratch using features built into WPF and using Microsoft Expression Blend to create specific XAML representations. To implement document searches, we used regular expressions. The details of the implementation are property of Microsoft and I cannot disclose them in this thesis.

7.3.3 General System Description

Although explicit coordination overhead is lower, since team members are in immediate proximity, implicit sharing can relay rich information rapidly. Scott et al. (2004) have suggested that allowing team members to maintain a personal region of the tabletop allows them to negotiate which work is local and which collaborative. However, rigid boundaries may obstruct collaboration: we want to allow team members to flexibly move from loosely coupled to tightly coupled collaboration (to use the model articulated by Tang et al. (2006)).

Our design is also influenced by SearchTogether, a collaborative web-search tool (Morris and Horvitz, 2007). In SearchTogether, remote collaborators share query histories and result lists, allowing them to efficiently search the web in groups. Similarly, we

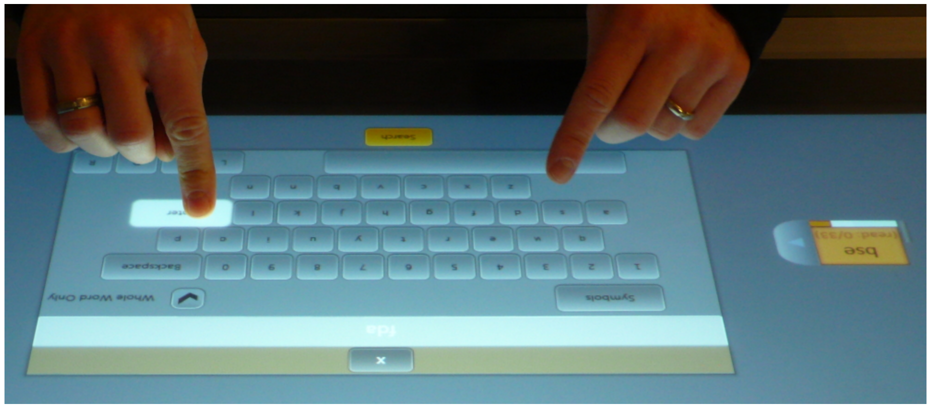


Figure 7.3: Interaction starts with a search. Each team member is assigned a colour, which is reflected in the search button (top) and keyboard (bottom).

want to make it possible for team members to see each others' query and reading histories. Cambiera is distinctive as a visual support for information foraging. We want to support information foraging with information visualization, allowing collaborators to explore different perspectives or subgoals from their own views of the data that are only connected by visual awareness cues. In practical terms, this means that we want to allow Ana and Ben to work independently. If they find something in common, or find that they are working on the same documents, they can work more closely.

Specific functionality of Cambiera is described in the following sections. We use these four questions about collaboration awareness to guide our discussion:

- Did another search also find my document?
- Has someone else issued my search?
- Has someone considered the same document?
- Has someone read the same document?

7.3.4 Presenting Search Results

The core of the system is a search tool for finding documents of interest. In general, an analyst starts off by pressing a coloured search button on their side of the table, which brings up an on-screen keyboard (Figure 7.3). A search issued from each of these keyboards results in a coloured *search box*, which contains the search results.

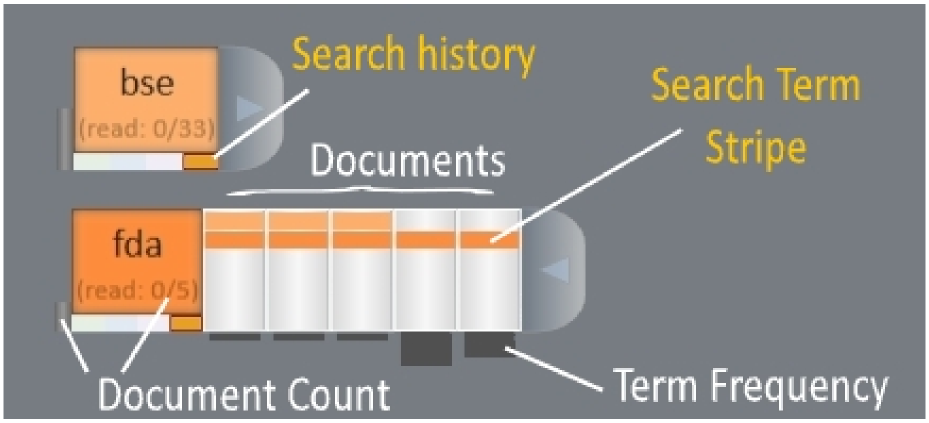


Figure 7.4: Initial search result overview. One closed *search box* (top), and one opened *search box* showing five result details (bottom).

These search results are first shown minimized (Figure 7.4, top). Clicking the arrow tab on the right side expands the set to show the individual results returned by the search (Figure 7.4, bottom). In the expanded view, each gray rectangle stands for one of the found documents. All documents are ordered by their publication date. Dark gray bars below each document indicate how many times the search term occurs in the document in order to give a hint about which documents may be interesting to explore, reminiscent of ScentedWidgets (Willett et al., 2007). On the left side of the search box, we see the search term and a written count of both the search results returned and the number of those that have already been read by any collaborator. A faint bar on the left redundantly encodes the number of documents returned relative to the total number of documents in the collection. The bar is drawn on a log scale to allow easier relative comparison of document count for small result sets. Like all other objects in the interface, the search boxes can be freely moved around and rotated in the workspace. When an analyst is done with a search, it can be deleted.

In our scenario, Ben issues a search for FDA (Figure 7.3); he has already searched for BSE. By looking at the returned *fda* search box, he immediately sees that only five documents were returned and opens the result list to see further details (Figure 7.4). By looking at the term frequency indicators, he can see that the last two documents include the term *fda* several times and decides to explore those first. Now he wants to find out more about the overlap of his two searches.

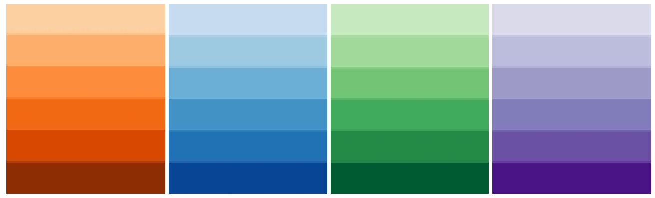


Figure 7.5: Colour scales to encode search terms. Each analyst's searches receive one hue of their base colour.

7.3.5 Did another search also find my document?

If a single document was found by two different search terms, then it may be more likely of particular interest. If two investigators are following different paths and stumble on the same results, that source may have particular salience. The following discussion shows how collaborative brushing and linking can help analysts to follow each others' searches. In Cambiera, collaborative brushing and linking requires that the system be able to track some information about which team member is interacting. While our surface cannot distinguish between different people's input (unlike the DiamondTouch (Dietz and Leigh, 2001)), the strategy carried out in this system helps maintain identity. Each team member gets a palette of colours that are all variations of one hue (Figure 7.5); each search gets a distinctive colour in that palette. In our example, Ben's searches are orange and always distinctly different from Ana's blue searches. Each document representation that is hit by a search is tagged with the distinct search colour. For example, in Figure 7.4, the background behind the word *bse* received a specific shade of orange to mark this search term. Each document in the search results received a coloured stripe in the colour of that keyword to indicate that this particular keyword was found in the document. By looking at the stripes (Figure 7.4), we can see that the first three documents in the *fda* set were found both by the results for *bse* and *fda* while the last two were only found by *fda*. This provides a simple way for doing AND/OR searches. The presence of several stripes indicates an AND combination while showing two search representations next to each other allows to see documents that contain either keywords. While our current implementation supports only simple keyword queries, the concept can easily be extended to more sophisticated queries.

By sliding a finger over the search results, the document under the finger is slightly enlarged for a simple lens effect and annotated with details. In Figure 7.6, we see Ben

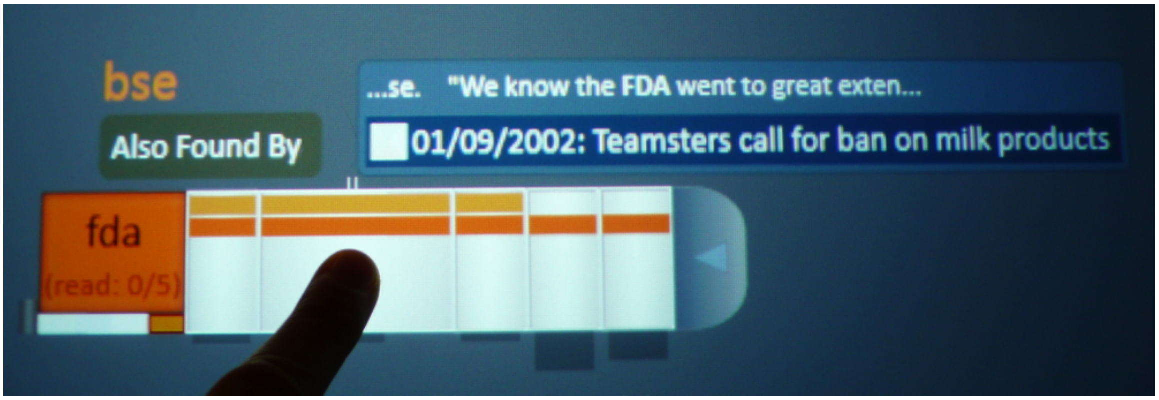


Figure 7.6: Detail-on-demand is shown for the document under the finger. It shows that "bse" also found this document (top-left), a document timestamp, title, and sentences that include the search term (white text, right).

looking at details for timestamp and document title, as well as the text surrounding the one occurrence of the word *fda* in the document text. Left of the title and sentence information, Ben can see a summary of search terms that have also found this document, just *bse* in this case. The colour is the same as the colour of the search representation, and stripes, representing the search term. The size of the word is representative of the number of times this specific search term occurs in the document, reminiscent of TagClouds (Viégas and Wattenberg, 2008).

When multiple team members synchronously issue searches in the system, the coloured stripes in the document representations are updated for all searches. The stripes are persistently linked across all views until a search is removed from the workspace. In our example, Ana has now begun her own searches. She is looking for political connections, trying to understand whether the mayor may have been involved in the BSE outbreak. She wants to see what the mayor and city hall have to say, and so she invokes a search for *luthor*, the name of the mayor, and *city hall*. Ana goes through her city hall list, finding a document that mentions BSE, Luthor, and city hall. The detail-in-context information now includes information for collaborative searches (Figure 7.7).

7.3.6 Has someone else issued my search?

As the analysts work, it may be important for them to know that they are currently looking at the same search or at a search a collaborator has previously looked at. Coloured

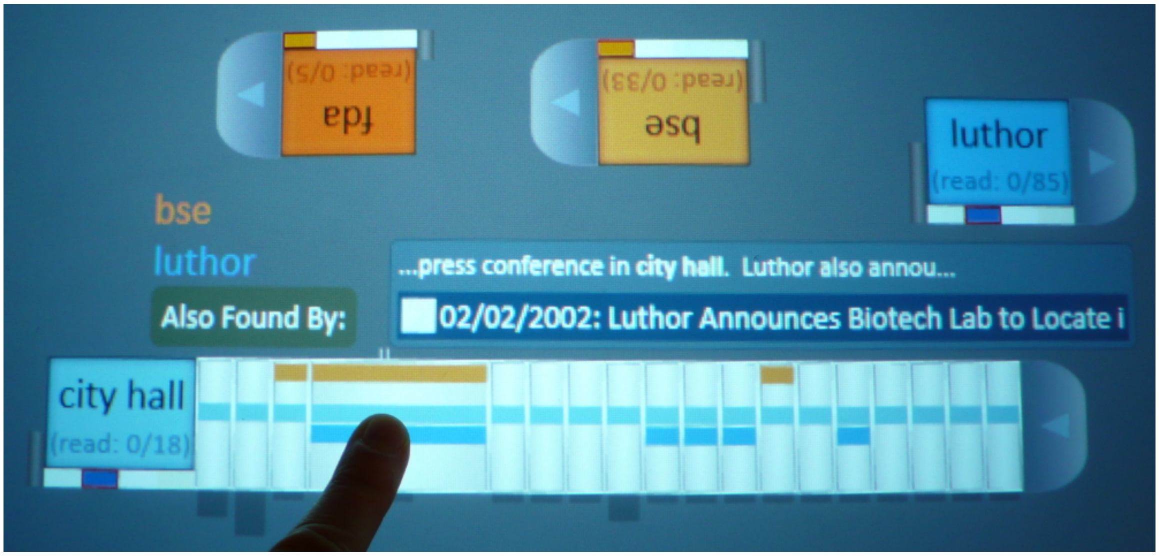


Figure 7.7: Different base-coloured stripes show when searches from other team members have found the same documents: Ana has search lists for "city hall" and "luthor".

rectangles under the search term show which analysts have also previously issued this search (Figure 7.8). A red border around the box indicates whether there is currently a representation of this search for this team member in the workspace. These four base colours used here are not included in the search term scales (Figure 7.5) in order to make them visually separate. In our example, both Ana and Ben follow their own hypotheses which finally lead them both to issue a search for *mad cow*. The orange and blue rectangles under the search term are both highlighted and receive a red border (Figure 7.8). The coloured stripe on the document itself is also split; this reinforces that the two analysts are both examining this word.

7.3.7 Has someone considered the same document?

Individual document representations can be dragged up and out of the search results, where they hover in the workspace (Figure 7.9). The original representation of the document in the search result list is highlighted by a red border to indicate that the document currently resides in the workspace. The basic representations of individual documents contain information on the document title and the publishing date. Stripes are again used to indicate a set of currently coloured keywords found in the document (Figure 7.10, left).

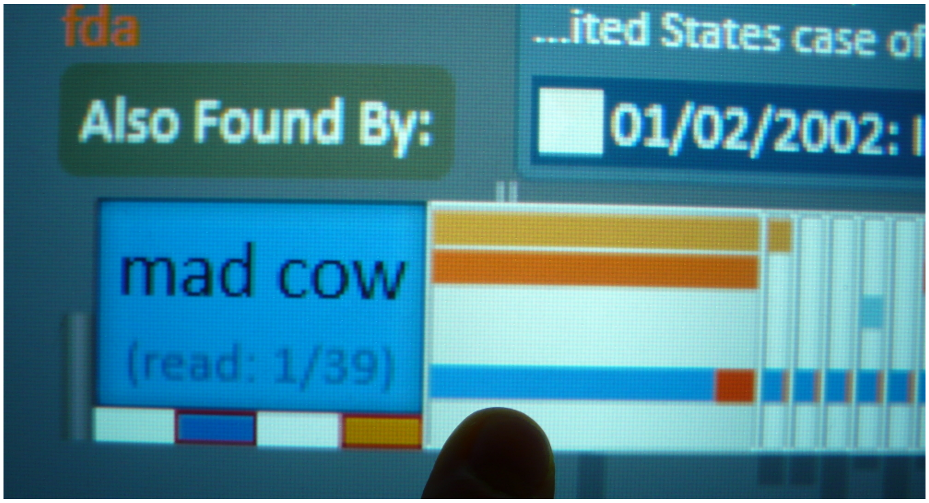


Figure 7.8: Ana and Ben have both searched for "mad cow." The search box has both blue and orange marks under it; just above the finger, the stripe that corresponds to the term is split and shows both their colours.

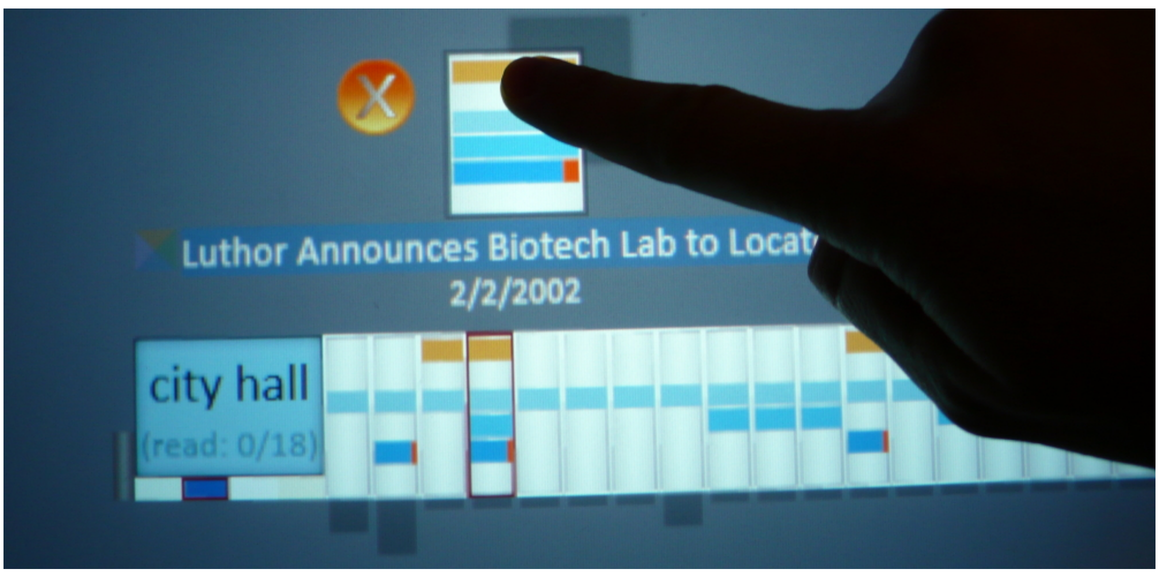


Figure 7.9: Ana drags a single result up and out of the search box, and so creates a floating representation of a document. Note that this representation shares the striping pattern of the search result.

In our scenario, Ana is interested in a specific document about an event at city hall, she pulls it out of the list (Figure 7.9) and can immediately see by the red border where the document resides in both the *city hall* and *mad cow* searches. When Ben sees the red indicator for a document in his BSE list, he can reconsider whether he had previously pulled the document in the workspace or start a conversation with Ana about the importance of the document.

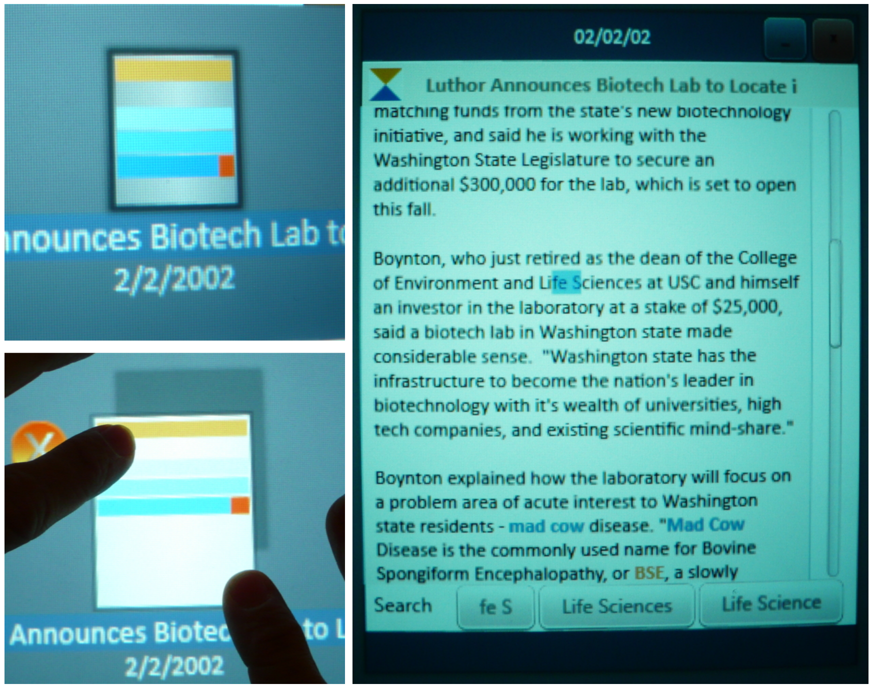


Figure 7.10: Minimized document representation (top left) and the full document reader (right). The reader is opened by resizing the minimized representation (bottom left).

7.3.8 Has someone read the same document?

The minimized representation of a document can be transformed into a full featured document reader by performing a two-handed resize gesture by pulling the document apart (Figure 7.10, bottom left). In the full document reader, search words are highlighted in their respective colour. If the search has been removed from the workspace, past search words are still bolded in black. This allows the analyst to see how previous searches have touched this document.

When a document is opened in the reader, its counterpart in the search representation receives a slightly darker background to indicate a read access, reminiscent of document read wear (Hill et al., 1992). The darker the background the more times a document has been opened. The colour becomes increasingly dark as the document is read more; however, the first colour step is distinctive enough that it is immediately visible whether a document has been read. Further reads are less finely grained to only reveal whether a document has been read or opened a few or several times as we expect the exact number to be less important. Figure 7.11 shows an example of two search representations in which individual documents have been read.

In our example, Ana has been concentrated on organizing her documents for a few minutes and was only peripherally aware that Ben had opened a document to read. As the document representations in her set change colour, she remains aware of which documents she or Ben have read.

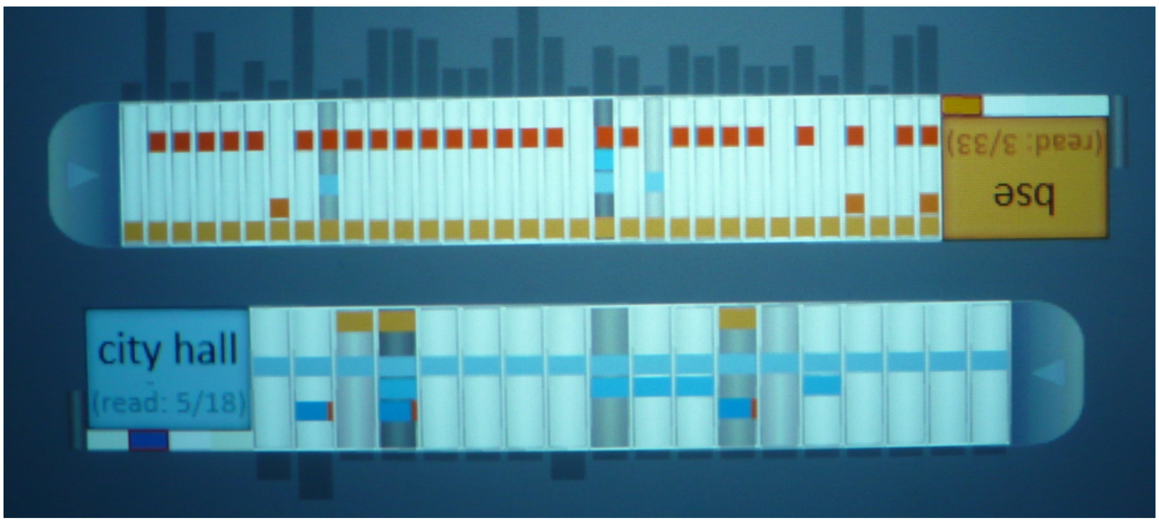
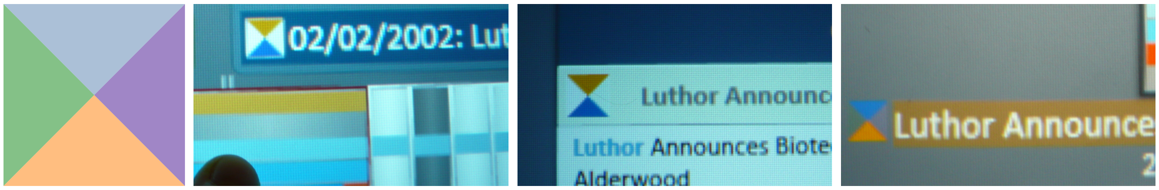


Figure 7.11: A darker background for individual documents indicates that a document has been opened in the document reader. A darker colour indicates repeated document access.

Information about what documents have been read and by whom is important for the synthesis of information from parallel analysis work (Robinson, 2008). In addition to the gray-scale encoding of read access in the document representations, we also designed a glyph to encode *who* had previously read a document. Figure 7.12 shows the basic glyph on the left. To its right are examples of the glyph that shows that both the orange and blue team member have read the document; their respective triangles are now opaque. The glyph is oriented to match the locations of the search buttons for each team member in the workspace; it is rotated appropriately to reflect the viewpoint of the owner of the object the glyph is embedded in. Figure 7.12b shows the glyph embedded in the detail-on-demand information of the search box, in Figure 7.12c it is



(a) Glyph. **(b)** In detail-on-demand. **(c)** In document reader. **(d)** In document title.

Figure 7.12: Icon representing who read a document. Each triangle stands for one analyst. The icon is embedded in three places. The three examples show documents that have been read by both the blue and orange analyst.

embedded in the title of the document reader, and in Figure 7.12d it is embedded next to the title of a document representation floating in the workspace. This last icon is also rotated towards the orange analyst.

7.3.9 Sharing Results

During their work, analysts may realize that one particular document or search may be more relevant for one of their collaborators. In this case, the documents and searches can easily be handed over by passing them to the respective colleague who can then drop it on his or her search button. This recolours the item and any interaction with it will now be issued from its new colour. In our case, Ben passes documents related to the politics at city hall to Ana while she does the same for Ben when she comes across something relevant to his BSE investigation.

7.3.10 Searches from Documents

We bring the interface full circle by providing functionality for analysts to issue new searches from the document reader, allowing them to pursue hypotheses through a chain of documents. On a finger's touch, the whole word under the finger is selected; dragging the finger selects a range of letters. In the bottom of the document reader (Figure 7.10, right) the analyst is then presented with several search options: both the exact selection and whole words contained in the selection. In the figure, the analyst has touched the word "Life" and moved his or her finger towards the word

"Sciences." The search suggestions for this selection include both the exact match "fe S", the contained words "Life Sciences," or the singular "Life Science."

7.4 System Summary

To elicit initial comments and observations we first demonstrated the tool to a security specialist at a large technology company, encouraging him to work with it for about 15 minutes. He found that it was easy to query the dataset, and felt that the direct interaction with the visual results allowed him to easily understand what he had found. He rapidly suggested a number of contexts in which the tool would be useful, ranging from time-critical search and analysis scenarios to large dataset exploration. He also suggested that the tool could be fruitfully adapted to distributed situations.

To get a sense about how our tool supports the visual analytics scenario discussed throughout the chapter, we conducted informal observations with two pairs of researchers using the aforementioned VAST contest dataset. Participants were seated each on the long sides of the table, were then given a 15 minute training session, worked on the task for 40 minutes, and then discussed their experience with us. Both groups had no difficulty reading or understanding the collaborative brushing and linking features. We saw these features used as intended: for broadening coverage of the information search, as a highlight for convergence, and as indicators to start an exploration. Interestingly, different members took advantage of the awareness features in different ways. One researcher was particularly concerned about common ground, and so made sure to read all documents that his partner had read. Another issued a broad set of queries, and then looked for overlapping terms between them. While the participants reported that they spent time monitoring each others' progress, they worked mostly in silence, only speaking occasionally to confirm a finding or discuss a document. Due to the given timeframe, participants did not solve the complete challenge but were able to find partial solutions. We used a relatively small surface table for this task which caused participants to lose overview of documents in the workspace that may be hidden under an opened document. For the size of table we used, we recommend no more than two concurrent participants for this task unless reading is performed on an external display. In terms of scalability, our system is currently most optimal for result sets containing up to 50 documents due to the physical display size. Each document in the dataset can be brushed with up to 24 concurrent searches due to the chosen colour scales that hold 6 colours per person. We have addressed this scalability issue by fading out old searches to gray—should more than 6 searches be added per person—and by adding interactive features that can re-colour an expired search. Using other visual variables than hue and saturation to distinguish searches and adding ways to refine and combine queries would be another way to address this issue.

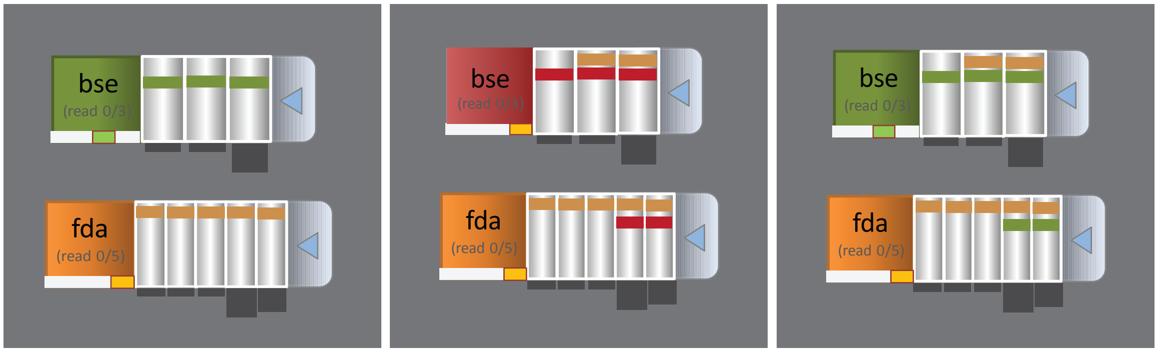
We expect that our collaborative brushing and linking features could easily be used on larger surfaces, in multi-display environment (MDEs), or with small modifications in distributed scenarios. As a next step we conducted a more formal study with the current tool to gain a richer picture of how pairs would interact with the tool in the context of a longer data analysis task.

7.5 Evaluation

In order to understand how small work teams would use Cambiera and its awareness features, how they would solve the tasks together, and how different coupling styles would manifest in this work, we[2] designed a study that would bring pairs of participants together to work on a visual analytics task using Cambiera. Motivated by the finding from Tang et al. (2006) of tool influence on coupling, we designed three experimental conditions that would vary the collaborative brushing and linking features of Cambiera (Figure 7.13):

None: Each team member's interactions with the data are only visible in his or her documents and search results. Each team member's searches and documents are tagged from a distinguishable colour palette. Overlaps between searches from both partners are not explicitly shown. We would expect to see partners in this condition work more loosely coupled as they are not specifically made aware of the overlap between their own and their partner's actions (Figure 7.13a).

[2] The following researchers were involved in designing and running this study: Petra Isenberg, Danyel Fisher, Meredith Ringel Morris, Kori Inkpen, and Mary Czerwinski. Thus, any use in the remainder of this chapter refers to this group of people.



(a) Condition "None": Each team member has a distinct colour. Search indicators do not overlap between both team members.

(b) Condition "Partial": Team members do not have a distinct colour. Search indicators overlap between both team members.

(c) Condition "Full": Team members have a distinct colour. Search indicators overlap between both team members.

Figure 7.13: Study conditions vary the types of awareness indicators that participants receive.

Partial: In this condition, team members share a single colour scale, so that they cannot explicitly tell which of them has issued a given query. Search overlap is shown but does not show who issued the search so team members have to rely on their memory to keep track of this information if it is necessary. In this condition, we would expect participants to receive awareness clues that would lead them to work more in a closely coupled fashion. However, we would expect participants to work carefully around each other's individual actions Figure 7.13b.

Full: The base state of Cambiera, as discussed above, each team member is distinguished by a base colour; team members can see each others' searches in their own search results and see which documents have been read and by whom. We would expect team members in this condition to work more closely coupled but also easily remain loosely coupled since one's own actions are easily distinguishable (Figure 7.13c).

7.5.1 Task

We based our experiment around the VAST 2006 Challenge, "Stegosaurus" (Grinstein et al., 2006), a visual analytics scenario that entails finding a hidden weapons-smuggling plot. Set in the fictional town of Alderwood, Washington, it requires teams to find and synthesize the results from searching through three hundred newspaper articles, plus several fact sheets and other articles, one map, four images, and one spread-sheet. The task begins with an introduction that gives a starting clue and suggests a first document to read. Buried among the three hundred articles are ten critical documents that are directly relevant to uncovering the plot. The plot of Stegosaurus hinges on identifying eleven connections between the people, places, and institutions in the story: the team must discover information about a terrorist group in South America, a dangerous chemical in Washington, and a smuggling ring that is shipping the chemical from Washington to South America (see Appendix A.4 for details). Cambiera does not currently support search of non-textual materials like maps and images. As such, we needed to pick a subset of the documents to be presented on the tabletop. We preloaded Cambiera with all of the newspaper and fact sheet articles. We also provided paper print-outs of the map and other images. Participants, thus, interacted with an archive of approximately 320 digital documents.

7.5.2 Participants

It is extremely difficult to obtain professional intelligence analysts for this sort of research; as with other visual analytics studies (Stasko et al., 2008), we instead recruited people who were familiar with data analysis. Participants were required to have a Master's (or more advanced) degree, and to have self-reported as enjoying solving puzzles or solving mysteries. The members of each pair knew each other and had worked together in some form; subjects included co-workers, friends, family members, and married couples. We recruited fifteen pairs of participants in our study. Participants ranged in age from 25–55; ten of the couples were mixed-gender; three were both women; and two were both male (14 male and 16 female participants in total).

7.5.3 Experimental Procedure

Each pair was assigned to one of the three conditions (*None*, *Partial*, *Full*), for a total of five pairs in each condition. Participants received a 15 minute tutorial on the various features of Cambiera using a sample dataset, during which they were encouraged to experiment with the features and ask questions freely. They were then introduced to the "Stegosaurus" problem with an introductory letter explaining the context. The external experimenter running the study was familiar with the dataset, and so was able to monitor the teams' progress. When teams stopped making progress entirely, as judged by the experimenter—visible by reading and re-reading distracter documents, or by reporting to be stuck—the experimenter provided assistance. During an assist, the experimenter did not provide new information, but rather asked the participants to clarify previous ideas that they had raised. Since our focus was on observing the group's collaborative interactions with each other and with the system, and less on performance outcomes, we wanted to ensure that pairs were able to make progress in the task, and that they continued working. For this reason, we decided to provide assists to teams who did not progress in the task. We tried to maintain consistency of assists by using only one external experimenter, who followed a written protocol. Participants reported their results verbally at the end of the study. We terminated the experiment when the team could produce a reasonably coherent story when asked for their hypotheses, and ended all experiments at one and a half hours. After the study, each members of 12 of the pairs filled out a questionnaire, resulting in 24 questionnaires in total (due to a technical error, three pairs did not fill out the questionnaire). Next, the experimenter debriefed the pair to understand how they approached the problem and to get their feedback on the technology.

7.5.4 Data Analysis

Sessions were video- and audio-recorded; in addition, screenshots at one-minute intervals and event logs were captured with timestamps for interactions with the tabletop. One experimenter took notes in real time on the group work. Another experimenter (the author of this dissertation) performed several video coding passes in order to get a rich understanding of the ways in which groups solved the task. To get an understanding for the types of details to code, I first engaged in thorough video coding pass of one session using an extended video coding tool I implemented (extended from the version mentioned in Chapter 4 and 6). During the coding I took full time-stamped transcripts of participants verbal communication, noted observations on work style (e. g., switches between parallel and joint work), task solving styles (e. g., what type of information was looked at and for what reason as well as current level of understanding of the problem), use of external material, observable occurrences of awareness feature use, hypotheses voiced, workspace organization styles, and tool problems. I coded this video with time-stamped interactions based on several codes extended from the set presented in Chapter 4. The full list of these initial codes can be found in Appendix A.4.2. Since coding on this level of granularity took about six hours for 90mins of video data, we discussed the most interesting findings from this coding and cut down the coding categories to a more manageable set. I then coded the collaboration styles behavior based on the code set presented by Tang et al. (2006) and took notes on roles adopted, how collaborative brushing and linking features were used (for spread or replication), how external information was used (shared or private notes), what types of group work breakdowns or conflicts occurred, and how often workspace items were shared.

During the first coding pass it became evident that the code set from Tang et al. (2006) had to be extended to accommodate the different study situation. Whereas the original code set was developed for a situation in which participants shared the same representation, our participants could also work with individual information items. Therefore, we extended the code set to more clearly distinguish when people shared views of the data and when they shared the same information items. More details about the extended code set are included in the results section of this chapter. During the second coding pass, I took detailed time-stamped notes on when participants switched to different types of collaboration styles using the extended code set with the video coding tool I implemented. I also took extended notes on which facts and documents teams found. The code set for the facts and documents is not included here as the dataset we used is part of a training dataset for intelligence analysts and we had agreed not to publish the answers.

I also engaged in a coding of the post-session interview based on written transcripts of participants' answers. The transcripts that were the basis of this coding were made by one of the two experimenters running the study. The coding of these transcripts resulted in higher-level categories of participants answers including: awareness (e. g., which information was missing or seen as helpful), work styles (e. g., strategies, roles, sharing, and collaboration) as well as tool features commented on (e. g., liked and missing features). The interaction log data was parsed and statistically analyzed by two other experimenters.

Together the detailed analysis of the field notes, log files, and screenshots provided a rich understanding about how teams solved the given task. The most interesting findings in relation to this dissertation work are presented next.

7.5.5 Findings

The results of our exploratory study showed that the tool and the tabletop collaborative setting allowed participants to approach the problem quickly and effectively. All participants immediately immersed themselves in the task and made use of the various features Cambiera offered. In this section, we present more detailed findings on how participants solved the task, worked with Cambiera, and how participants engaged with each other.

Task Completion

To find necessary information, groups issued an of average 50 ($\sigma = 14$) searches with 42 ($\sigma = 11$) distinct search terms. These led the groups to open an average of 90 ($\sigma = 26$) text documents in the document reader per session. Of these, the group opened an average of 58 of them more than once. Eleven of the 15 groups found all ten critical documents; the remainder missed one or two. Despite finding most of the critical documents, we observed a wide variance in how well pairs were able to connect the facts they had found, ranging from making three connections to all eleven ($\mu = 7.73, \sigma = 2.58$). Teams also required varying degrees of assistance ranging from one group that only connected four facts, but had five assists, to a group that made all eleven connections, with no assists. The average completion time for all experiments was 72 minutes ($\sigma = 12$).

Influence of the Experimental Condition

In reviewing the fifteen pairs, we found that the conditions did not impact how the tool was used overall. We did not find statistically significant differences for the number of searches performed, number of documents read, time spent on the task, or numbers of documents or searches passed between participants. This was surprising because we had hypothesized that the different conditions would influence how closely participants worked and that this would impact how the tool was used. For example in the *None* condition, participants did not see others' searches in their own result lists. We, therefore, expected them to pass documents back and forth more than in the other conditions. In other conditions, we expected participants to find overlaps more easily and see which documents others had read, leading to fewer documents passed to the partner. In Figure 7.14, we illustrate the mean number of instances of documents and searches shared by teams over the three conditions. A one-way Analysis of Variance (ANOVA) did not reveal a significant difference between the number of document exchanges or search exchanges, although the data trended in that direction ($F(2, 12) = 2.7, p = 0.11$ and $F(2, 12)2.5, p = 0.12$, respectively).

Measuring success is difficult in complex tasks like the one we tested. We can measure how many critical documents participants found and how many assists they received over the time of the trial. However, creating a joint metric out of these factors is not necessarily meaningful, as different types of assists were given to groups. On average our experimenter gave two assists per group. Three teams required no assists, found all critical documents, and connected 10 or 11 facts. These three teams were each in a different experimental condition. We believe that the complexity of the task and study situation, as well as individual and group variability, were contributing factors for the lack of measurable difference in experimental condition.

Despite this lack of measurable influence on tool use, eight of the ten pairs in the *Partial* and *None* condition spontaneously commented that they would have preferred additional awareness features during our post-session interview. During the debriefing sessions, eight of the ten pairs in *Full* and *Partial* identified this awareness information as a substantial benefit of Cambiera. Three teams in *Partial* and all five teams in *None*—that is, pairs that had incomplete or no awareness information—spontaneously suggested that it would be desirable to "add a feature" to allow the team to see each other's searches, and to code them to show who had done the search. In other words, these pairs expressed a desire for the features provided in condition *Full*. We note that participants found ways to work around the lack of awareness information, for example, by taking on different roles (e. g., one reading, one searching) and by sharing information verbally. Participants' qualitative feedback on the usefulness of the features

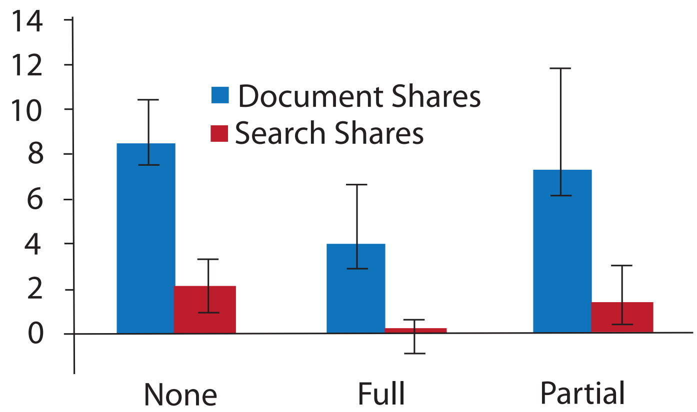


Figure 7.14: We did not find statistically significant differences between the average number of times that documents and search results were explicitly shared between pairs in the three different experimental conditions.

and their desire to have these features makes awareness support a promising direction for further investigation.

Observations of Group Work Styles and Strategies

To better understand whether, how, and when participants switched between phases of parallel (or loosely coupled) and joint (or closely coupled) work, we coded the coupling styles for each group inspired by the code set by Tang et al. (2006). For each code, we observed which of the analysis processes from Chapter 4 were common in each coupling style in order to further analyze how different information analysis processes manifest in phases of joint and parallel work.

We extended this original code set from Tang et al. (2006) by four codes. In the formulation, distinct codes specify whether the two team members were working on the same problem or different problems, on the same area or opposite areas of the table, and whether one team member was passively viewing. In addition, the original set coded every interaction as loosely or closely coupled. The coding scheme from Tang et al. (2006) with our modifications, can be found in Table 7.1. Our modifications are annotated with a (*).

Coupling	Code	Description	Most Common Analysis Processes
C	DISC (*)	**Discussion:** Conversation about the tool, task status, or work strategies (Figure 7.15).	Establish Task Strategy, Discuss Collaboration Style, Validation
C	SPSA	**Same Problem, Same Area:** Both participants work on the same information artifact on the table (Figure 7.16).	Browse, Parse, Select, Operate, Validate
C	VE	**View Engaged:** One participant is actively working with an artifact; the other is actively viewing, possibly commenting, but not touching (Figure 7.17).	Parse, Operate, Validate
C	SPDA-SI (*)	**Same Problem Different Area—Same Information:** Both participants are reading the same document, on their own sides of the table.	Browse, Select, Parse, Operate, Validate
C	SPDA-SSP (*)	**Same Problem Different Area—Same Specific Problem:** Both participants are working from a shared set of documents to solve the same problem (e. g., search through all found police reports for a clue about the injured driver).	Browse, Select, Parse, Operate, Validate
L	SPDA-SGP (*)	**Same Problem Different Area—Same General Problem:** Both participants agreed to find clues to a common question (e. g., "what happened to the driver") but are taking different approaches and searching for different terms. Private note taking is common.	Browse, Select, Parse, Operate
L	DP	**Different Problem:** The two team members work on separate areas of enquiry often taking notes on their findings.	Parse, Browse, Select, Operate
L	V/D	**Viewing / Disengaged:** One person is working on the problem; the other is watching passively or disengaged.	Parse, Browse, Select, Operate (one person)

Table 7.1: Video codes, adapted from (Tang et al., 2006). (*) indicates a category that is not in (Tang et al., 2006). In the first column, C indicates that we will refer to this as a close coupling and L as loose coupling.

Discussion (DISC):

Our first new code, *DISC* (discussion) focuses on participants' conversations about the task or tool. Discussion occurred always in closely coupled collaboration. We observed four main types of discussions in our study:

1. Discussions about how found facts can be connected and formed into a coherent story. This type of discussion included analysis processes *Discuss Collaboration Style*, *Validation*, and *Establish Task Strategy* from Chapter 4 which in this previ-

Figure 7.15: Example of participants discussing the tool (left) and discussing their current strategy (right).

ous study were also mostly performed in closely coupled collaboration. Discussions of collaboration styles were typically very brief and mostly evolved fluidly throughout the task, similar to what we observed in Chapter 4.

2. Discussions about which strategies could or should be applied to find new or missing facts or prove current hypotheses. In this type of discussion, team members further *established task strategies*, albeit with a focus on tool usage (i. e., how should the tool be used to find necessary information).

3. Discussions about how the tool should be used or interpreted. This type of discussion involved the *Clarify* process from Chapter 4. We mostly observed participants engage in clarifications about the encoding of awareness features.

Figure 7.15 gives an example of two of the three types of discussions we observed in our study.

Same Problem Same Area (SPSA):

When working on the same problem in the same area of the workspace (SPSA), the two team members shared the same information item (document or search result list) and would read a document (*Operate* code from Chapter 4), use the *Browse* process to look through the same search result list, or *select* a set of documents from a shared search result list. The reading of a common document was often accompanied by *validation* of a question, hypothesis, or interpretation that one team member had previously raised about the content of the document. Figure 7.16 gives an example of two pairs reading the same document together.

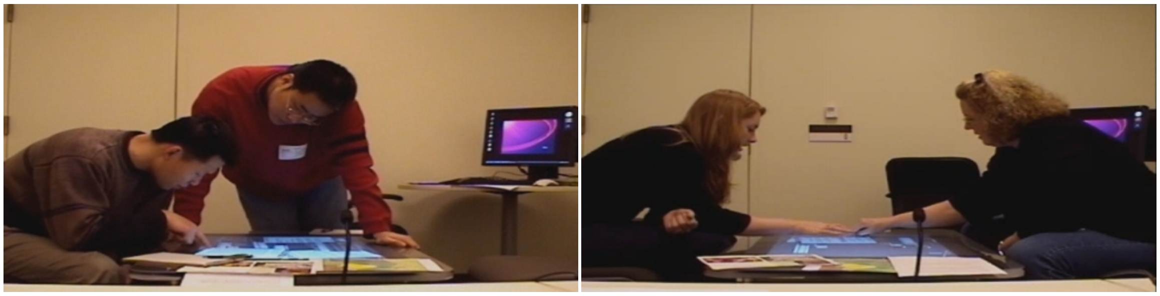


Figure 7.16: Example of participants working with the same information in same areas of the workspace.

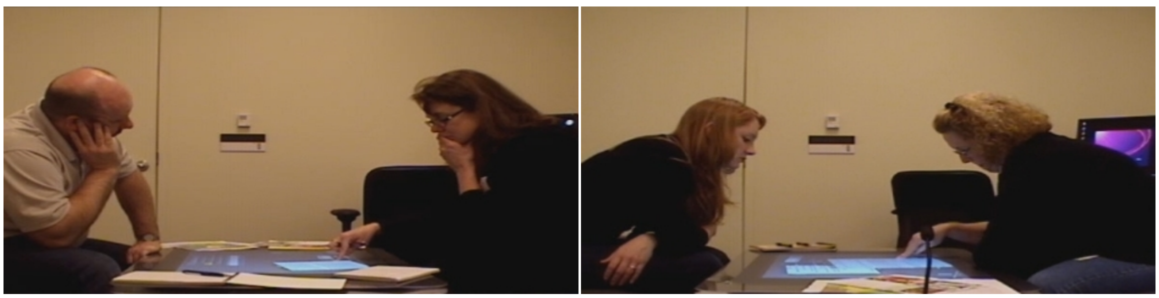


Figure 7.17: Example of two pairs where one team member is actively listening to the other team member without interacting with the workspace.

View Engaged (VE):

During phases in which pairs were *View Engaged*, one participant would interact with a workspace artifact, while the other was actively watching or listening. The main type of view engaged behaviour involved one participant reading a document aloud (the *Operate* process from Chapter 4) while the other was actively listening. This was a commonly observed behaviour when *parsing* the task sheet. Often, the other person was listening to a document being read in order to *operate* on (e. g., understand, actively connect, interpret) or *validate* some facts that the person reading the document had found. Figure 7.17 gives an example of two pairs in a VE phase.

Same Problem Different Area (SPDA):

Our next three codes split the original *SPDA (same problem different area)* code. Tang et al. (2006) describe the SPDA code as "two team members working on the same problem on opposite sides of the table." In contrast, our participants did not coordinate their interactions over a full screen shared, spatially fixed, representation but were interacting with representations which could be shared, duplicated, and moved. We found that we needed more degrees of refinement in our scenario, since the original SPDA code could not precisely cover the more diverse range of information that participants worked with in this study. However, differentiating what kind of information (e. g., same documents, same topics, different topics) can help to inform how closely or loosely coupled participants worked. Hence, we split the SPDA code in three sub-codes describing interactions with different types of information on the table surface:

SPDA-SI Same Information: Two team members reading the same document at the same time but in different areas of the workspace using their own document readers. Participants were engaged in the *Operate* process while reading the document to retrieve new facts or in *Validate* when checking previously found information. Participants were also coded as being in SPDA-SI when they *browsed* through the same search result set and *selected* information from the same result list representation.

SPDA-SSP Same Specific Problem: Both team members are working to solve the same specific problem. They have previously decided what problem to solve (e. g., to find out what happened to the driver involved in a car crash). Activities involve searching to find a specific document (e. g., a document about a car crash at a bank) or reading different documents from a common search result list looking to find a specific piece of information or fact (e. g., reading obituaries to find out whether a specific person had died). In SPDA-SSP, participants typically engaged in the *Operate* process but also *browsed* search results, *selected* new information to read, *parsed* the initial task sheet, or *validated* a partner's results.

SPDA-SGP Same General Problem: Both team members are working on the same general problem but taking different approaches. For example, both partners have agreed on finding out whether the Boynton lab is involved in the mystery; one searches for FDA investigations and the other searches for its connections to a BSE outbreak. Similar to SPDA-SGP, participants were mostly engaged in *Operate*, *Browse*, *Select*, or *Parse* processes.

We coded SPDA-SI and SPDA-SSP as closely coupled (C in Table 7.1) and SPDA-SGP as loosely coupled (L in Table 7.1). Visually, all three codes looked similar to Figure 7.18—

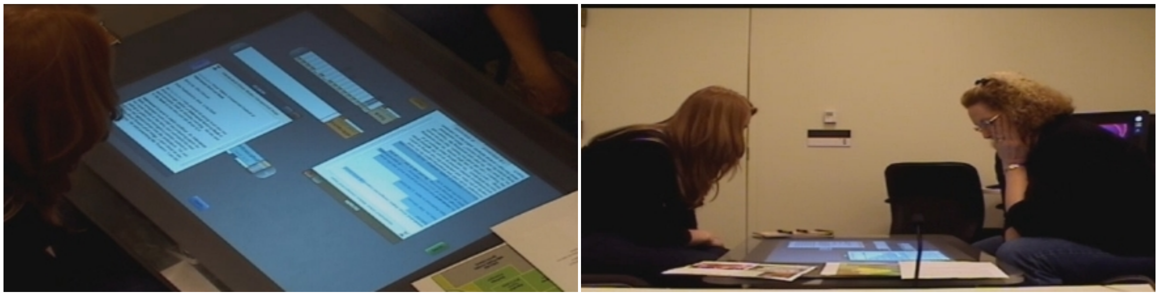


Figure 7.18: Example of two participants working on the same problem in different areas of the workspace. Here, both team members are reading the same document (SPDA-SI), trying to understand the information contained in relation to their previous findings.

Figure 7.19: Examples of two participants on different problems (left) and one participants viewing the other person without actively engaging in the task (right).

an example of the SPDA-SI code. As both participants were working in different areas of the workspace, the context of what participants were working on and what information they were looking at was deduced from video and audio coding and from looking at the time-stamped log files of the pair's interactions with the workspace.

Different Problem (DP), Viewing (V):

When partners did not agree on a common goal, their interactions were coded as *DP (different problem)*, for example, when one decided to search for all the farms on the map to look for clues and the other looked for chemicals with flower smells. When one partner was not actively listening but watching the partner work, we coded their interactions as *V* (viewing). DP and V were coded as "loosely coupled" collaboration. Figure 7.19 gives an example of both types of work. In *DP* and *V*, the working team member was engaged in any of the processes: *Operate*, *Browse*, *Select*, or *Parse*.

Temporal Sequence of Work Phases:

In order to get a better sense of when participants were working more closely or loosely coupled, we generated sequence graphs of the codes described above. Table 7.2 shows which work phases—according to the codes above—our participants engaged in, while Table 7.3 shows the temporal sequence of phases in which participants worked closely or loosely coupled.

In Table 7.3, we see the wide variance of times in which groups were closely coupled. Groups spent anywhere from 32% of their time (Group 7) to 92% of their time (Group 1) in close collaboration. The experimental condition did not have an influence on which coupling styles groups adopted. Overall, we observed a tendency for groups to work closely coupled, with eleven (of 15) groups spending over half of their time in closely coupled collaboration. We refer to those eleven groups as closely coupled; the remaining four groups are loosely coupled. Loosely coupled teams spent a large amount of their time (43%, on average) trying to answer a common general question such as "what is the involvement of Boynton laboratories," working on their own part of the workspace with separate search results and documents (SPDA-SGP). Closely coupled groups spent longest working with the same result sets in the SPDA-SSP style (24%, on average), but also spent an average of 23% of their time in SPDA-SGP. Frequently, this separate work led to groups switching to one of the closely coupled styles to discuss intermediate results, read documents together, interpret found facts, or offer/ask for help. In addition, our coding revealed that our teams showed high task engagement, with very little time spent in V (viewing) or D (disengaged) styles.

Closely and loosely coupled groups read documents differently. In closely coupled groups, team members would often read (*operate on*) documents aloud (VE), would rotate an open document towards the other person to allow them to read it together, or would read a document alone and then summarize the findings for the other person. In loosely coupled pairs, reading was mostly done in parallel, and when something interesting was found, private notes were taken; Occasionally, the document was passed to the other person to read for themselves. In these groups, sharing of information was less often spontaneous, and more often initiated by prompting: one group member would ask the other, "so what have you found?"

Group	Condition	% of Time Closely Coupled	Coupling Styles
7	None	32%	
5	Partial	33%	
8	Partial	39%	
3	Full	46%	
4	None	50%	
6	Full	54%	
9	Full	56%	
11	Partial	57%	
14	Partial	60%	
13	None	68%	
1	None	69%	
12	Full	75%	
2	Partial	90%	
15	Full	91%	
10	None	92%	

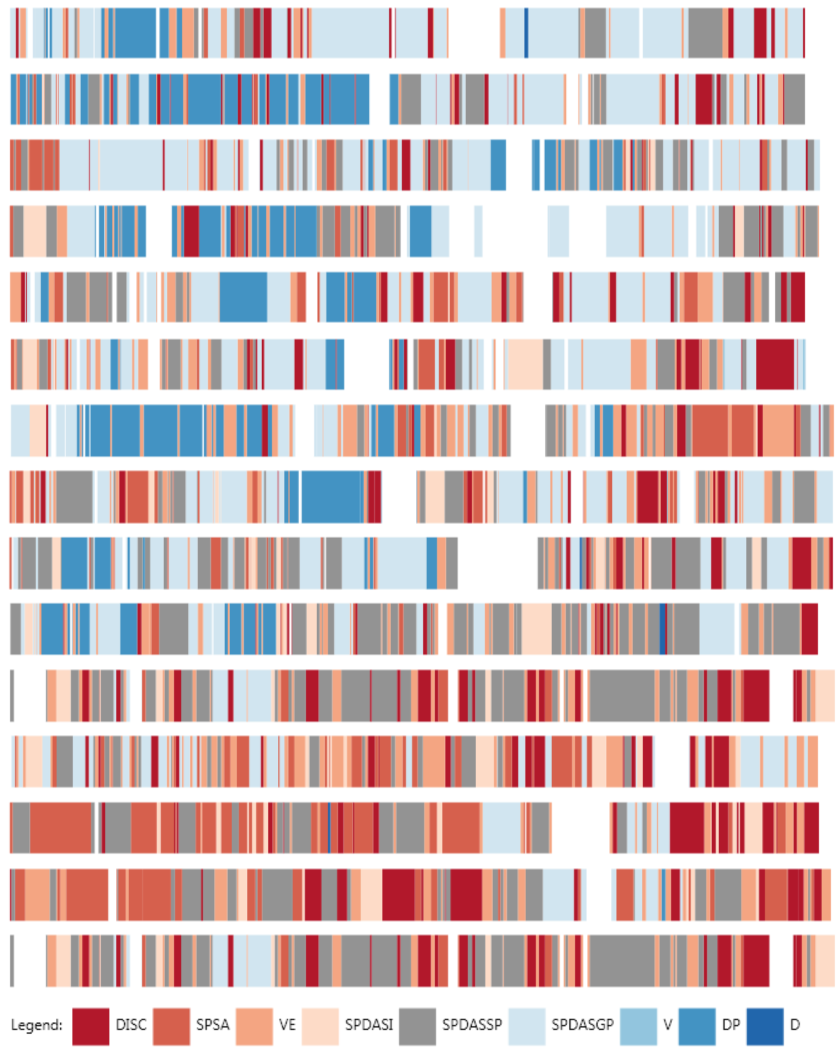


Table 7.2: Sequence diagrams of closely and loosely coupled phases that pairs engaged in. Red encodes phases of mostly closely coupled collaboration, while blue encodes phases of loosely coupled collaboration. Gray indicates the SPDA-SSP code which was coded as closely coupled. White indicates phases in which groups had stopped working (e. g., for interaction with the experimenter).

Group Work Strategies and Awareness Features

Watching the groups work, we observed that our participants found ways to work around the features that were missing in their experimental condition. For example,

Group	Condition	% of Time Closely Coupled	Coupling Styles
7	None	32%	
5	Partial	33%	
8	Partial	39%	
3	Full	46%	
4	None	50%	
6	Full	54%	
9	Full	56%	
11	Partial	57%	
14	Partial	60%	
13	None	68%	
1	None	69%	
12	Full	75%	
2	Partial	90%	
15	Full	91%	
10	None	92%	

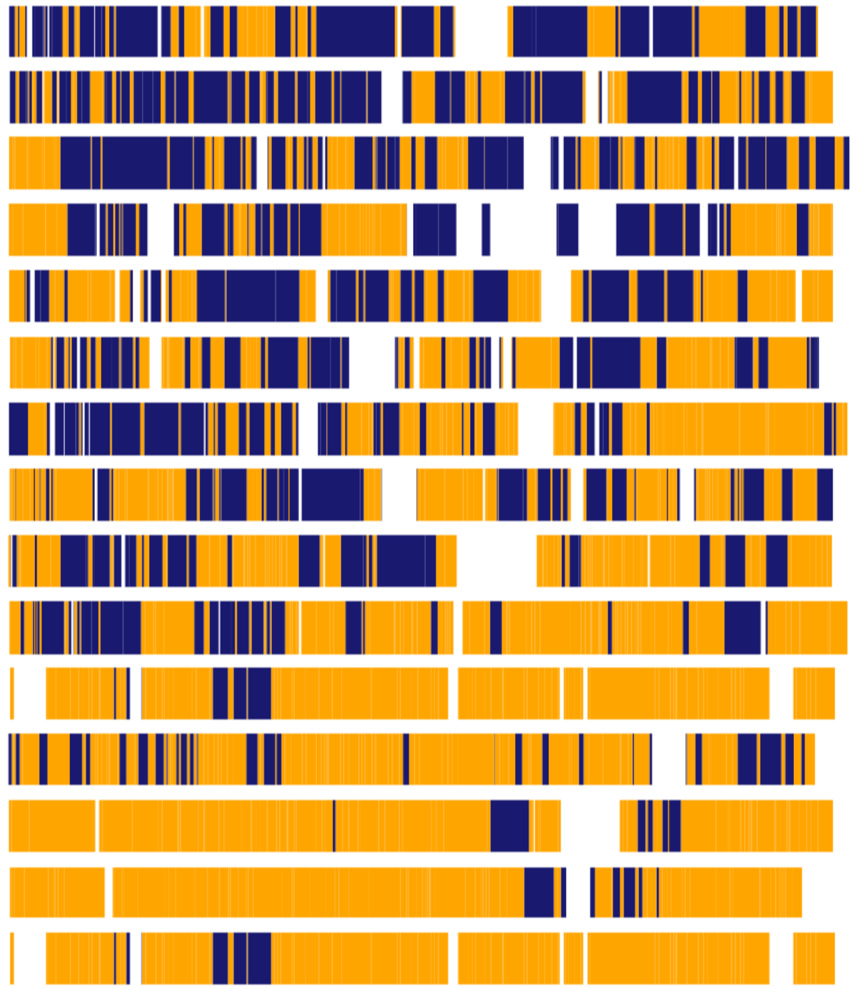


Table 7.3: Sequence diagrams of coupling styles that each pair engaged in. Blue encodes phases of loosely coupled collaboration and yellow phases of closely coupled collaboration. White indicates phases in which groups had stopped working (e. g., for interaction with the experimenter).

we saw teams passing searches back and forth frequently in the *None* condition, which allowed them to see search overlap more easily. Some groups in the *Partial* and *Full* conditions, however, worked in a way that made the search overlap features less effective. For example, some groups cleaned up their search results right after they were used, which circumvented the usefulness of the stripes. Other groups searched along completely different paths which led to infrequent search stripe overlaps. In order to illustrate how different strategies impacted the ways that teams worked, and how they utilized Cambiera, we discuss three different teams' strategies in detail.

Group 2: Working Closely Together

Group 2 was in the *Partial* condition, meaning that they did not have distinct individual colours. The two participants in this group were close friends and co-workers. They found all ten critical documents, made all eleven connections with no assists, and solved the complete task correctly in 70 minutes. Both participants worked closely coupled for 94% of the time. They had a clear work strategy: they searched and browsed the results in parallel but read interesting articles together, sharing 13 documents. They were able to rapidly identify connections between facts, and moved through the study very efficiently. Since these participants worked very closely with frequent communication throughout, they generally had a very good awareness of what each other were searching for. Each made sure that their partner read important documents that they had found, so Cambiera's awareness features were less critical for this group. As a result, the pair used query colouring less to track their own searches, and more as a way of finding documents that looked to be information-rich. In particular, they preferentially read documents that had several stripes (found by several searches), and that had not been read before.

Group 5: Failing to Combine Knowledge

Group 5 was also in the *Partial* condition. Both participants were friends, co-workers, and experienced puzzle solvers. They had coordinated multi-hundred-person puzzle-solving competitions, and so felt very confident about their ability to solve the mystery. However, within the 69 minutes of their work they found all critical documents but only found five connections and required 3 assists. Both participants reported that they were accustomed to working separately, trying to figure out a puzzle on their own. As a result, they adopted a loosely coupled work style, working closely coupled only 33% of the time. After reading the initial document together, they each chose a part of the problem that was of interest to them. As they worked, they would look for documents that might help their current approach, and periodically mentioned their thoughts aloud. They spent substantial portions of the task unaware of each others' work. It was not until the experimenter asked for a status update that they began to realize what information they were missing. Table 7.2, middle, illustrates the fact that Group 5 spent much of the study working separately on different problems (DP), periodically checking in. During the task, the group issued 78 searches; they later reported that they had gotten lost in the sheer quantity of results they had created. They did not make visible use of the awareness information that they did have, since they spent most of the time working on different problems where the search and read overlap was minimal. After the experiment, they asked for better awareness information that would let them know who had searched for a term and who had read a document, in order to be better aware of the difference between their own and their partner's work.

Group 10: Role Taking

Group 10 was a pair of History graduate students from a nearby university, who had enjoyed working together on a number of projects in the past. Daphne and Charles were in the *None* condition, which meant that search results did not contain information about which of the partner's searches found the same documents. The group made ten of the eleven connections, required no assists, and solved the complete story in 66 minutes. Daphne and Charles chose a working style that allowed them to work closely together. They overcame the constraints of the *None* condition by separating into roles. Daphne would select articles that she found interesting, and pass them to Charles. Charles would read them in more detail, sometimes opening them to fill half the screen so that they could read them together. As the study progressed, Daphne continued to conduct searches; Charles continued to sort and organize results. Charles would suggest ideas, which Daphne would search for. Their choice of distinct roles meant that the lack of links in their colours was irrelevant; they had effectively the same view as a group in the *Partial* condition.

These very different strategies meant that we were unable to detect a significant relationship between experimental condition and success. Teams found ways to work around the lack of links, and some teams that were linked disregarded that information.

7.5.6 Discussion

A key finding in our experiment was that while groups were presented with different variants of Cambiera in the three experimental conditions, this did not have a significant influence on performance. Instead, we found that there was substantial individual variation between teams: different teams approached the problem very differently. These differences were far more linked to their work styles than the experimental manipulations we chose. Watching the groups work during the experiment, we observed that our participants found ways to work around the awareness features that were missing in their condition. In this way, we share some of the surprise of Tang et al. (2006), who found that team members would work closely together even in conditions that were meant to separate them. However, our qualitative data showed that 8 of the 10 teams in the *None* and *Partial* conditions requested more awareness features. This could indicate that having to work around the lack of features was perceived as extra effort for the participants which could be alleviated by integrating better awareness indicators into the result visualization.

Cambiera was designed to support a range of coupling styles. Our results also show that a broad selection of coupling styles were adopted by our participants and that participants frequently and fluidly switched between these styles. More groups worked closely coupled than not; the ease of sharing information across the table could account for this dominance of closely coupled work. In relationship to Chapter 4, we noticed that the earlier analysis processes could also be identified in this study. We further found that similar to the earlier study, participants tended engage in certain analysis processes (browse, parse, operate, select) more commonly in coupling styles which were coded as loosely coupled, while other processes (discuss collaboration style, establish task strategy, clarify, and validate) were more common in closely coupled collaboration.

7.6 Chapter Summary

In this chapter I introduced Cambiera, a system designed to specifically to explore the notion of awareness indicators in co-located collaborative analysis activities. The problem addressed with this system was derived from findings of studies in Chapter 4 and

Chapter 6 which showed that team members would divide up their work with information visualizations and specifically liked to select and operate on specific data items individually. Specifically, during the study on CoCoNutTrix in Chapter 6, participants asked for better awareness information of what had been worked on and by whom and about which data items individual team members had already formed a decision. With Cambiera, we integrated specific awareness support into the visual display of search result lists. We defined collaborative brushing and linking as an awareness mechanism which has the goal to inform team members of each others' work during phases of parallel work. We note that there are several important differences from interactive brushing and linking. Unlike interactive brushing and linking, the collaborative form links views through social information about items being worked on. Also, interactive brushing and linking can be ephemeral: a team member can select a node on one chart, and see the highlight propagate to another instantly. In contrast, in an awareness tool, collaborators may not notice changes instantly; changes may have to be persistent or propagate slowly. Finally, interactive brushing and linking usually connects distinct visualizations with the same underlying data whereas collaborative brushing and linking adds on additional meta information. To help maintain common ground, Cambiera uses the same type of representation throughout. While it may be possible to use different representations, we have not yet experimented with that.

In our design, we have emphasized persistent colourings in order for team members to share common ground. In previous sections, we have outlined four different forms of collaborative brushing and linking:

- *search stripes*, to help team members see other search terms,
- *document read wear* to show what documents have been read and by whom,
- *red highlights* around documents to cue that the document is visible in the workspace, and
- *search boxes* which show who has repeated the same search.

Cambiera is designed for a synchronous, co-located surface computing system. While the concepts of shared awareness information are well known for distributed settings, co-located tasks may benefit from unobtrusive awareness cues. In a study of 15 pairs using Cambiera, we modified its collaborative brushing and linking features to try to understand the influence of these awareness features on group work and the use of the tool. We found that the awareness features did not have a measurable impact on group work behaviour or success in the task. Instead, we found that the results were swamped by the wide variety of coupling styles people adopted. Interestingly, however, pairs did ask for the brushing and linking features when they did not have access to them. In general, with design refinements, the concept of collaborative brushing and linking appears to have been partially validated in this research and in particular the qualitative responses of participants and our observations of people working around the lack of features in conditions that did not provide them. However, several questions remain open. First of all, the awareness features of Cambiera may have been more helpful in certain coupling styles. For example, groups who were more closely coupled could have been paying closer attention to each others' work and may have noticed the subtle awareness clues more easily. This hypothesis may be very difficult to test in practice as it is hard to force a specific coupling style on a group in a complex and long task such as ours. Secondly, we found that pairs who were more closely coupled were more successful at making connections within the data. We look forward to other studies that can help untangle the complex dependencies between coupling, awareness features, and success in an analysis. Figure 7.20 gives on final overview of the tool.

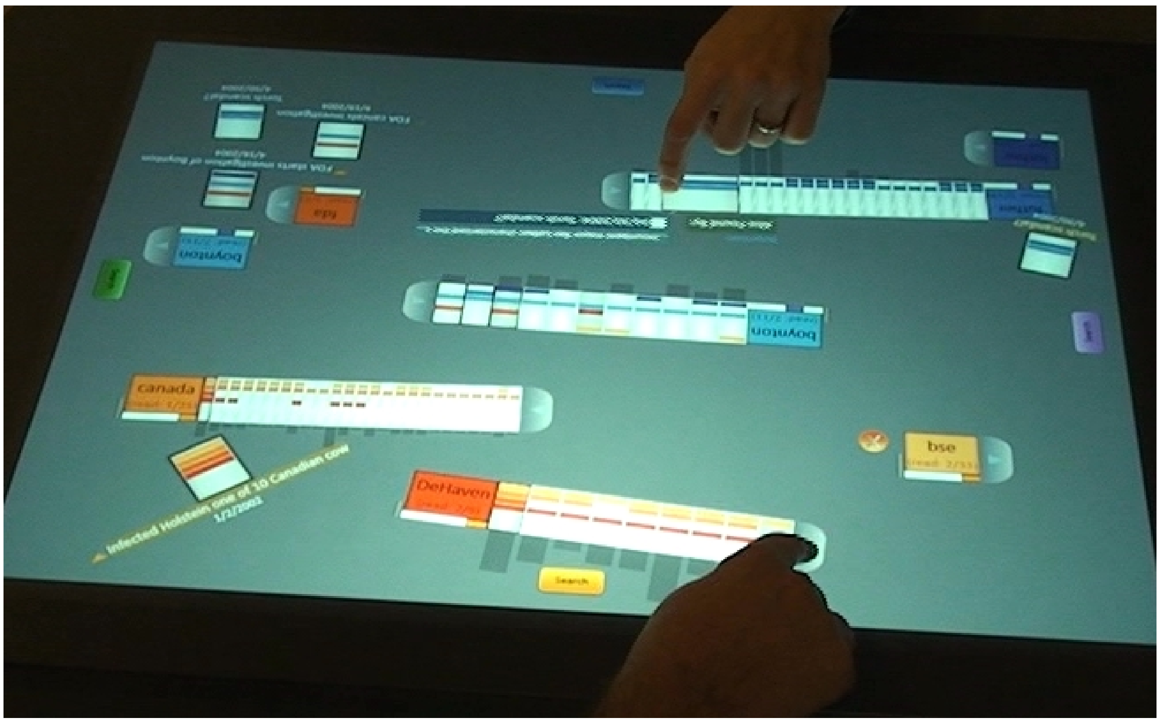


Figure 7.20: Overview of the workspace during an analysis session. Both analysts have arranged several search boxes and documents in the space related to their current hypotheses.

Chapter 8

Conclusions

In this dissertation, I have explored issues and challenges in the design of information visualization systems for collaborative co-located and synchronous work on shared large displays. In this chapter, I take a final look at the issues raised in the case studies and literature analysis performed in the earlier chapters. First, I summarize my research challenges and contributions towards the addressed research challenges. I then extend the first set of design considerations derived in Chapter 3 by summarizing the findings from later chapters. I conclude with a discussion of further research directions in the area of co-located collaborative data analysis work with information visualizations.

8.1 Research Objectives

My work focused on providing a richer picture of the design challenges for collaborative information visualization systems for large shared displays and synchronous work. In a group setting, the use of co-located collaborative technology needs to support a process of social interaction around the data. Ideally, this process helps the group to arrive at a common understanding of the data through collaborative interpretation, analysis, discussion, and interaction. This means that by using these tools, groups should be able to gain more than the simple combination of each team members' individual insight from the data by taking advantage of knowledge sharing, interpretation, and discussions around the data. However, in order for team members to arrive at a common understanding it is necessary to allow them to transition from individual data analysis phases to joint data analysis phases in which the data is discussed in more closely coupled collaboration. In information analysis tasks, team members may need to be able to work on their own sub-projects, in which tentative hypotheses or exploration paths can be followed and rejected. Desire for private work may be in tension with their desire to benefit from the team's shared effort. Therefore, I specifically focused on the challenges of designing collaborative interfaces that support both individual as well as joint work practices around data. My goal was to identify the needs and requirements for transitioning between both work phases in data analysis tasks with the overarching goal to help the group to more easily arrive at a common understanding of the data. In context of this research goal, I looked at three main research challenges. The next section summarizes the progress towards addressing the individual challenges.

8.2 Progress on Thesis Contributions

The main objective of this thesis was to provide a richer understanding of how to best design information visualization systems for co-located collaborative work, with a focus on the problem of supporting individual as well as joint data analysis phases. I completed this objective by addressing three research challenges. Next, I provide summaries on how I addressed each research challenges.

8.2.1 Applicability of Related Work to Collaborative Data Analysis

Challenge 1: We do not have a clear understanding of how related work in CSCW and Information Visualization areas applies to the specific problem of supporting co-located data analysis.

I have addressed this research challenge by conducting a large literature analysis (Chapter 3) of work from computer supported cooperative work, information visualization advice, and empirical work that looks directly at collaborative use of information visualization.

This analysis was the first to summarize a set of design considerations for co-located collaborative data analysis work for large displays. The design considerations were grouped into those related to: setting up a collaborative environment (display & input), supporting social interaction around data, and designing information visualizations for co-located collaboration. The studies conducted in Chapters 4, 6, and 7 further confirmed earlier work in other task domains or collaborative settings that stated the tendency for team members to switch between closely and loosely coupled work phases (e. g., Elwart-Keys et al. (1990); Mandviwalla and Olfman (1994); Gutwin and Greenberg (1998); Tang et al. (2006)). The research studies conducted in Chapter 4 further partially confirmed earlier work on data analysis processes. In particular, the work of the Sensemaking (or Knowledge Crystallization) Cycle (Card et al., 1999) showed to be highly applicable. It predicted some of our findings as to the temporal flexibility of collaborative data analysis processes. On the other hand, the study contradicts previous work (Mark and Kobsa, 2005) which supported a more linear process model. The study conducted in Chapter 6 on CoCoNutTrix confirmed previous work (Tse et al., 2004) which argued that conflicts of interaction are often successfully resolved socially without the need to implement specific control policies.

In summary, research on other collaborative contexts (e. g., distributed collaboration), other collaborative tasks (e. g., collaborative drawing, document editing), as well as information visualization advice (e. g., the Sensemaking Cycle) for single analysts has applicability to co-located collaborative data analysis. This dissertation contributes a richer understanding of which work applies and contributes empirical evidence confirming previous research from other work and task domains.

8.2.2 Understanding Collaborative Data Analysis Practices

Challenge 2: We do not well understand the data analysis practices and processes of small teams. How do they analyze information together and how are information visualizations used in this context?

I have addressed this research challenge by conducting several studies of co-located collaborative data analysis. In the first study (Chapter 4), we used paper-based visualizations to study how individuals and small teams would analyze data unencumbered by any specific implementation of a collaborative system. The study revealed several analysis processes which included different types of data analysis activities that team members engaged in (Browse, Parse, Discuss Collaboration Style, Establish Task-Specific Strategy, Clarify, Select, Operate, and Validate). Some of these processes were more evident during specific collaboration styles than others. Specifically, Clarify, Strategize, Discuss Collaboration Style, and Validate were predominantly performed in joint (closely coupled) work phases with frequent verbal exchange and sharing of artifacts while activities involving the Select, Operate, Parse, and Browse process were frequently performed in individual, parallel work with own information items and limited direct communication. We, thus, saw that participants often performed detailed work that involved careful scrutiny of the information artifacts in parallel work phases, while team members more often asked the help or opinion of others when making plans of how to proceed, clarifying how to read a visualization, or validating intermediate or final results. The study of CoCoNutTrix in Chapter 6 further looked at types of interactions between collaborators over one shared visualization. Again, we observed team members switching between parallel and joint work phases. Even though participants had to share the same visualization, their interactions rarely conflicted with one another. However, as participants did often pick separate parts of the visualization to work on in parallel, they often lost the overview of what had been worked on by the group. While the visualization itself could capture (through representation changes) parts of the teams' analysis activities, participants requested additional awareness information. This need for additional awareness information was further studied in Chapter 7. This final study showed again, the variability in work styles that participants engaged in. Teams spent between 32% and 92% of their time closely coupled. The large variability in work styles, coupled with variability in task solving strategies and, thus, tool use, meant that we did not find significant differences of tool use behaviour between groups in the different awareness conditions. However, qualitative feedback from teams without awareness or partial awareness indicators showed the need to embed such features into data representations for co-located collaborative work for this type of task.

In summary, these three studies provide evidence of the flexibility in which teams engage in complex data analysis tasks with information visualizations. The results from these studies contribute to our understanding of temporal flexibility in data analysis processes and activities (Chapter 4) and provide evidence of the variability of closely and loosely coupled collaboration during these tasks (Chapters 4, 6, and 7). Further, the results show that this switching is important to support group discussions and validations, and hence, group knowledge building. Further, the thesis contributes a set of design considerations that can help support this flexibility and provides evidence how well the flexibility was supported in the two studied systems (Chapters 6 and 7).

8.2.3 Designing Collaborative Data Analysis Systems

Challenge 3: We do not know how to design collaborative information visualization systems for co-located work. In particular, we do now know how we can support team members transitioning from parallel to joint work phases.

I addressed this research challenge by presenting three different system designs for co-located collaborative data analysis. Each had different data characteristics, group sizes, hardware configurations, or analysis tasks. Based on the initial set of design considerations derived in Chapters 3 and 4, CoTree (Chapter 5) was developed as a first system that translated these initial guidelines into a co-located collaborative information visualization system. Specifically, it supports parallel work by providing individual and unconnected views of the data, mechanisms for flexible workspace organization (including view rotations, scaling, and positioning), and flexible temporal work processes. Collaborators could use these features to rearrange and resize items in the workspace to transition to more joint work.

With CoCoNutTrix (Chapter 6), we took a different approach, starting from a system that would support joint work and introducing mechanisms for parallel work. We retrofitted a social network analysis system (NodeTrix by Henry et al. (2007)) for collaborative work. While it included one shared large representation, parallel work was made possible by introducing synchronous independent inputs, gestural interaction, and by minimizing global data operations. Team members could synchronously interact with the representations as each data item could be individually operated on, while still being connected to the remainder of the graph representation.

Finally, Cambiera (Chapter 7) was designed as a new collaborative visual analytics tool for the analysis of large text document collections. Similar to CoTree and CoCoNutTrix, it allowed team members to work in parallel on individual information items and search result visualizations but in contrast included more direct indications of overlap between parallel work activities. In Cambiera, search results and individual document representations are augmented with awareness information to encourage transitioning to more closely coupled work phases.

The studies of both CoCoNutTrix and Cambiera showed that the respective systems supported different working styles which justified the design choices made for each system.

In summary, the dissertation contributes three specific system designs for the support of co-located data analysis activities. The designs cover a range of different data (hierarchical, social network, and textual documents), different work environments (a large two-touch Smart DViT table, a wall display with four independent indirect inputs, and a small multi-touch table from Microsoft), and different analysis tasks (comparison of hierarchical data, social network analysis, intelligence analysis). Together, the three designs contribute to our evolving understanding of how to support the flexible work styles that are common in collaborative data analysis work.

8.3 Thesis Contributions

With this thesis I have enriched our understanding of the problems and challenges in designing information visualizations for co-located collaborative work. Below I separate out the main components of my contributions.

8.3.1 Major Contributions

- **The first set of design considerations for information visualization systems in co-located shared screen settings**. To the best of my knowledge, the design considerations derived in this dissertation are the first collection of such considerations for co-located collaborative data analysis with information visualizations. While such a collection can always be extended and refined through further studies, this first collection can help to guide interested tool designers towards the development of their own first prototypes. Furthermore, it can encourage researchers to study the role of their own tools in other collaborative data analysis scenarios, with new datasets, and different tasks and to extend and refine these considerations.

- **An analysis of information analysis processes and their common occurrences in phases of joint and parallel work**. This study taught us that, as visualization tool designers, we should allow for individuals' unique approaches toward analysis, and support a more flexible temporal flow of activity. It also taught us that it may be worthwhile to consider specific support for analysis processes that may be more commonly performed in closely coupled collaboration with joint information artifacts.
- **Three specific system designs.** Based on three examples, I demonstrate how co-located collaborative systems can be designed to support group analysis. I show how both parallel and joint work phases can be supported in these different systems and present experimental findings that assess two of the presented designs. These three examples can serve as inspirational examples to others who are interested in building their own tools. The study methods used to assess two of the designs can help to refine our knowledge of how such systems can best be evaluated.

8.3.2 Minor Contributions

- A definition of collaborative visualization that encompasses many different collaborative scenarios around shared visualizations.
- Discussion of design challenges in adapting information visualizations for touch interaction. Different solutions are offered based on local menus, workspace menus, and gestural interaction.
- A new tree comparison mechanism based on spatial proximity, allowing any number of trees to be chained together and compared.
- Discussion of challenges and advice on retrofitting single-user information visualization tools for collaborative work.
- Introduction of the 'collaborative brushing and linking' concept as an awareness technique for collaborative work with information visualizations.
- Design of a visual analytics system for the analysis of large text document collections on a multi-touch surface.

8.4 Extending the Design Guidelines

The research presented in Chapters 4–7 extended the first set of design considerations derived from related literature in Chapter 3. Specifically in the context of co-located collaborative work, the following additional design considerations emerged.

Support for Temporal Flexibility of Analysis Processes: In all three studies conducted for this dissertation, groups' flexible work styles were evident. Chapter 4 showed that people in a team flexibly transition between types of data analysis processes, and that support of temporal flexibility of work processes is necessary for groups to work together fluidly. In order for team members to be able to work in parallel, systems need to be relatively unrestrictive in their parallel support of different data analysis activities. It is insufficient to support just one analysis activity at a time, since different team members employ different strategies to analysis, often in parallel, and pairs frequently shift activities depending on the current stage of the task or depending on recent findings that were made. For example, global operations which hinder every team member from continuing their parallel work may have to be minimized. Operations such as search for data and operations on the data may support people's work styles better when they can be performed in parallel (i. e., when data sources are accessible and browseable, no matter if important data calculations are being performed in parallel).

Support for Data Analysis in Closely and Loosely Coupled Collaboration: The studies described in Chapters 4, 6, and 7 provided evidence that teams frequently switch between joint and parallel work phases in collaborative data analysis tasks. Data analysis and, in particular exploratory analysis, often requires team members to react to emerging findings or hypotheses, to re-assess their analysis strategies and previous solutions, and to communicate, validate, and clarify information to varying degrees with the other team members. Hence, the support of a wide variety of work styles and collaboration strategies is a challenge that must be addressed for the design of collaborative data analysis tools. With the study from Chapter 4 we found that specific data analysis processes were more common during joint and some were more common during parallel work phases. It is worthwhile to consider how best to support the activities, for example, of validation and clarification in the context of closely coupled work in which team members often shared their current results and compared information they had already found. Similarly, browsing and operating on the data was often a parallel activity, which may benefit from features which support personalization and individual activity.

Integration of Awareness Indicators: The study on CoCoNutTrix in Chapter 6 pointed out that team members were often concentrated on detailed work in separate areas of the workspace and could not adequately keep track of which information others had worked on and made decisions about. Even though team members shared the same screen, the nature of the task and the size of the group made it difficult to keep track of who had worked on what and which decisions had been made about the data. The need to support awareness for co-located data analysis work was further studied with Cambiera in Chapter 7. In this study, we saw that team members highly valued the available awareness information or missed it when it was not present. Further, we noted that teams that spent more time closely coupled were better able to make connections in the data and join their individual findings. This finding suggests that collaborative data analysis work may benefit from an integration of additional awareness information in co-located settings.

Design for Collaborative Reasoning and Sensemaking: While the systems presented in this thesis allowed for collaborative data analysis and exploration, they only included limited support for collaborative reasoning and sensemaking. For example, annotating mechanisms can not only provide information on what has been worked on but also support reasoning about the data and collaborative sensemaking. Other possible features could include ways to extract relevant information for the group to consider next (based on what has already been explored), and ways to automatically summarize semantic information (that the group has considered). In the spirit of supporting closely coupled collaboration, such features may focus on the collaborative aspect of the work and compare and connect findings and facts other team members have extracted. This need for additional sensemaking and reasoning features was particularly evident in the Cambiera study (Chapter 7) where large amounts of information had to be connected and interpreted.

With these extensions, Table 3.1 from Chapter 3 can be extended as seen in Table 8.1

Consideration	Aspects to Consider
	Collaborative Environment
Display size	Socially appropriate work space size per person, establishment of private, group, and storage spaces
Display configuration	Accommodation of group's current work practices, tasks, and goals
Input Type	Impact of input type on possible interactions
Resolution	Input and display resolution
	Supporting Social Interaction
Communication	Explicit data referencing across different representational and viewing contexts, e. g., annotation; implicit awareness cues of changes to the data across different representational and viewing contexts, **support clarification, validation, data and strategy discussions with shared artifacts**
Coordination	When using individual data views, location and rotation as a coordination and communication tool, sharing of visualizations and views, multiple synchronous interactions with shared representations, **temporal flexibility of analysis processes**, **analysis processes within different collaboration styles**, **awareness of analysis activities and histories to encourage joint work**
	Designing Information Visualizations
Representation	Personal preferences, multiple representation types, awareness support, appropriateness of representation for work environment and social interaction, **integrated reasoning and sensemaking support**
Presentation	Arrangement of data items for group access, providing copies of the same data, accommodation of input methods, compensations for display resolution
View	Interpretability of data from multiple viewpoints and orientations
Interaction	Interactive response rates despite simultaneous interaction, collaborative interaction histories, conflict reduction arising from global changes to data or view, fluid interaction, **temporal flexibility of analysis processes**, **support operation, selection, browsing, and parsing as parallel activities**

Table 8.1: Summary of the design considerations for co-located collaborative data analysis environments derived in this dissertation. In bold are the extensions to the table presented in Chapter 3.

8.5 Generalizability

In the studies conducted as part of this dissertation I have mostly taken a qualitative approach looking at people's interactions with one another and with physical or digital items in their workspace. Every study of this type is influenced by the observer's interests, their perspective, biases, training, and knowledge which in turn may change and evolve during the course of an analysis (Jordan and Henderson, 1995; Corbin and Strauss, 2008). In order to support my initial observations I captured each analysis session and went over the video and audio data several times. To further support my observations, I then used systematic data analysis approaches as outlined for each study (see Chapter 4, 6, and 7). These systematic approaches involved coding the data (e. g., using affinity diagrams) and counting of observations from the video- and audio-logs. In addition, I made use of a variety of available data sources for each study beyond the observations: interview data (and their transcripts), data logs (Chapter 6 and 7), and questionnaires to triangulate my findings. By looking at these different sources of data I proceeded inductively and generated rich descriptions about general patterns found. These descriptions formed the basis for more general observations about how people work with information visualizations in groups, how their work patterns do not follow rigorous temporal orders, and how several features such as awareness indicators or support for specific work styles in closely or loosely coupled collaboration can further support this work. These more general observations should not be seen as prescriptive statements but rather as considerations for application to other work contexts, datasets, groups, or analysis tasks. While my initial finding from Chapter 4 about the temporal and work style dynamics of groups was further confirmed in the studies of Chapter 6 and 7, I cannot generalize them to all other data analysis scenarios. However, there are now three examples of this type of behavior and this may indicate that this type of behavior may extend to other (likely not all) data analysis scenarios. Similarly, I drew conclusions about the successfulness of the CoCoNutTrix and Cambiera tool for supporting the collaborative analysis tasks asked of the participants in the studies. Both tools successfully supported collaborative data analysis work in that most of the teams reached satisfying solutions and no major conflicts arose (see (Gutwin and Greenberg, 2000)). Of course, it is difficult to separate out the tool from the people who use it. The successful completion of the tasks was not only due to the tool and its design but certainly also due to factors such as teams' motivations and abilities to work with one another and with the tool. Studies such as the ones presented in this thesis serve to provide rich descriptions of the scenario under study and help to identify different types of behaviors and interactions. Further studies in this research area will help to see in which other scenarios the findings from this thesis apply or do not apply.

8.6 Future Research

This research raises many new questions for collaborative data analysis. While I have given several considerations for the design of collaborative information visualizations, I looked at only a small subset of possible collaborative work scenarios, hardware, data, and tasks. There are many interesting open research directions in this domain and the following sections summarize just a few of them.

8.6.1 Study of Data Analysis Practices in Context

While post-implementation evaluation of information visualizations are becoming more common, as of yet, there are few empirically-based information visualization papers that focus on describing theories of visual information analysis practices in a real world context. However, having a clearer understanding of how people work together collaboratively with information visualizations, what type of information they share, in which format, and which activities rely on the input of several team members can be used to hypothesize about how people will interact with technology, to inform interface design, and to guide evaluation of created designs (Isenberg et al., 2008b).

For example, an observational study conducted to inform design and evaluation was performed by Tang and Carpendale (Tang and Carpendale, 2007). Their study looked at the information exchange required during nurses' shift changes. The results provided a decomposition of the types of information being exchanged in various media, and potential avenues for computer support including information visualizations. In scenarios such as these, it is important to understand current practices, as new information visualizations may offer improved efficiencies in some measured criteria, but their net impact on patient care needs to considered. For example, if a highly time-efficient digital system all but eliminates temporal overlap during hand-overs then some important verbal exchanges could be lost. Hence, studying collaborative practices with information visualizations is important to inform our understanding of when and where these systems can be used, which specific types of tasks they should best support, and how these systems can be integrated into the daily work routines of the individual team members.

8.6.2 Evaluation of Collaborative Data Anlaysis

The success of an information visualization is strongly connected to the mental model that a person can make about the data by viewing the visualization (Spence, 2007b). However, our understanding of how this mental model formation works is still very limited and we know even less about how a group forms an understanding or insight of a dataset. In my opinion, the goal of using a collaborative information visualization system should be to provide the group with an environment that enriches their data analysis activities beyond what they could come up with as separate individuals. A collaborative data analysis scenario should, therefore, support group insight formation. However, as already hinted at in Chapter 6, measuring group insight (or even individual insight as pointed out by Plaisant (2004) and Saraiya et al. (2005)) is difficult to measure. Work by Stahl (2006) closely relates to this problem. His book describes research on group cognition in the computer supported collaborative learning domain. Stahl argues for attempting to capture the richness of the learning phenomenon by observing the collaboration taking place. He mentions that collaborative knowledge building is a complex and subtle process that cannot adequately be captured with statistical methods and I believe the same to be true for collaborative building of insight. Similar to the problems inherent in evaluating single-user information visualizations, we do not have a clear idea about how to evaluate the possible additional insights or the group learning effect that can be achieved by using such a system. How do you capture group insight or learning? Is the group even important for the construction of insight in the individual? If so, how do we find out? (Stahl, 2006) proposes to observe team members' conversations about data discoveries. Where do they agree or disagree, augment or confirm each other? The advantage of observing collaborative formation of insight vs. insight made by a single person is that group members may have to make these processes visible to each other and may also make it visible to the observer. As more collaborative systems are built for co-located data analysis, different methods may have to be tried to evaluate of these systems. In this thesis, we have taken a mostly observational, qualitative approach to analysis which let us gain a richer understanding of the overall use of a tool and the behaviours of team members during the collaboration. This worked well for CoCoNutTrix (Chapter 6) to describe how different components of collaborative work were evident, but did not work well when trying to find specific differences for tool features in the Cambiera study (Chapter 7). Here, team members' individual approaches were too varying to measure any significant differences, yet, providing evidence of the usefulness of such features may be more convincing to promote their adoption in practice. How to address the variability in group work styles, collaborative tasks, data, and scenarios for any controlled experiment of collaborative data analysis tools is a challenge that is yet to be addressed.

8.6.3 Extending the Work to Other Contexts

In this thesis, only a limited subset of possible collaborative data analysis scenarios were addressed. Natural extensions include the following:

Small vs. Large Teams: Data sets nowadays are becoming increasingly large and complex. It is possible that the analysis of such data requires larger teams than the 2–4 team members that were considered here.

Joint Synchronous and Asynchronous Settings: Collaborative work is not always started from a planned or arranged meeting time but often happens ad-hoc. Imagine a scenario where two people are working on a shared large display in an office on a data analysis task. It may be quite frequent that another person happens to notice the collaboration by overhearing a comment that he/she is interested in. In cases such as this one, it may be beneficial if a collaborative system could support a person entering a collaboration at a later point in time, show which data items team members had already looked at, see what they have already edited or made decisions about, and which are left open to decide without interrupting the work flow of those already in a collaboration.

Privacy in Co-located Collaborative Analysis: Certain datasets that benefit from collaborative analysis may include data that only subsets of the team members have access rights to. The access could restrict who can see information or who can modify information. How to present and include information with restricted access rights to a collaborative data analysis team is an open question. Possible solutions could include anonymizing or abstracting the restricted information or displaying pointers to further information. This further information could be relayed to external screens (notebooks, PDAs, etc.) where it could be viewed or interacted with in private.

Communication vs. Data Analysis: In this dissertation, I have merely looked at the issues regarding data analysis on shared displays. However, information visualizations are increasingly used to communicate findings or concepts to larger groups of people. What requirements arise for the design of information visualizations to guide learning and understanding the information is an open research question.

8.7 Conclusion

With this dissertation, I have moved our understanding of the design requirements and challenges for co-located collaborative data analysis another step forward. Over the course of this dissertation I have seen increasing interest in using large shared displays for data analysis work (about 90 participants joined the 2009 VisWeek workshop, that I co-organized). In particular, the first design considerations derived here will help practitioners to start their own designs and to extend and refine the considerations to other domains, tasks, and scenarios.

BIBLIOGRAPHY

Maneesh Agrawala, Andrew C. Beers, Ian McDowall, Bernd Fröhlich, Mark Bolas, and Pat Hanrahan. The Two-User Responsive Workbench: Support for Collaboration Through Individual Views of a Shared Space. In *Proceedings of the Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 327–332, New York, NY, USA, 1997. ACM Press.

Robert A. Amar and John T. Stasko. Knowledge Precepts for Design and Evaluation of Information Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 11(4):432–442, 2005.

Saleema Amershi and Meredith Ringel Morris. CoSearch: A System for Co-located Collaborative Web Search. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 1647–1656, New York, NY, USA, 2008. ACM Press.

Cheong S. Ang, David C. Martin, and Michael D. Doyle. Integrated Control of Distributed Volume Visualization Through the World-Wide-Web. In *Proceedings of the Conference on Visualization (VIS)*, pages 13–20, Los Alamitos, CA, USA, 1993. IEEE Computer Society.

Vinud Anupam, Chandrajit Bajaj, Daniel Schikore, and Matthew Schikore. Distributed and Collaborative Visualization. *Computer*, 27(7):37–43, 1994.

Georg Apitz and François Guimbretière. CrossY: A Crossing-based Drawing Application. In *Proceedings of the Symposium on User Interface Software and Technology (UIST)*, pages 3–12, New York, NY, USA, 2004. ACM Press.

Kevin Baker, Saul Greenberg, and Carl Gutwin. Heuristic Evaluation of Groupware Based on the Mechanics of Collaboration. In *Engineering for Human-Computer Interaction*, volume 2254 of *LNCS*, pages 123–139. Springer Verlag, Berlin, Heidelberg, 2001.

Patrick Baudisch, Edward Cutrell, and George Robertson. High-Density Cursor: A Visualization Technique that Helps Users Keep Track of Fast-Moving Mouse Cursors. In *Proceedings of the Conference on Human-Computer Interaction (INTERACT)*, pages 236–243, Amsterdam, The Netherlands, 2003. IOS Press.

Benjamin B. Bederson, Jon Meyer, and Lance Good. Jazz: An Extensible Zoomable User Interface Graphics Toolkit in Java. In *Proceedings of the Symposium on User Interface Software and Technology (UIST)*, pages 171–180, New York, NY, USA, 2000. ACM Press.

Hrvoje Benko, Andrew D.Wilson, and Patrick Baudisch. Precise Selection Techniques for Multi-Touch Screens. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 1263–1272, New York, NY, USA, 2006. ACM Press.

Jacques Bertin. *Semiology of Graphics: Diagrams Networks Maps*. The University of Wisconsin Press, Madison, USA, 1983.

Eric A. Bier, Steve Freeman, and Ken Pier. MMM: The Multi-device Multi-user Multi-editor. In *Proceedings of CHI*, pages 645–646, New York, 1992. ACM Press.

Susan E. Brennan, Klaus Mueller, Greg Zelinsky, IV Ramakrishnan, David S. Warren, and Arie Kaufman. Toward a Multi-Analyst Collaborative Framework for Visual Analytics. In *Proceedings of the Symposium on Visual Analytics Science and Technology (VAST)*, pages 129–136, Los Alamitos, CA, USA, 2006. IEEE Computer Society.

Andreas Buja, John Alan McDonald, John Michalak, and Werner Stuetzle. Interactive Data Visualization using Focusing and Linking. In *Proceedings of Visualization (VIS)*, pages 156–163, Los Alamitos, CA, USA, 1991. IEEE Computer Society.

Stuart Card, Jock D. Mackinlay, and Ben Shneiderman, editors. *Readings In Information Visualization: Using Vision To Think*. Morgan Kaufmann Publishers, San Francisco, CA, USA, 1999.

Stuart K. Card and David Nation. Degree-of-Interest Trees: A Component of an Attention-Reactive User Interface. In *Proceedings of the Conference on Advanced Visual Interfaces (AVI)*, pages 231–245, New York, NY, USA, 2002. ACM Press.

John M. Carroll, Mary Beth Rosson, Gregorio Convertino, and Craig H. Ganoe. Awareness and Teamwork in Computer-Supported Collaborations. *Interacting with Computers*, 18(1):21–46, 2006.

Chaomei Chen. *Information Visualization: Beyond the Horizon*. Springer-Verlag, London, 2nd edition, 2006.

Ed Huai-Hsin Chi and John T. Riedl. An Operator Interaction Framework for Visualization Systems. In *Proceedings of the Symposium on Information Visualization (InfoVis)*, pages 63–70, Los Alamitos, CA, USA, 1998. IEEE Computer Society.

Yoo-Joo Choi, Soo-Mi Choi, Seon-Min Rhee, and Myoung-Hee Kim. Collaborative and Immersive Medical Education in a Virtual Workbench Environment. In *Proceedings of the Conference on Knowledge-Based Intelligent Information and Engineering Systems (KES)*, volume 3683/2005 of *Lecture Notes in Computer Science*, pages 1210–1217. Springer-Verlag, 2005.

Mei C. Chuah and Steven F. Roth. Visualizing Common Ground. In *Proceedings of the Conference on Information Visualization (IV)*, pages 365–372, Los Alamitos, CA, USA, 2003. IEEE Computer Society.

Herbert H. Clark. *Using Language*. Cambridge University Press, Cambridge, UK, 1996.

Herbert H. Clark and Susan E. Brennan. Grounding in Communication. In Ronald M. Baecker, editor, *Readings in Groupware and Computer-Supported Cooperative Work: Assisting Human-Human Collaboration*, chapter Groupware Design and Evaluation Methodologies, pages 222–234. Morgan Kaufmann Publishers Inc., San Francisco, 1993.

Juliet Corbin and Anselm Strauss. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Sage Publications, Thousand Oaks, CA, USA, 2008.

John W. Creswell. *Qualitative Inquiry and Research Design: Choosing Among Five Traditions*. Sage Publications, Inc., Thousand Oaks, London, New Delhi, 1998.

Catalina M. Danis, Fernanda B. Viégas, Martin Wattenberg, and Jesse Kriss. Your Place or Mine? Visualization as a Community Component. In *Proceedings of the Conference* *on Human Factors in Computing Systems (CHI)*, pages 275–284, New York, NY, USA, 2008. ACM Press.

Paul Dietz and Darren Leigh. DiamondTouch: A Multi-User Touch Technology. In *Proceedings of the Symposium on User Interface Software and Technology (UIST)*, pages 219–226, New York, NY, USA, 2001. ACM Press.

Dmitry Dimov and Brian Mulloy. Swivel. Website, 2009. `http://www.swivel.com/` (last accessed: February, 2009).

Alan J. Dix, Janet E. Finlay, Gregory D. Abowd, and Russell Beale. *Human-Computer Interaction*. Prentice Hall, London, New York, and others, 1998.

Paul Dourish and Victoria Bellotti. Awareness and Coordination in Shared Workspaces. In *Proceedings of the Conference on Computer Supported Cooperative Work (CSCW)*, pages 107–114, New York, NY, USA, 1992. ACM Press.

Paul Dourish and Sara Bly. Portholes: Supporting Awareness in a Distributed Work Group. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 541–547, New York, NY, USA, 1992. ACM.

Paul Dourish and Matthew Chalmers. Running Out of Space: Models of Information Navigation. In *Proceedings of HCI, Panel Statement*, 1994.

Pierre Dragicevic and Jean-Daniel Fekete. Support for Input Adaptability in the ICON Toolkit. In *Proceedings of ICMI*, pages 212–219, New York, NY, USA, 2004. ACM Press.

Jim Durbin, J. Edward Swan II, Brad Colbert, John Crowe, Rob King, Tony King, Christopher Scannell, Zacharz Wartell, and Terry Welsh. Battlefield Visualization on the Responsive Workbench. In *Proceedings of the Conference on Visualization (VIS)*, pages 463–466, Los Alamitos, CA, USA, 1998. IEEE Computer Society.

Mary Elwart-Keys, David Halonen, Marjorie Horton, Rober Kass, and Paul Scott. User Interface Requirements for Face to Face Groupware. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 295–301, New York, NY, USA, 1990. ACM Press.

Jean-Daniel Fekete. The InfoVis Toolkit. In *Proceedings of the Symposium on Information Visualization (InfoVis)*, pages 167–174, Los Alamitos, CA, USA, 2004. IEEE Computer Society.

Danyel Fisher. Hotmap: Looking at Geographic Attention. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):6, 2007.

Clifton Forlines, Alan Esenther, Chia Shen, Daniel Wigdor, and Kathy Ryall. Multi-User, Multi-Display Interaction with a Single-User, Single-Display Geospatial Application. In *Proceedings of the Symposium on User Interface Software and Technology (UIST)*, pages 273–276, New York, NY, USA, 2006. ACM Press.

Clifton Forlines and Ryan Lilien. Adapting a Single-User, Single-Display Molecular Visualization Application for Use in a Multi-User, Multi-Display Environment. In *Proceedings of the Conference on Advanced Visual Interfaces (AVI)*, pages 367–371, New York, NY, USA, 2008. ACM Press.

Clifton Forlines and Chia Shen. DTLens: Multi-user Tabletop Spatial Data Exploration. In *Proceedings of the Symposium on User Interface Software and Technology (UIST)*, pages 119–122, New York, USA, 2005. ACM Press.

Linton C. Freeman. Visualizing Social Networks. *Journal of Social Structure*, 1(1), 2000.

Mathias Frisch, Jens Heydekorn, and Raimund Dachselt. Investigating Multi-Touch and Pen Gestures for Diagram Editing on Interactive Surfaces. In *Proceedings of the Conference on Interactive Tabletops and Surfaces (ITS)*, pages 149–156, New York, NY, USA, 2009. ACM Press.

Ian J. Grimstead, David W. Walker, and Nick J. Avis. Collaborative Visualization: A Review and Taxonomy. In *Proceedings of the Symposium on Distributed Simulation and Real-Time Applications (DS-RT)*, pages 61–69, Los Alamitos, CA, USA, 2005. IEEE Computer Society.

Georges Grinstein, Theresa O'Connell, Sharon Laskowski, Catherine Plaisant, Jean Scholtz, and Mark Whiting. VAST 2006 Contest–A Tale of Alderwood. In *Proceedings of the Symposium on Visual Analytics Science and Technology (VAST)*, pages 215–216, Los Alamitos, CA, USA, 2006. IEEE Computer Society.

François Guimbretiére and Terry Winograd. FlowMenu: Combining Command, Cext, and Data Entry. In *Proceedings of the 1Symposium on User Interface Software and Technology (UIST)*, pages 213–216, New York, NY, USA, 2000. ACM Press.

Carl Gutwin and Saul Greenberg. Design for Individuals, Design for Groups: Tradeoffs between Power and Workspace Awareness. In *Proceedings of the Conference on Computer Supported Cooperative Work (CSCW)*, pages 207–216, New York, NY, USA, 1998. ACM Press.

Carl Gutwin and Saul Greenberg. The Mechanics of Collaboration: Developing Low Cost Usability Evaluation Methods for Shared Workspaces. In *Proceedings of the Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises (WETICE)*, pages 98–103, Los Alamitos, CA, USA, 2000. IEEE Computer Society.

Carl Gutwin and Saul Greenberg. A Descriptive Framework of Workspace Awareness for Real-Time Groupware. *Computer Supported Collabortive Work*, 11(3–4):411–446, 2002.

Mark Hancock and Sheelagh Carpendale. Supporting Multiple Off-Axis Viewpoints at a Tabletop Display. In *Proceedings of the Workshop on Horizontal Interactive Human-Computer Systems (Tabletop)*, pages 171–178, Los Alamitos, CA, USA, 2007. IEEE Computer Society.

Mark Hancock, Miguel Nacenta, Carl Gutwin, and Sheelagh Carpendale. The Effects of Changing Projection Geometry on the Interpretation of 3D Orientation on Tabletops. In *Proceedings of the Conference on Interactive Tabletops and Surfaces (ITS)*, pages 157–164, New York, NY, USA, 2009. ACM Press.

Mark S. Hancock and Kellogg S. Booth. Improving Menu Placement Strategies for Pen Input. In *Proceedings of Graphics Interface (GI)*, pages 221–230, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 2004. Canadian Human-Computer Communications Society.

Jeffrey Heer and Maneesh Agrawala. Multi-Scale Banking to 45 Degrees. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):701–708, 2006.

Jeffrey Heer and danah boyd. Vizster: Visualizing Online Social Networks. In *Proceeings of the Symposium on Information Visualization (InfoVis)*, pages 33–40, Los Alamitos, CA, USA, 2005. IEEE Computer Society.

Jeffrey Heer, Stuart K. Card, and James A. Landay. Prefuse: A Toolkit for Interactive Information Visualization. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 421–430, New York, NY, USA, 2005. ACM Press.

Jeffrey Heer, Frank van Ham, Sheelagh Carpendale, Chris Weaver, and Petra Isenberg. Creation and Collaboration: Engaging New Audiences for Information Visualization. In Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, and Chris North, editors, *Information Visualization—Human-Centered Issues and Perspectives*, volume 4950 of *LNCS State-of-the-Art Survey*, pages 92–133. Springer Verlag, 2008.

Jeffrey Heer, Fernanda B. Viégas, and Martin Wattenberg. Voyagers and Voyeurs: Supporting Asynchronous Collaborative Information Visualization. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 1029–1038, New York, USA, 2007. ACM Press.

Julie Heiser, Barbara Tversky, and Mia Silverman. Sketches For and From Collaboration. In *Visual and spatial reasoning in design III*. Key Centre for Design Research, Sydney, 2004.

Nathalie Henry, Jean-Daniel Fekete, and Michael J. McGuffin. NodeTrix: A Hybrid Visualization of Social Networks. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1302–1309, 2007.

William C. Hill, James D. Hollan, Dave Wroblewski, and Tim McCandless. Edit Wear and Read Wear. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 3–9, New York, NY, USA, 1992. ACM Press.

Juan Pablo Hourcade, Benjamin B. Bederson, and Allison Druin. Building KidPad: An Application for Children's Collaborative Storytelling. *Software: Practice & Experience*, 34(9):895–914, 2004.

Darrell Huff. *How to Lie with Statistics*. W. W. Norton & Company, Inc., 1954.

Stéphane Huot, Cédric Dumas, Pierre Dragicevic, Jean-Daniel Fekete, and Gérard Hégron. The MaggLite Post-WIMP Toolkit: Draw it, Connect it and Run it. In *Proceedings of the Symposium on User Interface Software and Technology (UIST)*, pages 257–266, New York, NY, USA, 2004. ACM Press.

Edwin Hutchins. *Cognition in the Wild*. MIT Press, 1996.

IBM. Atlas for Lotus Connections. Software, 2007.

iCharts Inc. iCharts. Website, 2008. http://icharts.net/ (last accessed: April, 2009).

Innovations in Visualization Laboratory. Large display framework. Software, 2009. http://innovis.cpsc.ucalgary.ca/Software/LDF.

Petra Isenberg, Anastasia Bezerianos, Nathalie Henry, Sheelagh Carpendale, and Jean-Daniel Fekete. CoCoNutTrix: A Study in Collaborative Retrofitting for Information Visualization. *Computer Graphics and Applications: Special Issue on Collaborative Visualization*, 29(5):44–57, 2009. To appear September / October 2009.

Petra Isenberg and Sheelagh Carpendale. Interactive Tree Comparison for Co-located Collaborative Information Visualization. *IEEE Transactions on Visualization and Computer Graphics (Proceedings Visualization / Information Visualization 2007)*, 13(6): 1232–1239, 2007.

Petra Isenberg and Danyel Fisher. Collaborative Brushing and Linking for Co-located Visual Analytics of Document Collections. *Computer Graphics Forum (Proceedings of EuroVis)*, 28(3):1031–1038, 2009.

Petra Isenberg, Anthony Tang, and Sheelagh Carpendale. An Exploratory Study of Visual Information Analysis. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 1217–1226, New York, USA, 2008a. ACM Press.

Petra Isenberg, Torre Zuk, Christopher Collins, and Sheelagh Carpendale. Grounded Evaluation of Information Visualizations. In *Proceedings of the CHI Workshop Beyond Time and Errors: Novel Evaluation Methods for Information Visualization (BELIV)*, New York, NY, USA, 2008b. ACM Press.

Tobias Isenberg, André Miede, and Sheelagh Carpendale. A Buffer Framework for Supporting Responsive Interaction in Information Visualization Interfaces. In *Proceedings of the Conference on Creating, Connecting and Collaborating through Computing (C^5)*, pages 262–269, Los Alamitos, CA, USA, 2006a. IEEE Computer Society.

Tobias Isenberg, Petra Neumann, Sheelagh Carpendale, Simon Nix, and Saul Greenberg. Interactive Annotations on Large, High-Resolution Information Displays. In

Conference Compendium of IEEE VIS, InfoVis, and VAST, pages 124–125, Los Alamitos, CA, USA, 2006b. IEEE Computer Society.

T. J. Jankun-Kelly, Oliver Kreylos, Kwan-Liu Ma, Bern Hamann, Kenneth I. Joy, John Shalf, and E. Wes Bethel. Deploying Web-based Visual Exploration Tools to the Grid. *IEEE Computer Graphics and Applications*, 23(2):40–50, 2003.

T. J. Jankun-Kelly, Kwan-Liu Ma, and Michael Gertz. A Model and Framework for Visualization Exploration. *IEEE Transactions on Visualization and Computer Graphics*, 13(2):357–369, 2007.

Hao Jiang, Daniel Wigdor, Clifton Forlines, Michelle Borkin, Jens Kauffmann, and Chia Shen. LivOlay: Interactive Ad-Hoc Registration and Overlapping of Applications for Collaborative Visual Exploration. In *Proceeding of the Conference on Human Factors in Computing Systems (CHI)*, pages 1357–1360, New York, NY, USA, 2008. ACM Press.

Jinput. Java Game Controller API. Website (`https://jinput.dev.java.net/`), 2008. Accessed: May.

Brian Johnson and Ben Shneiderman. Treemaps: A Space-filling Approach to the Visualization of Hierarchical Information Structures. In *Proceedings of the Conference on Visualization (VIS)*, pages 284–291, Los Alamitos, CA, USA, 1991. IEEE Computer Society.

Greg Johnson. Collaborative Visualization 101. *ACM Siggraph - Computer Graphics*, 32 (2):8–11, 1998.

Brigitte Jordan and Austin Henderson. Interaction Analysis: Foundations and Practice. *The Journal of Learning Sciences*, 4(1):39–103, 1995.

Paul E. Keel. Collaborative Visual Analytics: Inferring from the Spatial Organization and Collaborative Use of Information. In *Proceedings of the Symposium on Visual Analytics Science and Technology (VAST)*, pages 137–144, Los Alamitos, CA, USA, 2006. IEEE Computer Society.

Yoshifumi Kitamura, Takashige Konishi, Sumihiko Yamamoto, and Fumio Kishino. Interactive Stereoscopic Display for Three or More Users. In *Proceedings of the Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 231–240, New York, NY, USA, 2001. ACM Press.

Don N. Kleinmutz and David A. Schkade. Information Displays and Decision Processes. *Psychological Science*, 4(4):221–227, 1993.

Valdis Krebs. Spread of Contagions via Contact Tracing. Website report (http://www.orgnet.com/contagion.html), 2008.

Russell Kruger, Sheelagh Carpendale, Stacey D. Scott, and Saul Greenberg. Roles of Orientation in Tabletop Collaboration: Comprehension, Coordination and Communication. *Journal of Computer Supported Collaborative Work*, 13(5–6):501–537, 2004.

Russell Kruger, Sheelagh Carpendale, Stacey D. Scott, and Anthony Tang. Fluid Integration of Rotation and Translation. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 601–610, New York, NY, USA, 2005. ACM Press.

Wolfgang Krüger, Christian A. Bohn, Bernd Fröhlich, Heinrich Schüth, Wolfgang Strauss, and Gerold Wesche. The Responsive Workbench: A Virtual Work Environment. *Computer*, 28(7):42–48, 1995.

Wolfgang Krüger and Bernd Fröhlich. The Responsive Workbench. *IEEE Computer Graphics and Applications*, 14(3):12–15, 1994.

John Lamping, Ramana Rao, and Peter Pirolli. A Focus + Context Technique Based on Hyperbolic Geometry for Visualizing Large Hierarchies. In *Proceedings of the Conference of Human Factors in Computing Systems (CHI)*, pages 401–408, New York, NY, USA, 1995. ACM Press.

Jason Leigh, Andrew E. Johnson, Maxine Brown, Daniel J. Sandin, and Thomas A. DeFanti. Visualization in Teleimmersive Environments. *Computer*, 32(12):67–73, 1999.

Mary D. P. Leland, Robert S. Fish, and Robert E. Kraut. Collaborative Document Production Using Quilt. In *Proceedings of the Conference on Computer-Supported Cooperative Work*, pages 206–215, New York, NY, USA, 1988. ACM.

Lewis W. F. Li, Frederick W. B. Li, and Rynson W. H. Lau. A Trajectory-Preserving Synchronization Method for Collaborative Visualization. *IEEE Transactions of Visualization and Computer Graphics*, 12(5):989–996, 2006.

Munir Mandviwalla and Lorne Olfman. What Do Groups Need? A Proposed Set of Generic Groupware Requirements. *ACM Transactions on Computer-Human Interaction*, 1(3):245–268, 1994.

Gloria Mark, Keri Carpenter, and Alfred Kobsa. A Model of Synchronous Collaborative Information Visualization. In *Proceedings of the Conference on Information Visualization (IV)*, pages 373–381, Los Alamitos, USA, 2003. IEEE Computer Society.

Gloria Mark and Alfred Kobsa. The Effects of Collaboration and System Transparency on CIVE Usage: An Empirical Study and Model. *Presence*, 14(1):60–80, 2005.

David Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. W.H. Freeman & Company, 1982.

Ryuichi Matsukara, Koji Koyamada, Yasuo Tang, Yukihiro Karube, and Mitsuhiro Moriya. VizGrid: Collaborative Visualization Environment for Natural Interaction between Remote Researchers. *Fujitsu Scientific & Technical Journal*, 40(2):205–216, 2004.

MayaViz. Comotion®. Website, 2007. http://www.mayaviz.com/ (last accessed March 2007).

Barbara Mirel. *Interaction Design for Complex Problem Solving: Developing Useful and Usable Software*. Morgan Kaufmann Publishers, San Francisco, CA, USA, 2004.

Meredith Ringel Morris and Eric Horvitz. SearchTogether: An Interface for Collaborative Web Search. In *Proceedings of the Symposium on User Interface Software and Technology (UIST)*, pages 3–12, New York, NY, USA, 2007. ACM Press.

Tomer Moscovich, Fanny Chevalier, Nathalie Henry, Emmanuel Pietriga, and Jean-Daniel Fekete. Topology-aware navigation in large networks. In *Proceedings of the Conference on Human Factors in Computing Systems*, pages 2319–2328, New York, NY, USA, 2009. ACM.

Tamara Munzner. Guest Editor's Introduction: Information Visualization. *Computer Graphics and Applications*, 22(1):20–21, 2002.

Tamara Munzner, François Guimbretière, Serdar Tasiran, Li Zhang, and Yunhong Zhou. TreeJuxtaposer: Scalable Tree Comparison Using Focus+Context with Guaranteed Visibility. *ACM Transactions on Graphics*, 22(3):453–462, 2003.

Miguel A. Nacenta, David Pinelle, Dane Stuckel, and Carl Gutwin. The Effects of Interaction Technique on Coordination in Tabletop Groupware. In Christopher G. Healey and Edward Lank, editors, *Proceedings of Graphics Interface (GI)*, pages 191–198, Mississauga, ON, Canada, 2007a. Canadian Information Processing Society.

Miguel A. Nacenta, Satoshi Sakurai, Tokuo Yamaguchi, Yohei Miki, Yuichi Itoh, Yoshifumi Kitamura, Sriram Subramanian, and Carl Gutwin. E-conic: a perspective-aware interface for multi-display environments. In *Proceedings of the Symposium on User Interface Software and Technology (UIST)*, pages 279–288, New York, NY, USA, 2007b. ACM Press.

Petra Neumann, Stefan Schlechtweg, and M. Sheelagh T. Carpendale. ArcTrees: Visualizing Relations in Hierarchical Data. In *Proceedings of the Eurographics / IEEE VGTC Symposium on Visualization*, Eurographics Workshop Series, pages 53–60, 319, Aire-la-Ville, 2005. Eurographics.

Andreas Noack. Energy-Based Clustering of Graphs with Nonuniform Degrees. In *Proceedings of the Symposium on Graph Drawing (GD)*, pages 309–320, Berlin, 2005. Springer-Verlag.

Chris North, Tim Dwyer, Bongshin Lee, Danyel Fisher, Petra Isenberg, Kori Inkpen, and George Robertson. Understanding Multi-touch Manipulation for Surface Computing. In *Proceedings of Human-Computer Interaction (Interact)*, pages 236–249, Heidelberg, 2009. Springer Verlag.

Upul Obeysekare, Chas Williams, Jim Durbin, Larry Rosenblum, Rober Rosenberg, Fernando Grinstein, Ravi Ramamurti, Alexandra Landsberg, and William Sendberg. Virtual Workbench—A Non-Immersive Virtual Environment for Visualizing and Interacting with 3d Objects for Scientific Visualization. In *Proceedinfs of the Conference on Visualization (VIS)*, pages 345–350, Los Alamitos, CA, USA, 1996. IEEE Computer Society.

Gary M. Olson and Judith S. Olson. Distance Matters. *Human-Computer Interaction*, 15 (2 & 3):139–178, 2000.

Tom Paper. Data360. Website, 2009. `http://www.data360.org/` (last accessed: February, 2009).

Kyoung S. Park, Abhinav Kapoor, and Jason Leigh. Lessons Learned from Employing Multiple Perspectives In a Collaborative Virtual Environment for Visualizing Scientific Data. In *Proceedings of the Conference on Collaborative Virtual Environments (CVE)*, pages 73–82, New York, USA, 2000. ACM Press.

David Pinelle, Carl Gutwin, and Saul Greenberg. Task Analysis for Groupware Usability Evaluation: Modeling Shared-Workspace Tasks with the Mechanics of Collaboration. *ACM Transaction of Human Computer Interaction*, 10(4):281–311, 2003.

Peter Pirolli and Stuart K. Card. Information Foraging. *Psychological Review*, 106(4): 643–675, 1999.

Catherine Plaisant. The Challenge of Information Visualization Evaluation. In *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI)*, pages 109–116, New York, NY, USA, 2004. ACM Press.

Zachary Pousman, John T. Stasko, and MIchael Mateas. Casual Information Visualization: Depictions of Data in Everyday Life. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1145–1152, 2007.

Atul Prakash and Michael J. Knister. Undoing Actions in Collaborative Work. In *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW)*, pages 273–280, New York, NY, USA, 1992. ACM Press.

Rajeev R. Raje, Michael Boyles, and Shiaofen Fang. Cev: Collaborative Environment for Visualization using Java RMI. *Concurrency - Practice and Experience*, 10(11–13): 1079–1085, 1998.

Luc Renambot, Byungil Jeong, Hyejung Hur, Andrew Johnson, and Jason Leigh. Enabling High Resolution Collaborative Visualization in Display Rich Virtual Organizations. *Future Generation Computer Systems*, 25(2):161–168, 2009.

Ronald A. Rensink. *McGraw-Hill Yearbook of Science and Technology*, chapter Change Blindness, pages 44–46. McGraw-Hill, New York, NY, USA, 2005.

Meredith Ringel Morris, Andreas Paepcke, Terry Winograd, and Jeannie Stamberger. TeamTag: Exploring Centralized versus Replicated Controls for Co-located Tabletop Groupware. In *Proceedings of Human Factors in Computing Systems (CHI)*, pages 1273–1282, New York, NY, USA, 2006. ACM Press.

Meredith Ringel Morris, Kathy Ryall, Chia Shen, Clifton Forlines, and Frederic Vernier. Beyond "Social Protocols": Multi-User Coordination Policies for Co-located Groupware. In *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW)*, pages 262–265, New York, NY, USA, 2004. ACM Press.

George G. Robertson, Jock D. Mackinlay, and Stuart K. Card. Cone Trees: Animated 3D Visualizations of Hierarchical Information. In *Conference Companion of the Conference of Human Factors in Computing Systems, CHI'91*, pages 189–194, New York, NY, USA, 1991. ACM Press.

Anthony C. Robinson. Collaborative Synthesis of Visual Analytic Results. In *Proceedings of VAST*, pages 67–74, Los Alamitos, CA, USA, 2008. IEEE Computer Society.

Yvonne Rogers and Siân Lindley. Collaborating Around Vertical and Horizontal Large Interactive Displays: Which Way is Best? *Interacting with Computers*, 16(6):1133–1152, 2004.

Daniel M. Russell, Mark J. Stefik, Peter Pirolli, and Stuart K. Card. The Cost Structure of Sensemaking. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 269–276, New York, NY, USA, 1993. ACM Press.

Kathy Ryall, Alan Esenther, Clifton Forlines, Chia Shen, Sam Shipman, Meredith Ringel Morris, Katherine Everitt, and Frédéric Vernier. Identity-Differentiating Widgets for Multiuser Interactive Surfaces. *IEEE Computer Graphics and Applications*, 26 (5):56–64, 2006.

Purvi Saraiya, Chris North, and Karen Duca. An Insight-Based Methodology for Evaluating Bioinformatics Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 11(4):443–456, 2005.

Stacey D. Scott, M. Sheelagh T. Carpendale, and Stefan Habelski. Storage Bins: Mobile Storage for Collaborative Tabletop Displays. *IEEE Computer Graphics and Applications*, 25(4):58–65, 2005.

Stacey D. Scott, M. Sheelagh T. Carpendale, and Kori M. Inkpen. Territoriality in Collaborative Tabletop Workspaces. In *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW)*, pages 294–303, New York, NY, USA, 2004. ACM Press.

Stacey D. Scott, Karen D. Grant, and Regan L. Mandryk. System Guidelines for Co-located Collaborative Work on a Tabletop Display. In *Proceedings of the European Conference on Computer-Supported Cooperative Work (ECSCW)*, pages 159–178, Dordrecht, The Netherlands, 2003. Kluwer Academic Publishers.

Chia Shen, Neal B. Lesh, Frederic Vernier, Clifton Forlines, and Jeana Frost. Sharing and Building Digital Group Histories. In *Proceedings of the Conference on Computer Supported Cooperative Work (CSCW)*, pages 324–333, New York, NY, USA, 2002. ACM Press.

Ben Shneiderman. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *Proceedings of the Symposium on Visual Languages (VL)*, pages 336–343, Los Alamitos, USA, 1996. IEEE Comp. Society.

Ben Shneiderman and Catherine Plaisant. *Designing the User Interface – Strategies for Effective Human-Computer Interaction*. Addison-Wesley, 5th edition, 2009.

Smallthought Systems. DabbleDB. Website, 2009. `http://dabbledb.com/` (last accessed: February, 2009).

Robert Spence. A Framework for Navigation. *International Journal of Human-Computer Studies*, 51(5):919–945, 1999.

Robert Spence. *Information Visualization*. Pearson Education Limited, Harlow, England, 2nd edition, 2007a.

Robert Spence. *Information Visualization*. Pearson Education Limited, Harlow, England, 2nd edition, 2007b.

Gerry Stahl. *Group Cognition*. MIT Press, 2006.

John Stasko, Carsten Görg, and Zhicheng Liu. Jigsaw: Supporting Investigative Analysis Through Interactive Visualization. *Information Visualization*, 7:118–132, 2008.

John T. Stasko and Eugene Zhang. Focus+Context Display and Navigation Techniques for Enhancing Radial, Space-Filling Hierarchy Visualizations. In *Proceedings of the Symposium on Information Visualization (InfoVis*, pages 57–65, Los Alamitos, CA, USA, 2000. IEEE Computer Society.

John Stewart, Benjamin B. Bederson, and Allison Druin. Single Display Groupware: A Model for Co-present Collaboration. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 286–293, New York, NY, USA, 1999. ACM Press.

James Tam and Saul Greenberg. A Framework for Asynchronous Change Awareness in Collaborative Documents and Workspaces. *International Journal of Human-Computer Studies*, 64(7):583–598, 2006.

Peter Tandler, Thorsten Prante, Christian Müller-Tomfelde, Brobert Streitz, and Ralf Steinmetz. ConnecTables: Dynamic Coupling of Displays for the Flexible Creation of Shared Workspaces. In *Proceedings of the Symposium on User Interface Software and Technology (UIST)*, pages 11–20, New York, NY, USA, 2001. ACM Press.

Anthony Tang, Melanie Tory, Barry Po, Petra Neumann, and Sheelagh Carpendale. Collaborative Coupling over Tabletop Displays. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 1181–1290, New York, NY, USA, 2006. ACM Press.

Charlotte Tang and Sheelagh Carpendale. An Observational Study on Information Flow During Nurses' Shift Change. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 219–228, New York, NY, USA, 2007. ACM Press.

John C. Tang. Findings from Observational Studies of Collaborative Work. *International Journal of Man-Machine Studies*, 34(2):143–160, 1991.

James J. Thomas and Kristin A. Cook, editors. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society, 2005.

Matthew Tobiasz. Lark: Using Meta-Visualizations for Coordinating Collaboration. Master's thesis, University of Calgary, Calgary, AB, Canada, 2010.

Matthew Tobiasz, Petra Isenberg, and Sheelagh Carpendale. Lark: Coordinating Co-located Collaboration with Information Visualization. *IEEE Transactions on Visualization and Computer Graphics (Proceedings Visualization / Information Visualization 2009)*, 15(6):1065–1072, 2009.

Edward Tse and Saul Greenberg. Rapidly Prototyping Single Display Groupware through the SDGToolkit. In *Proceedings of the Australasian User Interface Conference*, pages 101–110, Darlinghurst, Australia, 2004. Australian Computer Society Inc.

Edward Tse, Jonathan Histon, Stacey D. Scott, and Saul Greenberg. Avoiding Interference: How People Use Spatial Separation and Partitioning in SDG Workspaces. In *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW)*, pages 252–261, New York, NY, USA, 2004. ACM Press.

Edward Rolf Tufte. *The Visual Display of Quantitative Information*. Graphic Press, Cheshire, Connecticut, USA, 2001.

Fred Vernier, Neal Lesh, and Chia Shen. Visualization Techniques for Circular Tabletop Interfaces. In *Proceedings of the Conference on Advanced Visual Interfaces (AVI)*, pages 257–263, New York, NY, USA, 2002. ACM Press.

Fernanda B. Viégas and Martin Wattenberg. Timelines Tag Clouds and the Case for Vernacular Visualization. *Interactions*, 15(4):49–52, 2008.

Fernanda B. Viégas, Martin Wattenberg, Frank van Ham, Jesse Kriss, and Matt McKeon. Many Eyes: A Site for Visualization at Internet Scale. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):1121–1128, 2007.

Visible Certainty. Verifiable.com. Website, 2009. `http://verifiable.com/` (last accessed: April, 2009).

Stephen Voida, Julie Stromer, Matthew Tobiasz, Petra Isenberg, and Sheelagh Carpendale. Getting Practical with Interactive Tabletop Displays: Designing for Dense Data, "Fat Fingers," Diverse Interactions, and Face-to-Face Collaboration. In *Proceedings of Tabletop and Interactive Surfaces (Tabletop)*, pages 109–116, New York, NY, USA, 2009. ACM Press.

Colin Ware. *Information Visualization – Perception for Design*. Morgan Kaufmann Series in Interactive Technologies. Morgan Kaufmann Publishers, Amsterdam, Boston, Heidelberg, and others, 2. edition, 2000.

Stanley Wassermann and Katherine Faust. *Social Network Analysis*. Cambridge University Press, 1994.

Martin Wattenberg. Baby Names, Visualization, and Social Data Analysis. In John Stasko and Matt Ward, editors, *IEEE Symposium on Information Visualization (InfoVis)*, pages 1–7, Los Alamitos, CA, USA, 2005. IEEE Computer Society.

Chris Weaver. Is Coordination a Means to Collaboration? Panel Statement, Proceedings.of the Conference on Coordinated & Multiple Views in Exploratory Visualization (CMV), 2007. Zürich, SZ.

Webster. Webster's english dictionary. Website `http://www.cs.chalmers.se/~hallgren/wget.cgi?presentation`, 2007. Accessed: November.

Etienne Wenger. *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press, 1999.

Gerold Wesche, Jürgen Wind, Martin Göbe, Larry Rosenblum, Jim Durbin, Robert Doyle, David Tate, Robert King, Bernd Fröhlich, Martin Fischer, Maneesh Agrawala, Andrew Beers, Pat Hanrahan, and Steve Bryson. Application of the Responsive Workbench. *IEEE Computer Graphics and Applications*, 17(4):10–15, 1997.

Daniel Wigdor, Hao Jiang, Clifton Forlines, Michelle Borkin, and Chia Shen. WeSpace: The Design Development and Deployment of a Walk-up and Share Multi-Surface Visual Collaboration System. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 1237–1246, New York, NY, USA, 2009. ACM Press.

Daniel Wigdor, Chia Shen, Clifton Forlines, and Ravin Balakrishnan. Perception of Elementary Graphical Elements in Tabletop and Multi-surface Environments. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 473–482, New York, NY, USA, 2007. ACM Press.

Wesley Willett, Jeffrey Heer, and Maneesh Agrawala. Scented Widgets: Improving Navigation Cues with Embedded Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1129–1136, 2007.

Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. User-Defined Gestures for Surface Computing. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 1083–1092, New York, NY, USA, 2009. ACM Press.

Jason Wood, Helen Wright, and Ken Brodlie. Collaborative Visualization. In *Proceedings of the Conference on Visualization (VIS)*, pages 253–259, Los Alamitos, CA, USA, 1997. IEEE Computer Society.

Mike Wu and Ravin Balakrishnan. Multi-Finger and Whole Hand Gestural Interaction Techniques for Multi-User Tabletop Displays. In *Proceedings of the Symposium on User Interface Software and Technology (UIST)*, pages 193–202, New York, NY, USA, 2003. ACM Press.

Christopher C. Yang, Nan Liu, and Marc Sageman. Analyzing the Terrorist Social Networks with Visualization Tools. In *Proceedings of the Conference on Intelligence and Security Informatics (ISI)*, pages 331–342, Berlin, 2006. Springer Verlag.

Beth Yost and Chris North. The Perceptual Scalability of Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):837–844, 2005.

Raul Niño Zambrano and Yuri Engelhardt. Diagrams for the Masses: Raising Public Awareness–From Neurath to Gapminder and Google Earth. In *Diagrammatic Representation and Inference*, volume 5223/2008. Springer-Verlag, Berlin / Heidelberg, 2008.

Jiajie Zhang and Donald A. Norman. Representations in Distributed Cognitive Tasks. *Cognitive Science*, 18(1):87–122, 1994.

Appendix A

A.1 Chapter 4: Materials for Information Analysis Processes Study

A.1.1 Informed Consent Form

The consent form on the following pages was handed to participants before the beginning of the study.

Department of Computer Science

Consent Form

Evaluation of Collaborative Information Analysis

Investigators: Petra Neumann, Anthony Tang, Sheelagh Carpendale

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

Description of Research Project:

We are currently investigating how people solve tasks that involve the exploration of data in graphical form, such as charts or graphs. We would like to learn how people solve these tasks to better meet the needs of people who have to do these tasks regularly. For this purpose, we will give you information in the form of graphs and tasks to solve with this information. The study will involve two different scenarios and graphs for these scenarios.

We will be observing your actions, as well as videotaping you during the course of the session. Videotaping is mainly done because it is difficult for us to observe everything that might be important for our research. We often discover things by watching the videos later that we have overlooked during your session. This videotaping, however, is optional and you can still participate if you choose not to be videotaped. You will also be asked to complete a post-session questionnaire to further our investigation. It is estimated that your involvement will take approximately one hour, and you will be offered remuneration for your time.

There are no known harms associated with your participation in this research. No information that discloses your identity will be released or published without your specific consent to disclosure. However, you may request your name to be cited in cases where we use your comments in a publication based on the study. All data received from this study will be stored in a locked cabinet and such information that will be stored on a computer will only be accessible through the use of a password. All data will be stored for a period of time no longer than five years. Information will be carefully disposed of (shredding for hard copies and deleting for electronic copies) when this investigation is complete.

You will be able to withdraw from this study at any point. If this occurs, any data collected up to that point about you will be discarded. You are also able to refuse to answer whatever questions you prefer to omit.

Informed Consent: Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in this research project and agree to participate as a participant. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal professional responsibilities. You are free to not answer specific items or questions in interviews or on questionnaires. You are free to withdraw from the study at any time without penalty. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, contact:

Petra Neumann, Department of Computer Science, University of Calgary
Phone: (403) 210-9499, pneumann@cpsc.ucalgary.ca

Sheelagh Carpendale, Department of Computer Science, University of Calgary
Phone: (403) 220-6055, sheelagh@cpsc.ucalgary.ca

Anthony Tang, Department of Electrical and Computer Engineering, University of British Columbia
Phone: (604) 822-4583, tonyt@ece.ubc.ca

If you have any questions not satisfactorily answered by the primary researchers concerning your participation in this project, you may contact Bonnie Scherrer in the Research Services Office, University of Calgary at (403) 220-3782; email bonnie.scherrer@ucalgary.ca.

Circle One

I grant permission to be audio taped: *Yes:* ___ *No:* ___

I grant permission to be videotaped: *Yes:* ___ *No:* ___

I grant permission to have anonymized video or still images of me used in a publication: *Yes:* ___ *No:* ___

I grant permission to be quoted anonymously in a publication: *Yes:* ___ *No:* ___

Participant's Name (please print legibly)	Participant Signature
Investigator/Witness	Date

A copy of this consent form will be given to you to keep for your records if you request it. This research has the ethical approval of the Department of Computer Science and the University of Calgary.

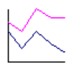
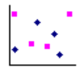
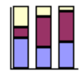
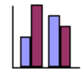
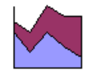
A.1.2 Questionnaire

The questionnaire on the following pages was handed to participants after completing the study tasks. Most relevant results from the questionnaire are discussed in Chapter 4.

Instructions: Please respond to all of the items listed below.

1) Were the types of graphs used in the study familiar to you? ☐ Yes ; ☐ No

If not, which ones were unfamiliar?

 ☐;  ☐;  ☐;  ☐;  ☐;  ☐;

2) How often do you analyze data similarly to how you analyzed it in the study (e.g. for school or work)?

1	**2**	**3**	**4**	**5**
Daily	Weekly	Monthly	Yearly	Never

What type of data do you analyze? ____

____

Do you analyze this data alone or together with others? ________________________________

____

3) Is there anything that could have helped you to better solve the tasks given?

____

____

4) How closely did you work with your partner(s) during the study (circle one)?

1	**2**	**3**	**4**	**5**	**6**	**7**
We worked together all the time			We worked both together and indepen-dently half of the time			We worked independently all the time

Comments: ____

5) How closely did you monitor the work of your partner(s) during the study (circle one)?

1	**2**	**3**	**4**	**5**	**6**	**7**
I was fully aware of my partner(s) activities			I was (not) aware of what my partner was doing half of the time			I had no idea what my partner(s) were doing

Comments: ____

6) Did you ever divide up a task?: Yes ☐, No ☐

If so, which questions did you divide up?

- Cereal Scenario: 1 ☐, 2 ☐, 3 ☐, 4 ☐
- Behaviour/Situation Scenario: 1 ☐, 2☐, 3☐, 4☐, 5☐, 6☐

How did you divide up the task?

____

Instructions: Please respond to all of the items listed below.

Why did you divide the task(s)?

____

7) Did the group work effectively as a team on the tasks (circle one)?

1	2	3	4	5	6	7
The group work was very effective			The group was neither effective nor ineffective			The group work was very ineffective

Why (not):____

8) How much do did you contribute to solve the tasks (circle one)?

1	2	3	4	5	6	7
I contributed the most to solving the tasks			I contributed equally to solving the tasks			I contributed the least to solving the tasks

Why (not):____

9) How satisfied are you with your work in the team to solve Scenario 1 (Cereal) (circle one)?

1	2	3	4	5	6	7
Very satisfied			Neither satisfied nor dissatisfied			Very dissatisfied

Why (not):____

10) How satisfied are you with your work in the team to solve Scenario 2 (Situations/Behaviours)?

1	2	3	4	5	6	7
Very satisfied			Neither satisfied nor dissatisfied			Very dissatisfied

Why (not): ____

11) Can you think of anything that could have helped to solve the tasks better as a group?

____

____

12) What is your age? __________________

13) What is your gender? [ ] Male [ ] Female

14) Are you currently a student? ☐ Yes; ☐ No

If yes, what is your major and minor?

____

A.1.3 Task Materials

Cereal Scenario

The following task instructions were given to participants:

Scenario:

As part of an effort to improve the marketing of its breakfast options, a consumer packaged goods company polled 880 people, noting their age, gender, and whether or not they have an active lifestyle (based upon whether they exercise at least twice a week). Each participant then tasted 3 breakfast foods and was asked which one they liked best. The company produced a number of charts of the data that have been given to you for analysis. You received:

- Charts on the demographics of the participants
- Charts on the preferred breakfast options for different types of participants

Tasks:

1. The manager of the company wants to know what the characteristics of the participants were. He does not like numbers. Please give a short and very general description of the characteristics of the participants according to age, gender, and lifestyle. (For example: "There are more x than y", "most people are ...", "few people are...").
2. The company wants to increase advertising for its products to groups that already like a particular breakfast option. Currently they are unsure what kinds of people prefer which breakfast option. Please make a recommendation to the company on who each breakfast option should be advertised to.
3. Make a good estimate if more females prefer oatmeal than active people prefer cereal?
4. Do more inactive people prefer oatmeal than people over 60? Do you think there might be a relationship between the lifestyle and age in terms of preference for oatmeal?

Please come to a consensus about the answers to each task and present the results to us once you have completed the task.

Figure A.1 shows the data material for this task.

Behaviour Scenario

The following task instructions were given to participants:

Scenario:

52 students were asked to rate the appropriateness of 15 behaviours in 15 different situations. Appropriateness was rated on a 10-point scale ranging from 0 = "extremely appropriate" to 9 = "extremely inappropriate". Your role is to analyze their responses.

Specifically you were given:

- Charts that relate the behaviours and their appropriateness in various situations, both as scatter plots and bar charts.
- Two overview charts.

Tasks:

1. Find at least two different pairs of behaviours that have similar ratings in at least three different situations.
2. Choose three situations and describe behaviours most appropriate for that situation according to the graphs.
3. Find two situations that have at least five behaviours with similar ratings.
4. Is it more appropriate to argue or belch in a park?
5. Where did people think it was most appropriate to laugh?
6. What behaviour in which situation was most inappropriate? What behaviour in which situation was most appropriate?

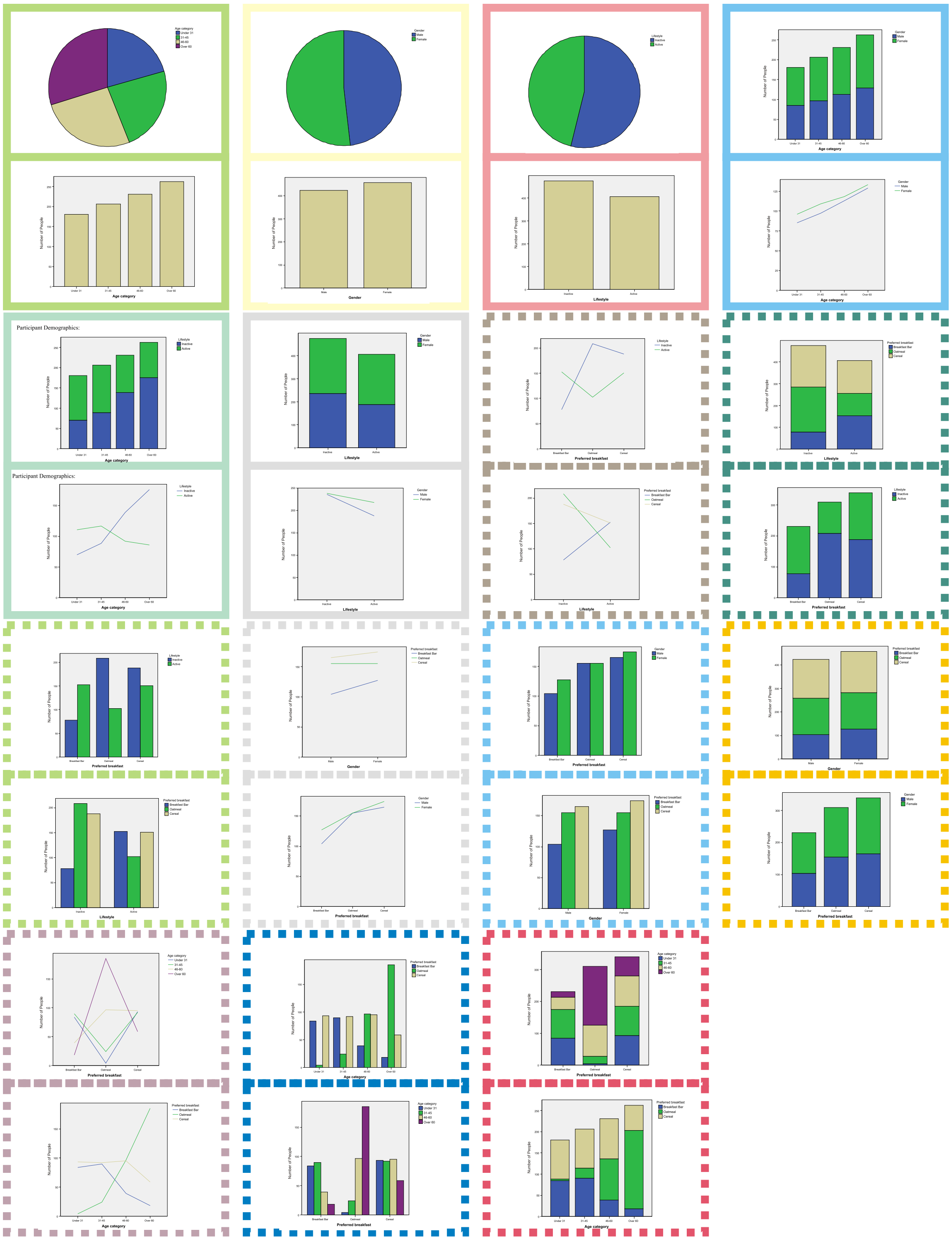


Figure A.1: Study Material for Cereal Scenario.

Please come to a consensus about answers to each task and present those results once you have completed every task.

Figure A.2 shows the data material for this task.

Figure A.2: Study Material for Behaviour Scenario.

	Singles	Pairs	Triples	Average
Parse	3.83	1.93	2.13	2.63
Operate	11.89	8.23	12.93	11.02
Select	3.62	3.43	4.06	3.70
Strategy	0.00	1.70	1.90	1.20
Clarify	0.43	0.51	0.98	0.64
Browse	1.16	0.49	0.46	0.70
Validate	0.04	1.28	3.14	1.49
Collab	0.00	0.09	0.30	0.13
Idle	0.03	0.07	0.25	0.12

Table A.1: Average time spent per analysis process.

A.1.4 Additional Analysis Results

Table A.1 shows the average time (in minutes) spent for both scenarios per identified analysis process in the study.

Figure A.3–A.8 show temporal sequence of processes used in each study group. Each bar stands for the temporal sequence of one participant per study group. The following legend was used to encode the processes:

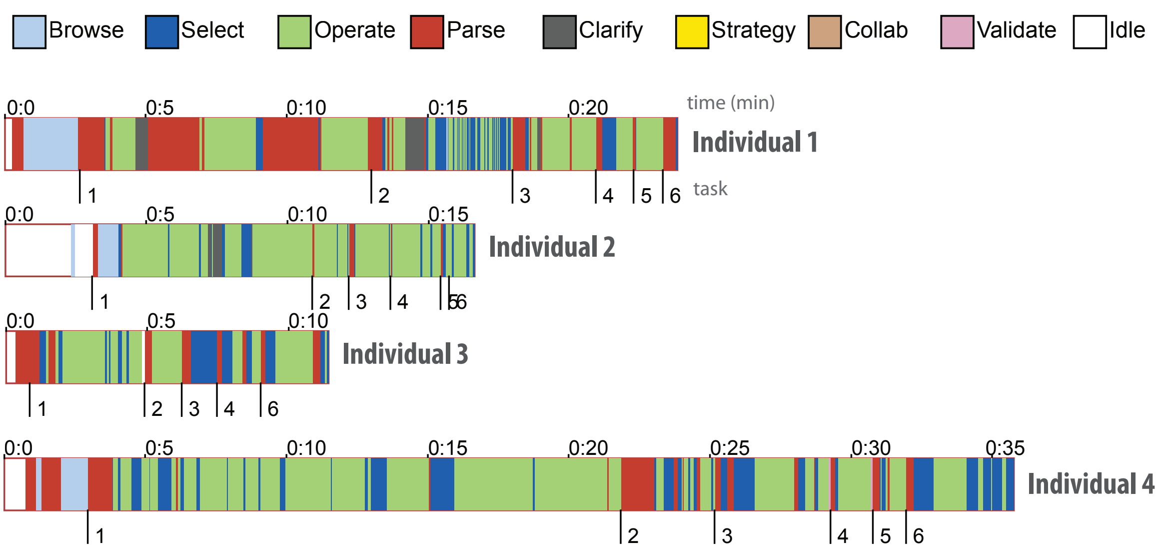


Figure A.3: Process sequences for individuals in the Behaviour Scenario.

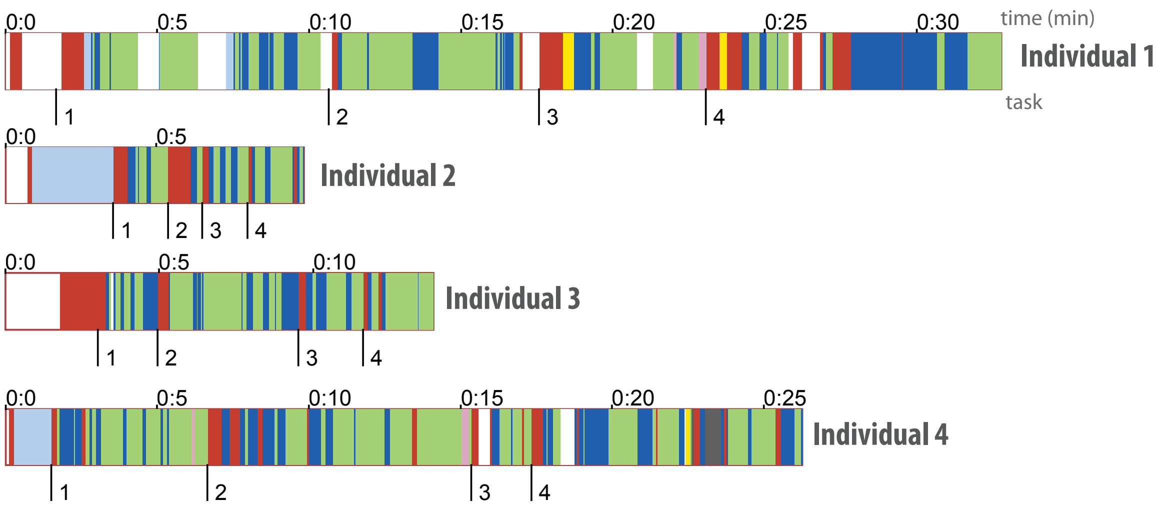


Figure A.4: Process sequences for individuals in the Cereal Scenario.

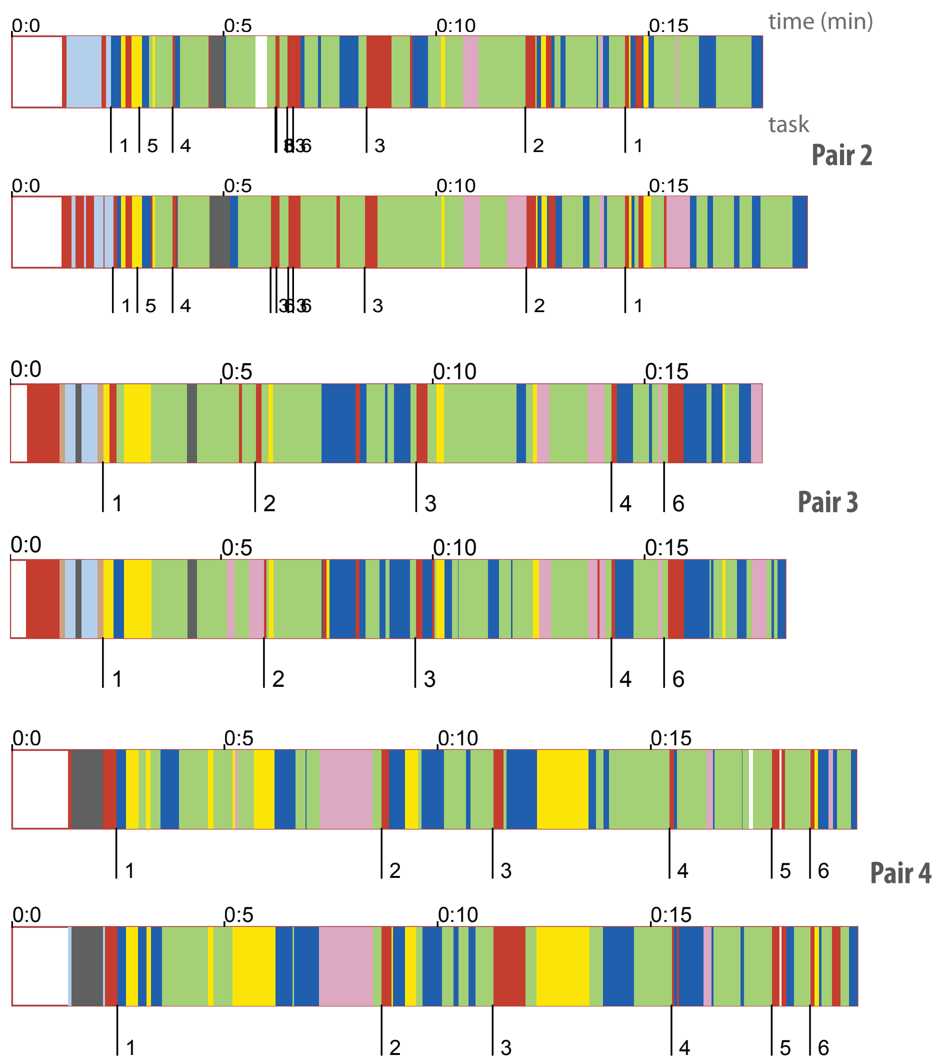


Figure A.5: Process sequences for pairs in the Behaviour Scenario. One pair did not consent to being videotaped and, hence, process sequences could not be collected.

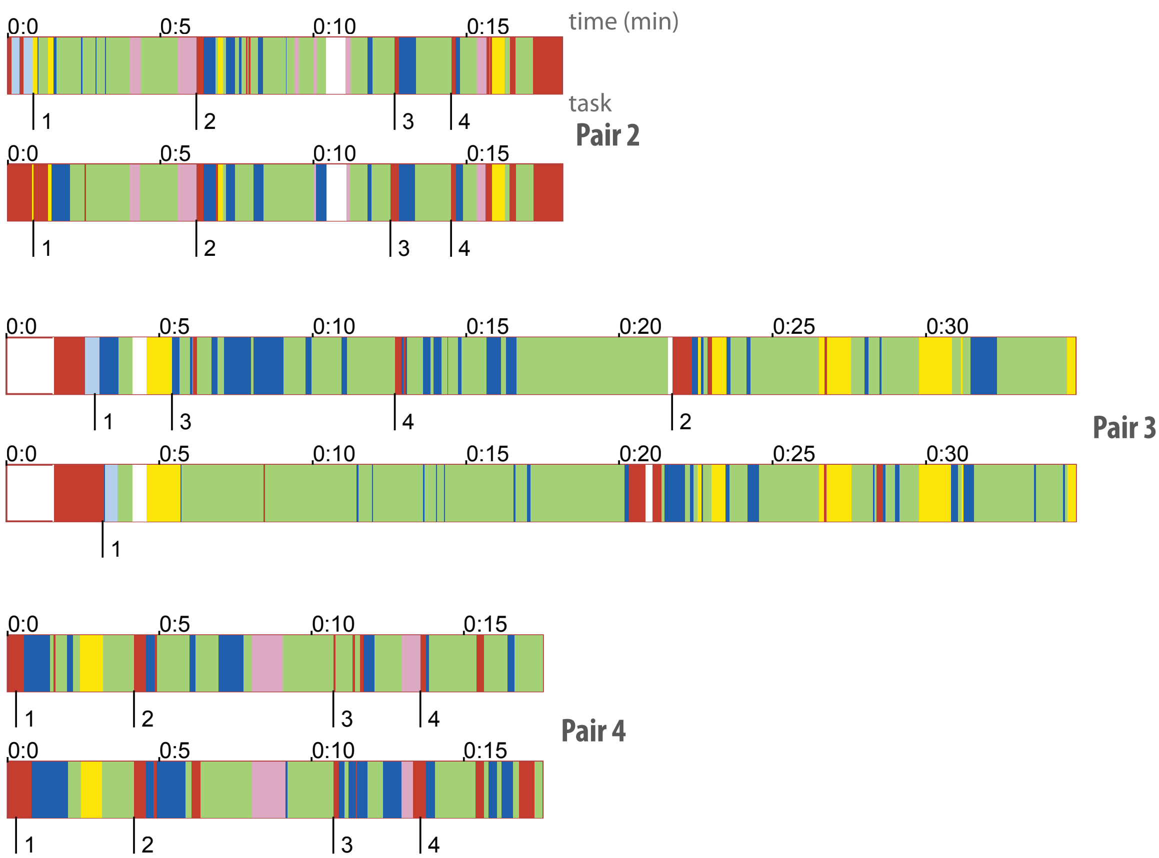


Figure A.6: Process sequences for pairs in the Cereal Scenario. One pair did not consent to being videotaped and, hence, process sequences could not be collected.

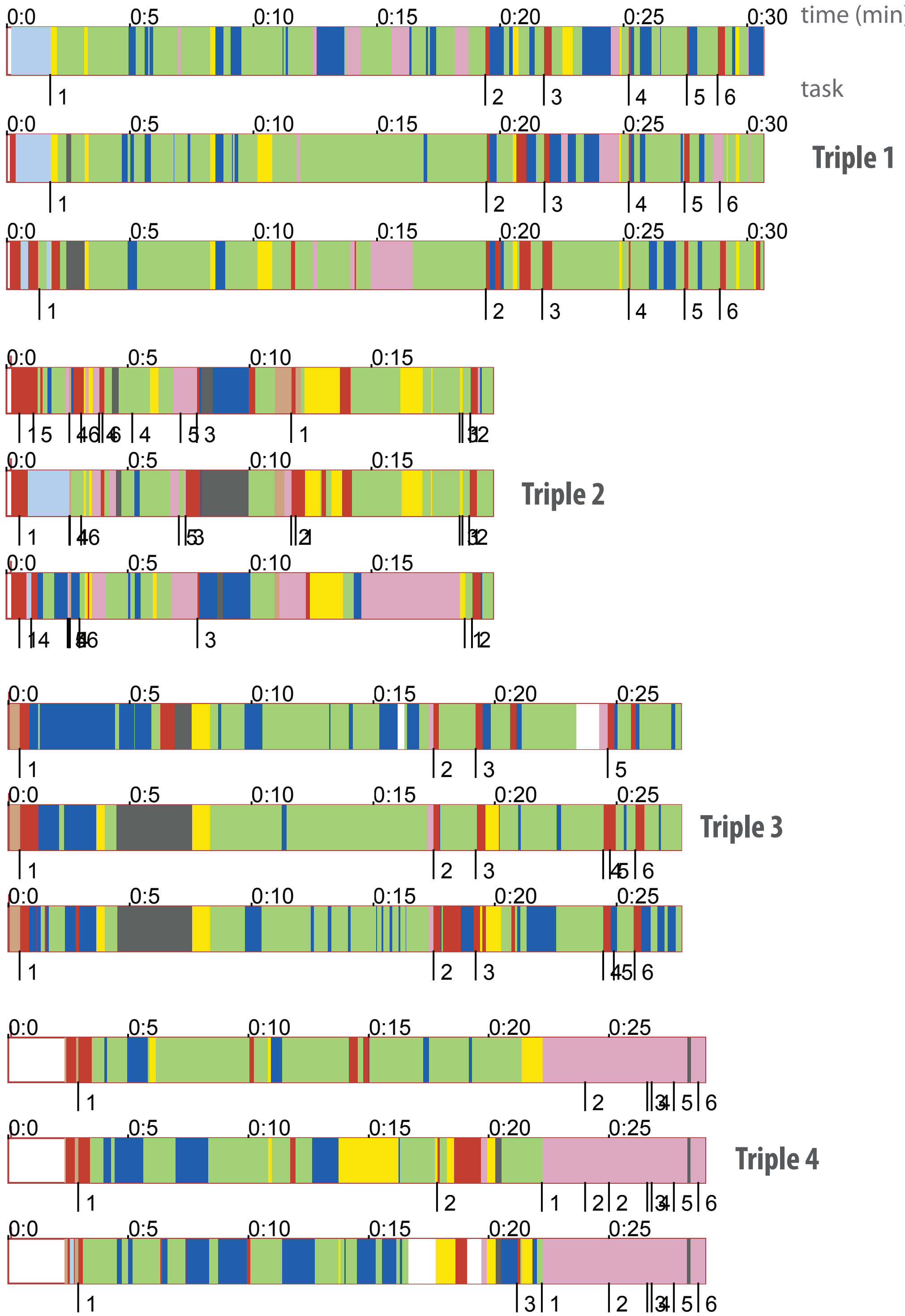


Figure A.7: Process sequences for triples in the Behaviour Scenario.

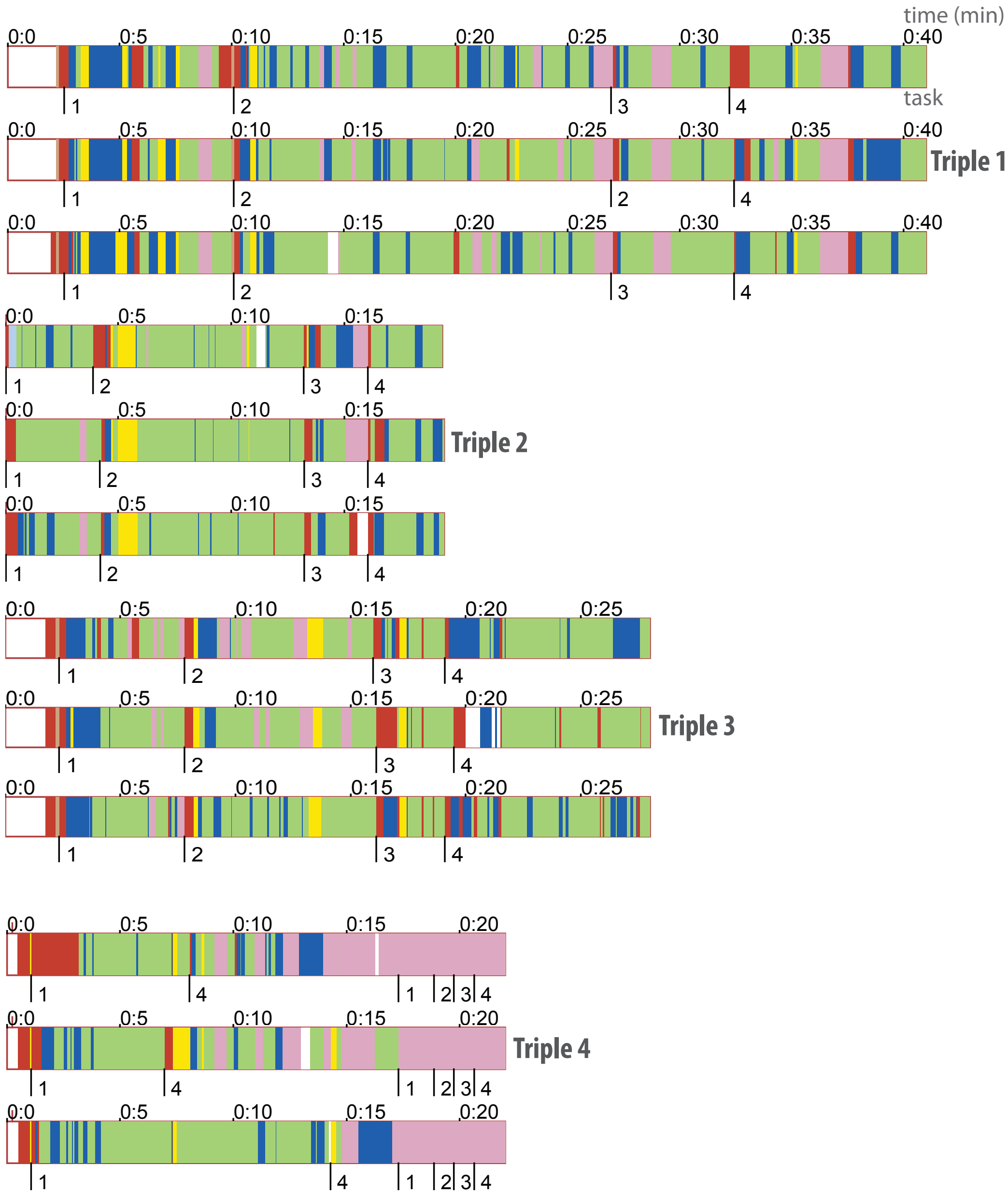


Figure A.8: Process sequences for triples in the Cereal Scenario.

A.2 Chapter 5: Implementation Details

CoTree is implemented using the Large Display Framework as discussed by Isenberg et al. (2006a). It is based on OpenGL and C++ code. In this section, I discuss the tree layouts components added to the framework as part of this project.

In order to improve rendering efficiency in the program, the tree layouts in CoTree are pre-calculated and stored once the first view of the layout is created. When a view of the particular tree layout is created in the interface, the pre-calculated layouts are accessed and mapped to the specific size and aspect ratio of the visualization plane on which the layout is rendered. Once a pre-calculated layout is drawn, it is stored in an OpenGL display list. If the aspect ratio of a visualization plane changes upon resize, the display list is recreated but the layout does not have to be recalculated. The details are discussed next for both the dendrogram and the radial space filling layout. Most of my code was adapted and integrated within a sequence of libraries by Matthew Tobiasz as well as extensions to the Large Display Framework (Innovations in Visualization Laboratory, 2009). The libraries and layouts are discussed in more detail in Tobiasz' MSc thesis (Tobiasz, 2010).

A.2.1 Dendrogram Layout

When pre-calculating the dendrogram layout, the x-coordinate of each node is stored within a 0–1 range, 0 for the leftmost node and 1 for the rightmost node. These x-coordinates can be easily mapped to the respective size of the visualization plane once the OpenGL display list is created. Listing A.1 shows the pseudocode for calculating this layout.

A.2.2 Space-Filling Radial Tree Layout

When pre-calculating the space-filling radial tree layout, I stored two radial coordinates per node. Its beginning angle and its end angle. The inside and outside radius of each node depends on the node's level and can be assigned during the generation of the OpenGL display list when the maximum radius is known. Listing A.2 shows

```
Dendrogram Layout Pseudocode()
{
 //ordered list of leaves
 List<Nodes> leafList = getAllLeafsFromTree();
 //space between leaf nodes
 double leafOffset = 1.0 / (leaftList.Count − 1.0);

 //Next, the node positions are assigned inside a post order tree traversal
 postOrderTreeTraversal(root);

 //Method to calculate the x position for each node.
 //Note that no y position needs to be assigned here.
 //The y position of the node depends on its level or leaf status,
 //it is therefore assigned during display list generation.
 postOrderTreeTraversal(Node n){

  for(int i = 0; n.getChildCount; ++n){
   postOrderTreeTraversal(n.getChildAt(i));
  }

  if(n.isLeaf){
      int leafPosition = getPositionInLeafList(n, leafList);
      n.X = leafPosition ∗ leafOffset;
   }
   else{
      double leftChildX = n.getXPositionOfLeftMostChild();
      double rightChildX = n.getXPositionOfRightMostChild();
      n.X = (rightChildX − leftChildX) ∗ 0.5 + leftChildX;
   }
 }
}
```

Listing A.1: Pre-calculation of the dendrogram layout.

code to calculate the angles for the nodes. In order to draw the cushion effect on the individual nodes, I turned on lighting and assigned a normal to each vertex of the node. Listing A.3 shows how the cushion effect can be achieved by turning the normals slightly away from the upright direction.

```
//Hashtables mapping node IDs to angles
Hashtable<int,double> startAngles = new List<double>();
Hashtable<int,double> endAngles = new List<double>();

CalculateAngles(Node n){
   //go through all nodes in pre-order traversal
   if(node.isRoot()){
      startAngles.insert(n.ID,0.0);
      endAngles.insert(n.ID,360.0);
   }
   else{
      //get the index of the node among its siblings
      int childNr = n.getSiblingIndex();
      int parentDegree = parent.getDegree();
      int parentId = n.getParent().getId();
      double startAngle,endAngle;

      //assign the starting angle
      if(childNr == 0){
         start = beginningAngles[parentId];
      }
      else{
         int nextOldestSiblingId = parent.getChild(childNr - 1).getId();
         start = endAngles[nextOldestSiblingId];
      }

      //assign the ending angle
      end = endAngles[parentId];

      if(childNr != (parentDegree - 1)){

         //calculate the end angle based on a node weight
         //for example, to make all child nodes the same size:
         double angleRange = end - beginningAngles[parentId];
         //simply divide the angle range of the parent by the number of its
             children
         double childAngle = angleRange / parent.getDegree();
         end = start + childAngle;
      }
      beginningAngles.insert(n.getID,start);
      endAngles.insert(n.getID,end);
   }

   for(int i = 0; i < n.getDegree(); ++i){
      calculateAngles(n.getChildAt(i));
   }
}
```

Listing A.2: Pre-calculation of the radial space filling layout (pseudocode).

```
void GLUtilities::fillCircleArcCushionGL(double radiusOuter,
                                         double radiusInner,
                                         double startAngle,
                                         double endAngle,
                                         double stepsize){

    int stepsPerArc = (int)(0.5 + (endAngle - startAngle) / stepsize);
    if (stepsPerArc < 1) stepsPerArc = 1;
    double angle = startAngle;
    double angleIncrement = (endAngle - startAngle) / (double) stepsPerArc;
    double shiftVectorFactor = 1.0;
    double shiftVectorFactorIncrement = -2.0 / stepsPerArc;

    glBegin(GL_TRIANGLE_STRIP);
    for (int i = 0; i <= stepsPerArc; ++i) {
        double sine = sin(angle);
        double cosine = cos(angle);

        // normal vector shift
        double cubicFactor = shiftVectorFactor *
                             shiftVectorFactor * shiftVectorFactor * 1.5;
        double xShift = -sine * cubicFactor;
        double yShift = cosine * cubicFactor;
        // outer normal vector
        Vector v(cosine + xShift, sine + yShift,2.7f);
        v.normalize();

        glNormal3f(v.x,v.y,v.z);
        glVertex2d(cosine * radiusOuter,sine * radiusOuter);

        Vector v_inner(xShift, yShift,1.0f);
        v_inner.normalize();
        glNormal3f(v_inner.x, v_inner.y, v_inner.z);
        glVertex2d(cosine * radiusInner, sine * radiusInner);
        angle += angleIncrement;
        shiftVectorFactor += shiftVectorFactorIncrement;
    }
    glEnd();
}
```

Listing A.3: Nodes drawn with a cushion effect (C++ and OpenGL code).

```java
public class MultiMouseEvent extends MouseEvent {

    private int mouseId = 0;

    public MultiMouseEvent(int mouseId, Component source, int id, long when,
         int modifiers, int x, int y, int xAbs, int yAbs, int clickCount,
        boolean popupTrigger, int button){
       super(source, id, when, modifiers, x, y, xAbs, yAbs, clickCount,
          popupTrigger, button);
       this.mouseId = mouseId;
    }

    public int getMouseId() {
       return mouseId;
    }

    public void setMouseId(int mouseId) {
       this.mouseId = mouseId;
    }
}
```

Listing A.4: A simple extension to a standard MouseEvent class to allow input ids to be captured per mouse.

A.3 Chapter 6: Implementation and Study Details

A.3.1 Implementation Details

This section includes brief sample code for the mouse event modifications made to retrofit NodeTrix. Listing A.4 shows a simple extension of a standard Java MouseEvent-class to allow input IDs to be captured per event. These IDs were used in all classes that reacted to mouse input. The MouseWheelEvent class was similarly retrofit. Listing A.5 and A.6 show two examples of modifications to code required for handling the selection of information items through picking and lasso techniques.

```
MouseReleased(MultiMouseEvent e){
  //identify which mice has fired the event
  Integer mouseId = getMouseId(e);

  //this example is for a left click
  if( (e.getModifiersEx()&InputEvent.BUTTON1_DOWN_MASK) != 0){
    //Find out what item was clicked on
    Item item = getItemPickedOn(e.getX(),e.getY());
    Integer itemId = new Integer(item.getId());

    //Find out if someone else already clicked on it
    //pickedItems is a hashtable mapping mouseIDs to itemIDs
    boolean contained = pickedItems.containsValue(itemId);
    //Either way add the item to the list of things that are clicked on
    pickedItems.put(mouseId, itemId);

    if(contained){
      mouseIdColumn.setExtend(itemId, -1);
      //to get the involved mice
      Vector<Integer> mouseIds = getMiceOnItem(aggregatedItemPicked,
          itemId);
      //check for mouse conflicts if you want
      if(mouseIds.size() > 1){
        conflictOccurred(MultiInputConflict.MATRIX_MULTITOUCH, mouseIds,
            "Matrix",itemId);
      }
    }
}
```

Listing A.5: An example showing how picking was modified to allow for multiple concurrent inputs.

```
LassoReleased(MultiMouseEvent e){
  //on each mouse move and drag, the coordinates of the mouse track are
     stored in a separate datastructure per mouse.
   Polygon lasso = nodetrixVis.getMouseTrack(mouseId);

   //check which objects the lasso encloses
   ArrayList selections = new ArrayList();
   selections = nodetrixVis.pickAll(lasso,selections);

   //since it is possible that someone could be dragging a lasso around a
      node that is currently being picked by someone else we can check for
       a conflict here
   IntArrayList subItems = new IntArrayList();
   for(int i=0; i<selections.size(); i++){

      //for each object in the lasso check if it's being dragged currently
      Item item = (Item)selections.get(i);
      subItems.add(item.getId());
      //is there a conflict?
      if(itemsPicked.containsValue(item.getId())){
         //yes, there is
      //prepare and send a conflict event or do something about it
         conflictOccurred(MultiInputConflict.LASSO_TOUCHEDNODE_INTERSECTION
             ,mouseIds,"Lasso␣around␣picked␣node",item.getId());
      }
   }
}
```

Listing A.6: Supporting multiple mice may introduce problems of multiple concurrent selections. This is one example how our code was modified to deal with multiple and concurrent lasso and node selection.

A.3.2 Materials for the CoCoNutrix Study

Informed Consent Form

The consent form on the following pages was handed to participants before the beginning of the study.

UNIVERSITY OF CALGARY

Department of Computer Science
Consent Form

Title of Investigation: Interfaces for organizing and sharing information
Investigators: Sheelagh Carpendale, Petra Isenberg, Anastasia Bezerianos, Nathalie Henry, Jean-Daniel Fekete

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

Description of Research Project:
We are currently investigating how software can support interactions between people as they organize and share information. To this end, you will be asked to use a tool we call "CoCoNutrix" which is a system for co-located sharing and organization of information shown in a social network. You will be shown different interaction techniques that offer a variety of support for organizing and sharing this information. This may involve such tasks as moving objects to different locations, passing digital objects between collaborators, and organizing digital objects for shared access and analysis. We will be observing and programmatically capturing your actions, as well as videotaping you during the course of the session. This videotaping is optional and you can still participate if you choose not to be videotaped. You will also be asked to complete a pre-session questionnaire to further our investigation. It is estimated that your involvement will take approximately one hour, and you will be remunerated for your time. This research is being conducted collaboratively with two researchers at the University of Calgary, two researchers at Université Paris-Sud, Orsay, France, and one researcher at NICTA, Sydney, Australia (see below). The data analysis will take place in Calgary. There are no known harms associated with your participation in this research. No information that discloses your identity will be released or published without your specific consent to disclosure. All data received from this study will be stored in a locked cabinet and such information that will be stored on a computer will only be accessible through the use of a password. All data will be stored for a period of time no longer than five years. Information will be carefully disposed of (shredding for hard copies and deleting for electronic copies) when this investigation is complete. You will be able to withdraw from this study at any point. If this occurs, any data collected up to that point about you will be discarded. You are also able to refuse to answer whatever questions you prefer to omit.

Informed Consent: Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in this research project and agree to participate as a participant. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal professional responsibilities. You are free to not answer specific items or questions in interviews or on questionnaires. You are free to withdraw from the study at any time without penalty. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, contact:

Sheelagh Carpendale, Department of Computer Science, University of Calgary
Phone: (403) 220-6055, sheelagh@cpsc.ucalgary.ca

Petra Isenberg, Department of Computer Science, University of Calgary
Phone: (403) 210-9499, pneumann@cpsc.ucalgary.ca

Anastasia Bezerianos, National ICT Australia (NICTA), Sydney, Australia
Phone: +61 2 8374 5569, a.bezerianos@nicta.com.au

Nathalie Henry, LRI, Université Paris-Sud, Orsay Cedex
Phone: +33 16915 3486, Nathalie.henry@lri.fr

Jean-Daniel Fekete, INRIA, Université Paris-Sud, Orsay Cedex
Phone: +33 1 69 15 64 94, Jean-Daniel.Fekete@inria.fr

Informed Consent: Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in this research project and agree to participate as a participant. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal professional responsibilities. You are free to not answer specific items or questions in interviews or on questionnaires. You are free to withdraw from the study at any time without penalty. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation.

If you have any concerns about the way you've been treated as a participant, please contact Bonnie Scherrer in the Research Services Office, University of Calgary at (403) 220-3782; email bonnie.scherrer@ucalgary.ca.

I grant permission to be audio taped: *Yes:* ___ *No:* ___

I grant permission to be videotaped: *Yes:* ___ *No:* ___

I grant permission to have anonymized video or still images of me used in a publication: *Yes:* ___ *No:* ___

I grant permission to be quoted anonymously in a publication: *Yes:* ___ *No:* ___

______________________________ ______________________
Participant Date

______________________________ ______________________
Investigator/Witness (optional) Date

A copy of this consent form will be given to you to keep for your records if you request it. This research has the ethical approval of the Department of Computer Science and the University of Calgary.

Questionnaire

The questionnaire on the following pages was handed to participants before completing the study tasks. Most relevant data from this questionnaire is presented in Chapter 6.

Questionnaire *(For experimenter: Participant Code __________)*

Thank you for taking part in our study. Please fill out the following questionnaire as accurately as possible. If you have any questions or concerns do not hesitate to ask the experimenter.

1. Age: _____

2. Sex: M ☐ F ☐

3. Handedness: Left ☐ Right ☐ Ambidextrous ☐

4. How often do you use a computer?

Several hours per day	At least once per day	At least once per week	At least once per month	Almost never
1	2	3	4	5

5. How would you rate your familiarity with using large displays?

Extremely familiar	Very familiar	Familiar	Somewhat familiar	Not familiar at all
1	2	3	4	5

Please elaborate:
(e.g. I use a projector once a week for our meeting and present slides from my laptop, we use a digital whiteboard in our class and I often write on it)

____

____

____

____

6. How would you rate your familiarity (working and social) with other members of your group?

Extremely familiar	Very familiar	Familiar	Somewhat familiar	Not familiar at all
1	2	3	4	5

Please elaborate: *(e.g. I work with Joe every day, I go for lunch with Jill once a week, but I've never met John before today)*

____

____

____

____

7. How often do you work closely with others?

Almost always	At least once daily	At least once a week	At least once a month	Almost never
1	2	3	4	5

8. How do you mostly work with others?

Face to face ☐; over the phone/Skype ☐; via email ☐; using cvs/svn ☐;

Other ________________

Please elaborate: *(e.g. I am a mathematician and I prove theorems together all the time in my office, I call my editor at least once a week to discuss the layout of the magazine, …)*

9. How would you rate your knowledge of the data you are analyzing today?

Expert	Very knowledgable	Knowledgable	Passing knowledge	No knowledge
1	2	3	4	5

Please elaborate: *(e.g. I know a lot of people in the dataset and their connections because…)*

10. How would you rate your experience with node/link diagrams?

Expert	Very knowledgable	Knowledgable	Passing knowledge	No knowledge
1	2	3	4	5

Please elaborate: *(e.g. in university I use node/link graphs to see protein connections, I sometimes see them in articles)*

11. How would you rate your experience with social network representations?

Expert	Very knowledgable	Knowledgable	Passing knowledge	No knowledge
1	2	3	4	5

Please elaborate: *(e.g. I am a social analysts and have done statistical analysis of social networks, I am on Facebook and I have create a "friends wheel" to see my social network)*

Possible Interview Questions

These questions will not be asked of every participant or group. They are intended to encourage participants to explain how they experienced the collaboration with Co-CoNutTrix. Other questions may be asked based on responses from the participants.

Content

1. Can you explain to us your findings and the communities you created?
2. Did you notice anything else about people/nodes besides who they co- authored with?
3. Did you base your decisions on prior knowledge, or just the graph co-authorship information?
4. Can you elaborate on the nature of communities you found (cliques vs crosses)? Where there situations that did not fit these categories and how did you deal with them?
5. Do you feel happy with the matrices/communities or believe you could have done better given more time, or with better knowledge of the domain?
6. Did you learn something from the task?
7. What outside information (not part of the graph) would/did you find useful to have? (ex. resources such as google)

Collaboration

1. Would you have preferred to do the task alone?
2. If not, would you have preferred to have your own personal representation and discuss things with partners?
3. Did you feel someone drove the process more than others?
4. Did you assume different roles? (including assigning someone as they main interactor)

5. Did you have any conflict during the collaboration, both in terms of communication and resources? (then ask for explicit conflict examples)

6. Would you have liked item ownership/locking and if so in which situations?

Awareness, history, etc

1. Did you at any point loose awareness of what other people were doing? Would you have liked more information? (ex. areas where people collaborative, highlighting links leaving actors)

2. Was the undo mechanism sufficient for your task? (or would you have wanted a per-person, or per-item undo)

3. Was it easy to remember what decisions you made, was there any confusion?

4. Did you need to keep a history of your process? (Shameless leading to history, if need history per person, global, etc. Also how to present it: incorporated on layout level using say halos, snapshots, etc).

5. Did you need any other type of annotation? (Annotation of actions -summaries- from the system, todo tags, free annotations) Should the annotations be persistent?

Process

1. Describe your strategy to complete the task. Did you give up on any strategies during the task?

2. If you were asked to do the same thing again would you follow the same strategy? (Both in terms of experience, but also if issues like speed delays arise during trial)

3. Did at some point let go of the mouse and just watched? Why? Did you still want to interact during that time?

4. When you created a community, did you revisit it often? (Would you have liked to mark it as done?)

Display space

1. Were you comfortable with the way you were sitting?
2. Where you compelled to interact with areas of the display right in front of you? (especially edge people)
3. Was there enough room for you to complete your task? If not what would you have preferred to have? (ex. shrinking, fisheyes, flipcharts)

Other feedback

1. What more needs to change between given people one mice and highlighted cursors and actions.

Affinity Diagram Categories

Data from interviews and observations were organized into one of the following categories:

- General group characteristics
- Process of creating communities
- Process of working together
- Criterion for creating groups
- How would they do the task again
- Things learnt from the dataset
- Patterns identified in the data
- Strategies used when people couldn't be placed
- Interesting observation
- Obvious advantages of NodeTrix to support group
- Comments on how changes in the global view affect or could affect group work
- Comments on individual views

- Happy with result
- Happy with level of knowledge
- Would have preferred to do task alone
- Perceived roles in group
- Perceived conflicts by participants
- Observed conflicts
- Comments on undo
- Data revisitation
- Awareness comments
- Comments on collaboration space
- Features requested

A.4 Chapter 7: Materials for the Cambiera Study

The evaluation was conducted as part of an internship at Microsoft Research, Redmond, WA, USA. Ethical guidelines of the host institution applied. Participants were informed of their rights verbally and no written consent form was collected.

A.4.1 Questionnaire

The questionnaire on the following pages was handed to twelve of the fifteen pairs participating in the Cambiera study. Due to a technical error, three pairs did not receive the questionnaire.

DATE: **NAME:**

"I was aware of what my partner was working on"

☐ Strongly Agree ☐ Agree ☐ Undecided ☐ Disagree ☐ Strongly Disagree

"I believe that my partner knew what I was working on"

☐ Strongly Agree ☐ Agree ☐ Undecided ☐ Disagree ☐ Strongly Disagree

How often did your partner have important information that would have helped you, but you didn't find out about it?

☐ Very often ☐ Often ☐ Sometimes ☐ Once in a while ☐ Never

How frequently did you pay attention to your partner's work?

☐ Very frequently ☐ Frequently ☐ Occasionally ☐ Rarely ☐ Never

How often did you work on the same question as your partner?

☐ Very often ☐ Often ☐ Sometimes ☐ Once in a while ☐ Never

How frequently did you look for the same information as your partner?

☐ Very frequently ☐ Frequently ☐ Occasionally ☐ Rarely ☐ Never

How often did you look across the table to see what your partner was doing?

☐ Very often ☐ Often ☐ Sometimes ☐ Once in a while ☐ Never

How frequently did you learn about what your partner was doing from your search results?

☐ Very frequently ☐ Frequently ☐ Occasionally ☐ Rarely ☐ Never

DATE: **NAME:**

Roughly, out of 100% of your time:

How did you learn about what your partner was doing?

They told me	(____ %)
I asked them	(____ %)
We did it together	(____ %)
I looked at their searches	(____ %)
I looked at their documents	(____ %)
I looked at my own search results	(____ %)

I gave my partner information by:

Telling them	(____ %)
Passing them searches	(____ %)
Suggesting searches	(____ %)
Suggesting articles	(____ %)
Passing them or showing them articles	(____ %)

I worked…

Alone, researching my own question	(___ %)
Alone, researching a shared question	(___ %)
Together, with my partner	(___ %)

A.4.2 Initial Coding Categories

Table A.2 shows the coding categories I used for coding the first video. It became evident that it was impractical to follow such a rigorous coding theme and the codes were reduced to the list discussed in **??**.

A.4.3 Group Task Successfulness

Table A.3 gives an overview of how many assists each group required, how many facts they connected and how many critical documents were found. Unfortunately, a requirement of using this specific dataset was not to publish the solutions to the task. Hence, the specific facts and documents that were coded cannot be included here but are available upon request.

A.4.4 Temporal Occurrence of Processes Per Group

Table A.4 shows time in seconds spent per collaboration style by each of the groups as coded in the second coding pass. Table 7.2 reflects these timings.

Code	Description
Reading	Reading a document
Parse	Parsing the task description or external material (like the map) or their own notes
Browse a Search	Browsing all the searches in the workspace or through one particular search by running fingers across
Organize	Organizing the workspace by moving documents around and aligning them
Discuss Doc	Discussing the contents of a particular document
Discuss Search	Discussion the results of a particular search
Share Fact	Sharing of a fact from a doc, requires the other person to send feedback
Combine Results	Explicit sharing by both partners of their individual results in order to combine them into the current set of group findings
Mumble	Mumbling about what one is doing without paying attention whether the other person sends feedback
Pass Doc	Passing a doc to the other person
Pass Search	Passing a search to the other person
Strategy	Discussing a strategy to solve the task (e.g. let's build a network diagram, let's look for x or y)
Data Question	Questions about the data (e.g. what does x mean)
Tool Question	Question about how to use the tool
Hypothesize	Stating a hypothesis about what is going on, this is different from sharing a fact
Keyboard Search	Issuing a search from the keyboard
Doc Search	Issuing a search from a document
Checkpoint	Hitting a checkpoint
Vis Awareness	Showing clear signs of making use of the awareness features built into Cambiera
Lack Awareness	Showing lack of awareness of what other person is doing or what is going on
Note Taking	Taking notes on external paper
Assist Partner	Helping the partner do something (e.g. if they don't know how to open a doc or highlight a word)
Tool Problem	A problem with the tool occurred, participants clearly don't know how to use/do something
Assist Experimenter	The experimenter intervenes and does something
Idle	I don't know what the person is doing or s/he isn't doing anything but watching or listening

Table A.2: Full code set used for the initial coding of one particular session.

Group	Condition	Assists	Facts Connected	Critical Docs Found
1	None	3	11	10
2	Partial	0	11	10
3	Full	3	5	8
4	None	2	8	10
5	Partial	3	5	10
6	Full	2	3	10
7	None	2	7	10
8	Partial	5	4	8
9	Full	3	7	10
10	None	0	10	10
11	Partial	4	10	10
12	Full	0	10	10
13	None	2	7	10
14	Partial	1	9	9
15	Full	0	9	9

Table A.3: For each group this table lists how many assists it received, how many facts were connected, and how many critical documents the group found.

Group	DISC	SPSA	VE	SPDA	SPDA-SI	SPDA-SSP	SPDA-SGP	V	D	DP
1	700	164	549	2183	242	1656	285	0	0	0
2	652	1513	335	1660	283	997	380	0	0	0
3	218	118	406	2084	183	809	1092	0	0	901
4	425	412	672	2213	8	655	1550	0	0	602
5	350	142	298	2010	0	594	1416	0	0	1324
6	594	256	649	2728	305	628	1795	23	0	215
7	323	66	449	2799	0	480	2319	0	22	397
8	220	556	406	3725	85	841	2799	0	0	477
9	209	610	1215	1533	128	557	848	0	0	1250
10	265	551	885	2197	202	854	1141	0	0	76
11	494	504	691	2774	283	876	1615	6	0	537
12	459	603	815	1631	317	449	865	0	0	0
13	331	139	785	3654	414	2211	1029	0	42	706
14	354	279	447	3342	245	1583	1514	0	16	418
15	554	792	541	1184	165	770	249	0	0	16

Table A.4: Temporal occurrence of collaboration styles per group in the study