Universität Konstanz
Department of Computer and Information Science

Master Thesis for the degree

Master of Science (M.Sc.) in Information Engineering

Is web technology ready for NUI? Evaluation of a cross-browser and multi-display system for collaborative web search

by

Johannes Zagermann

(01/775157)

1st Referee: Prof. Dr. Harald Reiterer

2nd Referee: Prof. Dr. Marc H. Scholl

Supervisor: Roman Rädle

Konstanz, March 2nd, 2015

Abstract

Collaborative group tasks such as searching, organizing, or problem solving in general are usually facilitated by shared group spaces. For these spaces, tables are often used because of their physical affordances that are essential for group work activities (e.g., allow deictic references or face-to-face collaboration). Therefore, in CSCW, interactive tabletops have been part of various studies around group work activities. However, the influence of tabletop size on such activities has not yet been researched.

The TwisterSearch system was used as a tool to analyze the effects and influences of tabletop size on group work activities. Groups can work collaboratively on search tasks facilitated by personal devices to enable individual search and reading activities as well as a shared group device that allows to structure and share information in an around-the-table situtation.

In an exploratory study the three different tabletop sizes 10.6″, 27″, and 55″ were used as shared group spaces in a between-subjects design. The focus of the evaluation lay on groups' activities as well as their communication and collaboration behaviors. The results reveal differences in communication and collaboration behaviors, as for example groups working on the midsize tabletop tend to discuss longer and more frequently. Both larger tabletops invite their users to work more collaboratively and thereby structure information on the shared group space more actively. The smallest tabletop lets groups rather evolve different roles. These trends are reflected in the outcome as groups working on both larger tabletops show higher individual and group scores.

Zusammenfassung

Aufgaben wie Suchen, Strukturieren und das Lösen von Problemen sind häufig Kernbestandteile von Gruppenarbeit. Diese Aktivitäten finden oft an Gruppenarbeitstischen statt, die durch ihre physischen Eigenschaften dazu einladen, sich zu unterhalten und dabei die Diskussion durch Mimik und Gestik zu unterstützen. Deshalb sind diese Gruppenaktivitäten und interaktive Tische Bestandteil vieler Studien im Bereich CSCW. Der Einfluss der Größe des interaktiven Tisches auf Gruppenarbeit wurde allerdings noch nicht untersucht.

Mit dem TwisterSearch-System wurde ein Werkzeug entwickelt, mit dem sich Effekte und Einflüsse der Tischgröße auf kollaborative Suchaufgaben untersuchen lassen. Hierbei findet die individuelle Suche und das Lesen von Artikeln auf persönlichen Geräten statt. Ein interaktiver Tisch bildet einen gemeinsamen Arbeitsbereich, der es ermöglicht, Informationen gemeinsam zu strukturieren und sich mit Gruppenpartnern auszutauschen.

In einer explorativen Studie wurden drei unterschiedliche Tischgrößen (10.6″, 27″ und 55″) als gemeinsamer Arbeitsbereich eingesetzt. Der Fokus der Evaluation lag auf den Gruppenaktivitäten und deren Kommunikations- und Kollaborationsverhalten. Die Ergebnisse zeigen, dass beispielsweise Gruppen, die am mittleren Tisch arbeiten dazu neigen länger und öfter zu diskutieren. Die beiden größeren Tische laden die Benutzer eher dazu ein, zusammen zu arbeiten und dabei die Informationen auf dem gemeinsamen Arbeitsbereich zu strukturieren. Der kleinste Tisch hingegen führt eher dazu, dass Gruppenmitglieder verschiedene Rollen einnehmen. Diese Tendenzen spiegeln sich in den Lösungen der Gruppenarbeit wider, da die Gruppen, die an den größeren Tischen gearbeitet haben, als Gruppe und individuell bessere Ergebnisse erzielt haben.

Erklärung

Ich versichere hiermit, dass ich die anliegende Arbeit mit dem Thema:

Is web technology ready for NUI? Evaluation of a cross-browser and multi-display system for collaborative web search

selbständig verfasst und keine anderen Hilfsmittel als die angegebenen benutzt habe. Die Stellen, die anderen Werken dem Wortlaut oder dem Sinne nach entnommen sind, habe ich in jedem einzelnen Falle durch Angaben der Quelle, auch der benutzten Sekundärliteratur, als Entlehnung kenntlich gemacht.

Konstanz, 2. März 2015

Johannes Zagermann

Conventions

Throughout this thesis the following conventions are used:

- The plural “we” will be used throughout this thesis instead of the singular “I”, even when referring to work that was primarily or solely done by the author.
- Links to websites of mentioned products, applications or documents are shown in a footnote at the bottom of the corresponding page.
- References follow the Harvard citation format.

Contents

List of Figures

List of Tables

1 Introduction

Tables play an important role in traditional group work activities. Their physical affordances like their size, height, or their design evolved over decades to fit their particular purposes (Hajizadehgashti 2012). They can serve as a common group workspace to share, coordinate, or structure content. Undeniably, tables invite all members of a group to comfortable sit around them. This leads to several key advantages: Multiple people can work together and beyond that it allows them to seamlessly switch between closely coupled collaboration and loosely coupled parallel work (Isenberg et al. 2010). Further, group members can help each other, see their individual contributions, and have face-to-face communications. Latter of course supports rich gestures, deictic references, and even subtle mimics like facial expressions.

Figure 1: Traditional group work

The procedure of group work and usage of different materials and tools depends strongly on the given topic. There are tasks that cannot be solved solely by the use of pen and paper as shown in Figure 1. Often group work requires consulting additional digital sources to obtain materials from sources like digital libraries or the Internet. However, as soon as technologies are involved into group work processes, benefits of traditional around-the-table collaboration begin to fade. Figure 2 for example shows a typical German high school class doing group work while the task requires students to search the Web for facts. Clearly, group members are separated and constrained to work loosely coupled without benefiting from the aforementioned traditional advantages (Rädle et al. 2013).

Figure 2: Group work supported by desktop computers

At least in German schools, desktop computers are still prevalent when it comes to group work tasks that involve computing technology. Of course, schools start to equip pilot classes with one tablet per student but still, they cannot afford to equip entire classrooms with multiple yet expensive interactive surfaces such as interactive tabletops. Tablets, however, might be able to circumvent the need for interactive tabletops but we do not know yet if they can keep up with the advantages of traditional tables. As tables play an important role in traditional group work activities, so do interactive tabletops in CSCW scenarios (Geyer et al. 2011; Isenberg et al. 2010; Morris et al. 2010; Wallace et al. 2011). To our knowledge only few (Hajizadehgashti 2012; Jakobsen 2014; Ryall et al. 2004) have studied the influence of physical properties of interactive tabletops. However, the effect of interactive tabletop size on group work activities needs yet to be researched. Therefore, a controlled lab experiment is used to study the effects and influences of tabletop display size on group work activities.

2 Theoretical Foundations

This chapter forms the theoretical background of this thesis. As the developed Twister-Search system can be described as a visual information seeking system for groups, it makes use of different models of search and theories of co-located collaboration. Since the developed system does not only consist of a single application running on a personal computer, the user interface cannot be described as a classic WIMP[1] user interface but as a post-WIMP user interface. Post-WIMP user interfaces adapt the interaction with computers to the known interaction with the real world by using "pre-existing knowledge of the everyday, non-digital world." (Jacob et al. 2008). Distributed user interfaces build a subset of these post-WIMP user interfaces. They distribute parts of their interfaces e.g. to different individuals, various devices or places. For these reasons this chapter provides insights into the way people search and work together as well as the characteristics of post-WIMP and distributed user interfaces as they are described in literature.

2.1 Models of Search

The way people search and use computer-supported search systems depend strongly on their purpose. It's a broad field that differs from people who are actively searching local libraries for specific scientific literature to persons who are looking for recipes in the Internet to people who are interested in a certain kind of music or movies. Thus, there are a variety of search engines with different requirements that meet people with various necessities and procedures.

There are several models of search in literature that help to understand different search behaviors. First, we will describe two frameworks developed by Ben Shneiderman (Framework for Mega-Creativity (Shneiderman 1998; Shneiderman 2000) and his four-

---

[1] WIMP is an abbreviation for Windows, Icon, Menus and Pointer.

phase framework for textual search (Shneiderman et al. 1997)). As a second step, we will take a look at the canonical model of social search by Evans and Chi (Evans & Chi 2008; Evans & Chi 2010) as this is a basis of the TwisterSearch system.

2.1.1 Framework for Mega-Creativity

In his framework for mega-creativity (Shneiderman 1998) Shneiderman describes that generators of excellence (so-called genexes) support creativity as well as innovation. They are digital tools that can help people in certain domains to generate ideas and be creative. His concept is built on four foundational believes (FB1 - FB4):

- **New knowledge is built on previous knowledge (FB1):** Creative persons can build new knowledge by their existing understanding of a domain. Knowledge can be combined to gain new knowledge or ideas. Previous knowledge has to be filtered and searched by the help of search services, information visualization or similar (Shneiderman 1998).
- **Powerful tools can support creativity (FB2):** Powerful tools can support innovation by providing templates or facilitating exploratory processes. Rapid prototyping, simulation, combination of ideas or exploration in general help to be creative (Shneiderman 1998).
- **Refinement is a social process (FB3):** Individuals create new ideas that are shared with peers. Their suggestions and consultations can be used to refine ideas (Shneiderman 1998).
- **Creative work is not complete until it is disseminated (FB4):** New knowledge can be shared with others to trigger discussions, create new ideas and gain new knowledge (Shneiderman 1998).

These foundational believes can be found in various sources of information. Kirby Ferguson, a New York-based writer and filmmaker, investigates the elements of creativity in his four-part video series *Everything is a Remix* (Ferguson 2012). His empirical research on creativity focuses on music, movies and media in general as well as the evolution of inventions like today’s computers. Ferguson defines three basic elements of creativity that are linked to Shneiderman’s FB1:

Figure 3: Basic elements of creativity (Ferguson 2012)

Ferguson shares Shneiderman's belief that new knowledge is based on previous knowledge and therefore describes his three basic elements of creativity with the help of various examples:

- **Copy:** Various musicians like Bob Dylan started their career with covering (and thereby copying) songs from other famous artists. Thereby they got new knowledge that they were able to use later in their career. The famous writer Hunter S. Thompson started his writing career by re-typing the novel *The Great Gatsby* from F. Scott Fitzgerald to get a feeling of how to write a good novel (Ferguson 2012).
- **Transform:** Thomas Edison originally did not invent the light bulb. His first patent was about improvements of electric lamps. He then tried 6000 different materials that finally shaped the known light bulb (Ferguson 2012).
- **Combine:** Johann Gutenberg invented the printing press anno 1440 by combining pre-existing knowledge and tools like screw press, ink and paper. Henry Ford was able to produce the first mass market car by combining the idea of assembly lines, interchangeable parts as well as the automobile itself (Ferguson 2012).

Bill Buxton shows the process of sketching as a conversation of a sketch and the mind.

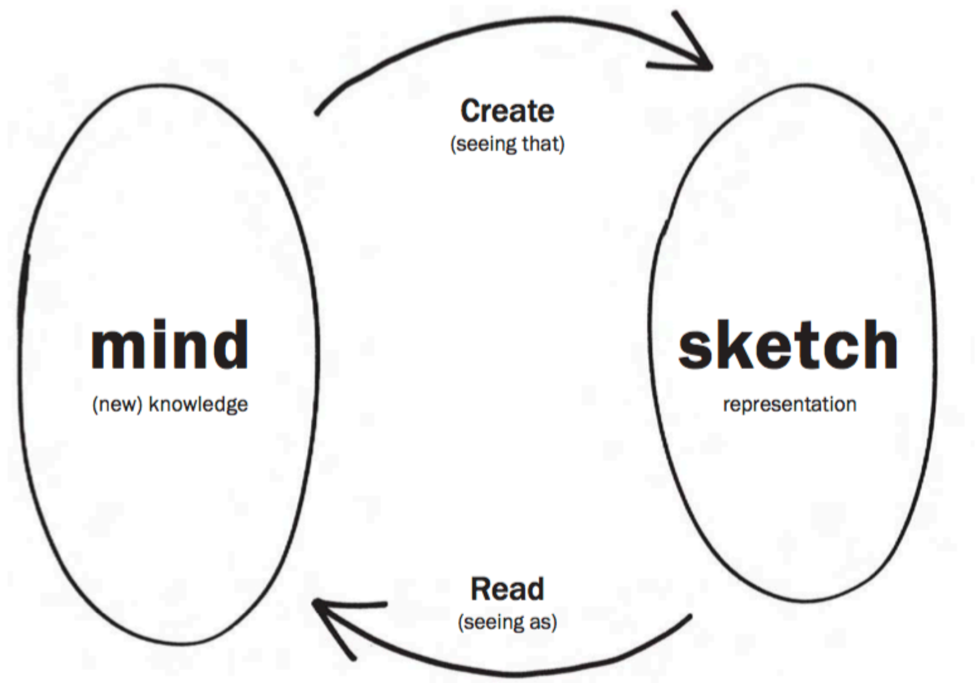


Figure 4: Sketching a a conversation between a sketch and the mind (Buxton 2007)

A sketch can be a quick and informal drawing of an idea, concept or prototype. It helps to understand a design space or a domain. Current knowledge is used to create a new sketch. This external representation of knowledge can be read again by the mind to gain new knowledge (Buxton 2007). This can be connected to Shneiderman's FB1. As Shneiderman proposes in FB2, powerful tools can also support the process of sketching to gain different insights or to explore design possibilities. These sketches can also be used to discuss ideas with peers (compare FB3) to redefine them. Shneiderman describes his foundational believes and the linked phases as non-linear but rather iterative. So does Bill Buxton. Buxton describes prototyping as an iterative incremental refinement that is based on previous knowledge and leads to a final product (see Figure 5 and compare FB1 and FB3).

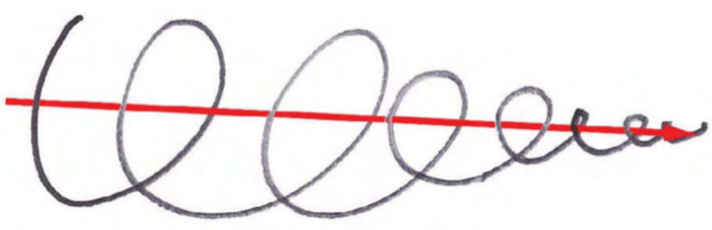


Figure 5: Prototyping as an iterative incremental refinement (Buxton 2007)

Buxton also describes design as an exploration and comparison of different alternatives. They are based on previous ideas and knowledge and each of them can be compared, explored and eliminated with peers to find a solution (see Figure 6 and compare FB1 and FB3).

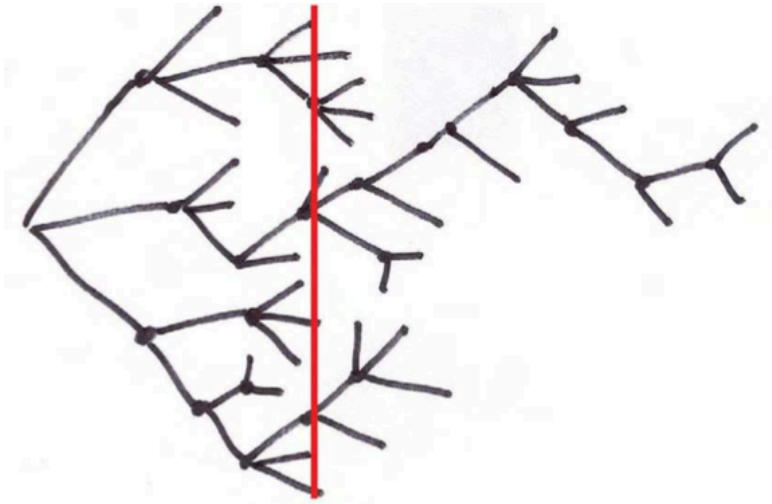


Figure 6: Design as exploration and comparison (Buxton 2007; Greenberg et al. 2012)

Shneiderman describes four different phases of his framework of mega-creativity (Shneiderman 2000) that are linked to his four foundational believes:

- **Collect (FB1):** Learn from previous works by searching and browsing in digital libraries or the Web (Shneiderman 2000).
- **Relate (FB3):** consult with peers in various stages for intellectual and emotional support (Shneiderman 2000).
- **Create (FB2):** explore, compose and evaluate possible solutions (Shneiderman 2000).
- **Donate (FB4):** disseminate results and contribute them to libraries or the Web (Shneiderman 2000).

2.1.2 Four-phase Framework for Textual Search

Shneiderman additionally describes another framework. His four-phase framework for textual search (Shneiderman et al. 1997) shows how to prevent confusing interfaces that lead to results with too many results or without results. This framework is directed to designers to create good and consistent search interfaces. The four phases are:

- **Formulation:** This is the expression of the actual search. Users have to know where to search, what to search and how to search (Shneiderman et al. 1997).

- **Action:** This is the start of the search. It can either be explicitly be pressing a search button or implicitly as the system could start to search while the user is still typing (Shneiderman et al. 1997).
- **Review of results:** In this phase, the user can read and browse the search results. Shneiderman talks about traditional lists and emphasizes the use of visualization techniques helping to find relevant results (Shneiderman et al. 1997).
- **Refinement:** The system supports the refinement of search queries, provides a search history or enables to share and save searches (Shneiderman et al. 1997).

2.1.3 Canonical Model of Social Search

Shneiderman shows different models that help to understand how individuals search. Evans and Chi define a canonical model of social search (Evans & Chi 2008; Evans & Chi 2010) that shares Shneiderman's view on search behavior but goes into more detail in terms of search phases and they consider social aspects of group search activities, too. Shneiderman's search phases concentrate on the duration of the actual search, whereas the canonical model of social search by Evans and Chi divides the search process into three different main phases as shown in Figure 7. In their model, the search starts before the actual search activity and ends after it. Therefore, they present their three phases: *Before Search*, *During Search* and *After Search*.

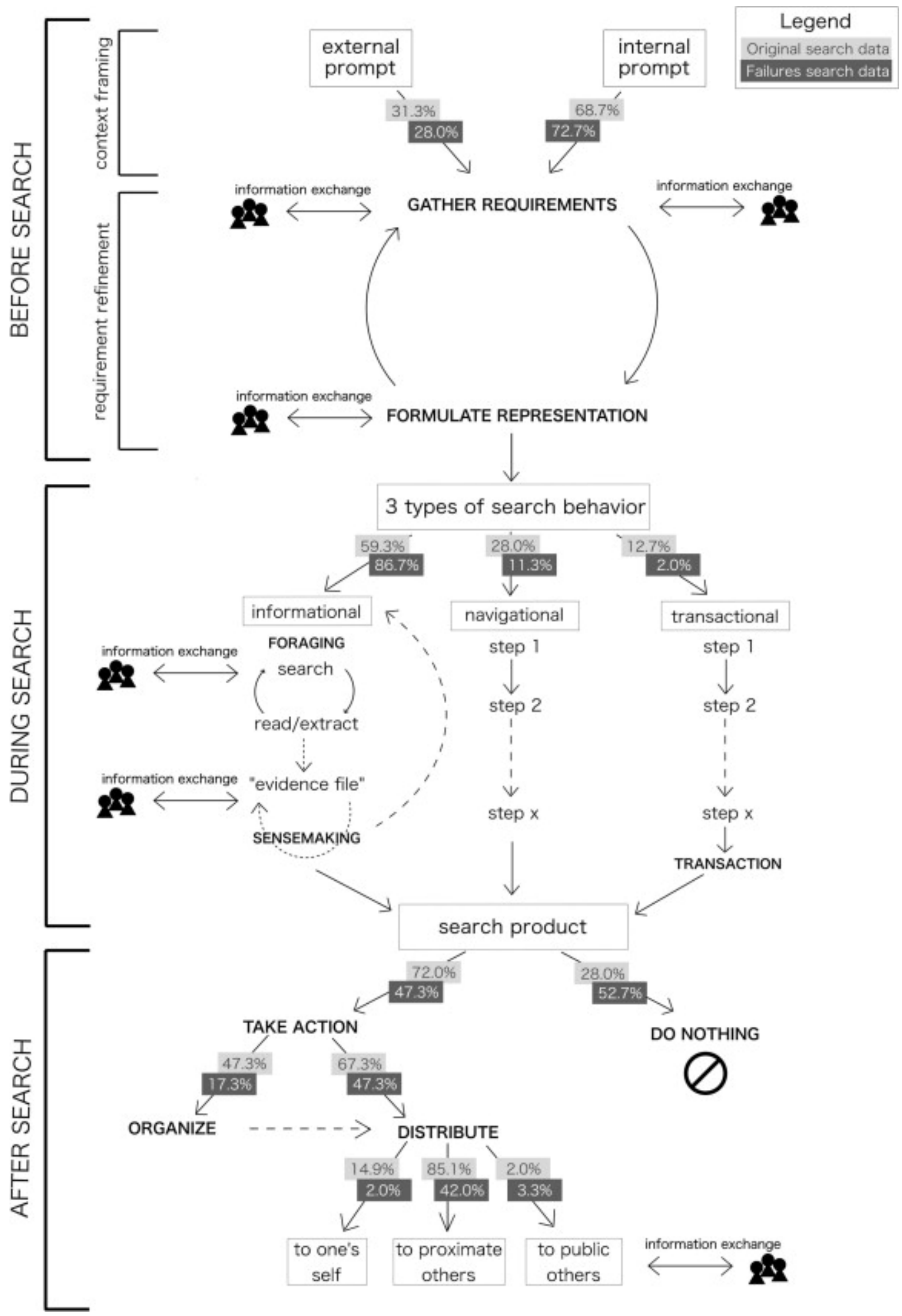


Figure 7: Canonical model of social search (Evans & Chi 2010; Evans & Chi 2008)

Before Search

Evans and Chi describe that the first phase of their social search model starts with a *context framing*. Context framing helps to define motives and information needs (comparable to Schön's framing of problems (Schön 1983)). They constitute what users want to know and how the search is approached. This context framing can be triggered from external requests such as from a boss, teachers or colleagues. This implies a later feedback or answers after the search activities ended. The self-initiated search however can be seen as a background to find information related to personal or work tasks. The gathered requirements from this first step can be discussed with other persons. The *Before Search* phase also includes the step of *requirement refinement*. By exchanging with others the information needs can be further refined by talking about other's knowledge concerning the given topic and to develop concepts, schemas or keywords to formulate search activities (Evans & Chi 2008).

During Search

The *During Search* phase describes activities that happen while users engage in traditional information seeking activities. Evans and Chi divide this into three different types of searches: *Transactional Search, Navigational Search* and *Informational Search:*

- **Transactional Search:** This type of search is characterized by no social interaction during the actual search but by pre-search social interaction. It can be seen as a routinely used search behavior where users want to know driving directions or weather information in well-known environments (Evans & Chi 2008).
- **Navigational Search:** This type of search shares the characteristics of social interaction with the *Transactional Search*. During this phase, users navigate trough familiar locations until the desired information is found. This is often recognized directly and thus the search stops immediately. The users often know where to find the specific information (Evans & Chi 2008).
- **Informational Search:** This type of search is characterized by an exploratory nature, combining foraging, sensemaking and social interactions in an iterative manner. Users may not be familiar with their desired information. The process of foraging is based on skimming, reading and extracting information from first search results. Discussions with peers help to reformulate search terms to refine their search iteratively. Foraging and sensemaking loops are tightly coupled. The act of foraging helps to find preliminary files that help to reflect and modify their

search. Whereas sensemaking and discussions help to find so-called evidence files that can be used for later activities (Evans & Chi 2008).

After Search

Evans and Chi's search model also describes activities that happen after the actual search task. The result of the *During Search* phase is often a search product that can be e.g. a presentation, documents or diverse information for instructions. This search product is often shared and distributed. Individuals save documents for later usage, share them with peers to get feedback or they send their search product to other persons involved in the different search phases. Additionally, the found documents and materials can be organized and saved for later usage. This can also be seen as a sensemaking activity (Evans & Chi 2008).

2.2 Co-located Collaboration

The previous chapter has shown that Evans and Chi as well as Shneiderman describe search as an activity that involves the social interaction with other individuals to a certain extent.

Multiple people can search for information together, share or edit it. This collaboration can have different manifestations. Johansen classifies various types of collaboration in a groupware time – space matrix (Johansen 1988) as shown in Table 1.

	Same Time	**Different Times**
Same Place	Face-to-face interaction	Asynchronous interaction
Different Places	Synchronous distributed interaction	Asynchronous distributed interaction

Table 1: Time – Space Matrix (Johansen 1988)

Table 1 shows the different types of collaboration. The different dimensions time and space can be combined in their various specifications to enable synchronous or asynchronous as well as co-located or distributed collaboration.

Multiple individuals working at the same place and same time collaborate co-located. Thereby the collaboration can be a mixture between individual and group-based activities. This is known as *mixed-focus collaboration* and individuals can change the way they work independently and often (Tang et al. 2006).

The different types of collaboration like individual or social interaction are known as coupling styles. Coupling styles are used to describe the collaboration and behavior of group members. Coupling thereby can be divided into closely and loosely coupled collaboration. Closely coupled interaction thereby involves more social interaction with other individuals than loosely coupled interaction (Tang et al. 2006).

Tang et al. found six different coupling styles for pairs in a collaborative tabletop-based task. They are:

- **Same problem same area:** Collaborators are actively working together
- **View engaged:** The pair is working together, but only one is actively manipulating information
- **Same problem, different area:** Collaborators are working on the same problem, but they focus on different areas
- **View:** One individual is working, the other one is watching
- **Disengaged:** One collaborator is disengaged and pays no attention to the task and the partner
- **Different problems:** Collaborators are working on different and unrelated problems

These coupling styles imply that collaborative system have to feature aspects concerning individual and collaborative work practices. The six named coupling styles are taken up in a subchapter of Chapter 3 Related Work.

2.3 Post-WIMP User Interfaces

Post-WIMP user interfaces show new approaches and can eliminate traditional input devices as mouse or keyboard to some extent by new interaction modalities. These modalities involve all senses in parallel by e.g. touch input and enable natural language communication for speech recognition and allow multi-user environments. Van Dam describes post-WIMP user interfaces as systems that contain "at least one interaction technique not dependent on classical 2D widgets such as menus and icons" (van Dam 1997).

Several frameworks exists that help to design and describe post-WIMP user interfaces that allow intuitive interaction as well as user-friendly behaviors of systems. Two of them are shown in the following sub-chapters: *Reality-based Interaction* and *Blended Interaction*.

2.3.1 Reality-based Interaction

Jacob et al. use their framework of *Reality-based Interaction* to describe interaction concepts regarding multiple themes of reality. Their goal is to "attempt to make computer interaction more like interacting with the real, non-digital world" (Jacob et al. 2008). The four themes of reality are *Naïve Physics, Body Awareness & Skills, Environment Awareness & Skills* and *Social Awareness & Skills.* They are shown in Figure 8.

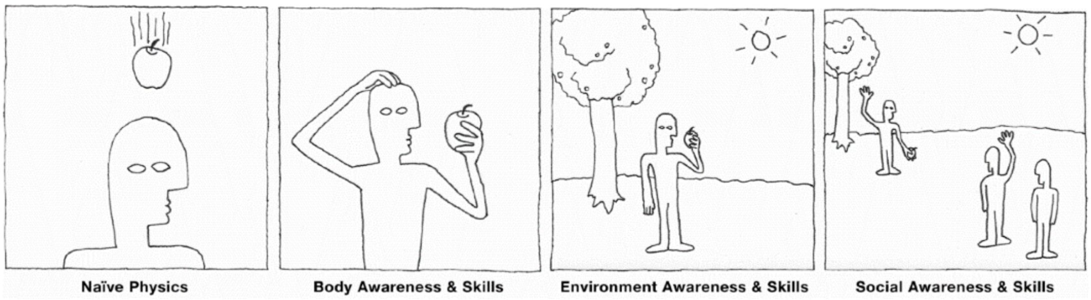


Figure 8: Themes of reality (Jacob et al. 2008)

- **Naïve Physics:** This theme describes the "common sense knowledge about the physical world" (Jacob et al. 2008). It includes concepts like gravity, velocity, movement or scaling that individuals are used to from interacting in the real world. Jacob et al. name tangible user interfaces (TUIs) as an example – using real-world objects for manipulating interactive interfaces
- **Body Awareness & Skills:** This theme builds upon the developed skills of humans and their understanding of their own body to coordinate movements of their body parts in order to achieve a desired action. They can be used e.g. in virtual reality environments.
- **Environment Awareness & Skills:** Individuals are aware of their physical presence in a spatial environment using their sense of orientation and spatial understanding to navigate in or to change an environment. They can be found in the field of virtual reality environments or augmented reality environments.
- **Social Awareness & Skills:** This theme addresses social awareness and skills concerning social interaction. This includes verbal and non-verbal communication or the ability to collaborate with others.

Jacob et al. suggest to base all interaction on these themes as they are based on pre-existing real world knowledge and skills that can be re-used. This reduces the mental demand to use interactive systems as the required skills are already developed by individuals. In addition to their themes of reality, Jacob et al. name the qualities of virtual environments:

- **Expressive Power:** The functionalities of a system / the variety of tasks users can perform.
- **Efficiency:** “users can perform a task rapidly” (Jacob et al. 2008).
- **Versatility:** systems can offer various tasks from different application domains.
- **Ergonomics:** “users can perform a task without physical injury or fatigue” (Jacob et al. 2008).
- **Accessibility:** handicapped persons are able to perform tasks with the system.
- **Practicality:** “the system is practical to develop and produce” (Jacob et al. 2008).

The real and the virtual world both have their specific qualities. Therefore, Jacob et al. suppose to use both qualities in a power vs. reality trade-off. Their goal is “to give up reality only explicitly and only in return for desired qualities” (Jacob et al. 2008). These means that systems should be designed in favor of real-world qualities as described as themes of reality by providing powerful computational power where useful.

2.3.2 Blended Interaction

Blended Interaction is a conceptual framework that helps to design, analyze and explain post-WIMP user interfaces. The goal is to combine aspects of the familiar real world with all its physical and social characteristics with their digital counterparts to achieve a “natural” way of interaction with digital environments. *Blends* build the foundation of this idea. The concept of blends is based on *conceptual integration* or *conceptual blending*, which is a theory in the field of cognitive science defined by Fauconnier and Turner (Fauconnier & Turner 2002). Fauconnier and Turner show that the conceptual system exists of basic-level concepts internalized by our mind. Those concepts can be combined to new and more complex concepts by using two different existing basic-level concepts as input variables. The output of this process is a *blend* that “has an emergent new structure that is not available from the inputs.” (Jetter et al. 2013). This idea can be compared with the ideas of Shneiderman’s *Framework for Mega-Creativity* and his foundational belief that new knowledge is created by existing knowledge (Shneiderman 1998).

The concept of *Blended Interaction* transforms this *conceptual blending* from the field of cognitive science to human-computer interaction. Thereby it is possible to explain how to blend the digital world with natural user skills and knowledge. These blends are a design trade-off between those two counterparts and help designers of post-WIMP user interfaces to find advantages of both of them and combine them to create easy-to-use system functionalities. Figure 9 shows a visual overview of the *Blended Interaction* framework.

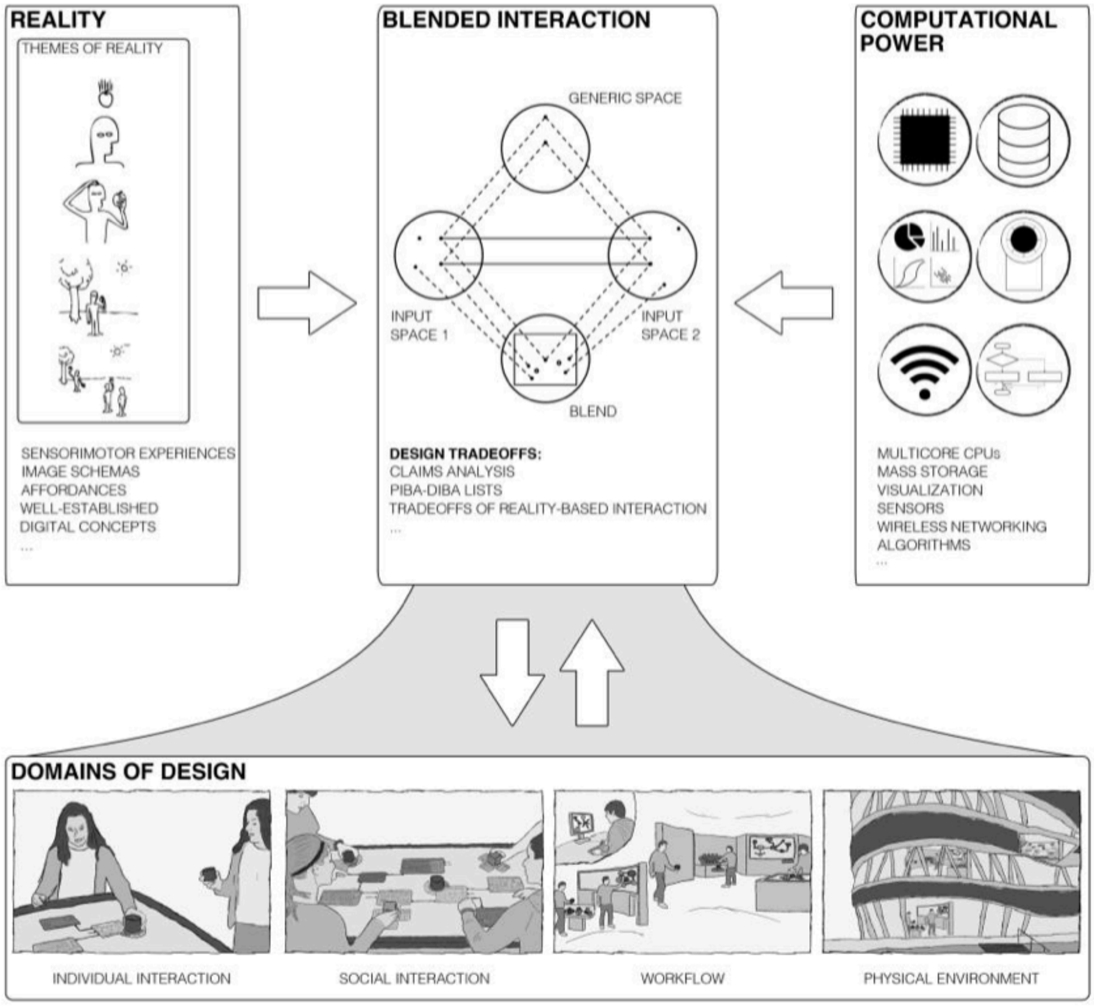


Figure 9: A visual overview of the framework Blended Interaction (Jetter et al. 2013)

As described, the framework *Blended Interaction* uses the concept of conceptual blending with the idea of blends as a central point. Blends need two different input spaces. Therefore, the framework uses aspects of reality shown on the left hand side and aspects of computational power shown on the right hand side of the figure.

While Jacob et al. (Jacob et al. 2008) describe their themes of reality as their natural counterpart to the digital world, Jetter et al. go beyond this idea of reality and add e.g. well-established digital concepts. These concepts are connected to the digital world but as smartphones or tablets are used on a daily basis and their usage is part of our everyday life, they have to be added to our understanding of reality (Jetter et al. 2013). On the digital side, the authors name the expressive power of digital computation. The both input sides are used to create blends that can be applied to four different domains of design: *Individual Interaction, Social Interaction, Workflow* and *Physical Environment.*

- **Individual Interaction:** Collaborative activities assume that all individuals have possibilities to interact with artifacts or parts of the interactive system. Therefore, *Individual Interaction* is the base for collaboration.
- **Social Interaction:** Communication between individuals helps to coordinate collaboration and activities of all individuals.
- **Workflow:** The single activities of individuals and sequences of social interactions are often guided by a higher level workflow with different phases that can be supported.
- **Physical Environment:** The named *individual* and *social interaction* as well as the *workflow* take place in a physical environment including furniture or form factors of the room or digital devices.

The created blends are applied to the named domains of design. In addition to this, the domains of design can even help to redefine different blends. Thus, the connection between blends and their applied domains can be described as a two-way conversation.

2.4 Distributed User Interfaces

Mark Weiser's vision of the computer for the 21st century (Weiser 1991) describes the usage of numerous computers with different specific purposes in various sizes like tabs, pads or boards. They all are interconnected and so parts of their user interface can be distributed on all of them to fit their purpose. Their distribution can even go beyond the realm of devices that Weiser describes as *ubiquitous* and *indistinguishable* from the real world and spread their user interfaces across multiple users, platforms, different spaces or points in time (Weiser 1991).

Today, Mark Weiser's vision of ubiquitous computing becomes more and more concrete. We use different computational devices in various sizes in our every day life. Smart watches check our heartbeat while going for a run and we can analyze the route, the burned calories and the distance later by ourselves or at the same time e.g. by a doctor on smartphones or laptops. We use interactive pens and enhanced paper technologies like provided by Anoto[2]. Commercial solutions like Evernote[3] enable us to round up creative ideas via physical notebooks or our personal devices at the same time.

---

[2] Anoto: www.anoto.com

[3] Evernote: www.evernote.com

The more computational devices disappear and are distributed the more they have to be interconnected to provide seamless interaction. Research in this field can be captured by the term of *distributed user interfaces* (Elmqvist 2011).

Elmqvist defines distributed user interfaces:

"A distributed user interface is a user interface whose components are ***distributed*** *across one or more of the dimensions input, output, platform, space, and time."* (Elmqvist 2011)

The different dimensions in detail:

- **Input:** Single computational device vs. several devices
- **Output:** Single display vs. several displays / devices
- **Platform:** Single computing platform vs. different platforms
- **Space:** co-located vs. remote
- **Time:** synchronously vs. asynchronously

Fisher et al. (Fisher et al. 2014) define eight challenges distributed user interfaces (DUIs) have to face:

- **Consistency:** "Each software instance running on a device involved in a DUI application must maintain a shared and consistent state using the network." (Fisher et al. 2014)
- **Synchronization:** "Synchronizing the actions of software components on different hosts is a core concern for general distributed systems." (Fisher et al. 2014)
- **Heterogeneous hardware:** "DUI applications must run on multiple different hardware platforms, yet leverage the unique capabilities of each platform and device." (Fisher et al. 2014)
- **Volatile device ecosystem:** "DUI applications must be robust against devices joining or leaving the shared environment at any point." (Fisher et al. 2014)
- **Limited resources:** "The storage, computational, or physical resources available for each device in a DUI application may be different, causing that resource to be limited to one or a few devices." (Fisher et al. 2014)
- **Data transfer:** "DUI applications need support for transferring *binary large objects* (blobs) – such as media, images, files, databases, documents, etc. – between devices." (Fisher et al. 2014)
- **Physical space:** "Physical space must be managed on a global level for a DUI application so that individual participating devices can make autonomous decisions." (Fisher et al. 2014)

- **Asymmetric functionality:** “Distributed applications often contain unique components that are asymmetrically divided between participating devices.” (Fisher et al. 2014)

3 Related Work

Chapter 2 has shown the theoretical background with fundamental theories of how we search and work individually as well as collaboratively in groups. Additionally, two frameworks are presented that describe and help to design novel user interfaces that can support these theories. The idea of distributed user interfaces builds a subset to them. In this chapter, related research works are presented that cover different aspects of this thesis. Different subchapters show works about collaborative group work and search systems and research regarding the influences of tables for group activities.

3.1 Collaborative Group Search

This chapter describes different research projects concerning collaborative group tasks and systems that support collaborative search.

3.1.1 WeSearch

WeSearch is a system that allows co-located collaborative web search activities like search itself, browsing and sensemaking for groups up to four individuals gathered around a multi-touch tabletop display. The system provides color-coded toolbars with virtual keyboards for each user that enable them to enter queries or entire URLS. Individual browser elements are used to display search results and parts of them can be stored via clips for later usage (Morris et al. 2010).

WeSearch is based on several design goals (Morris et al. 2010):

- Support **awareness** among group members
- Support **division of labor** among group members
- Enable the **shared search** to persist beyond a single session
- Support **sensemaking** as an integral part of the collaborative search process

- Provide facilities for **reducing the frequency of virtual-keyboard** text entry
- **Reduce clutter** on the shared display
- Address the **orientation challenges** posed by text-heavy tabletop applications.

The WeSearch system has several main features that enable collaborative search activities:

- **Toolbars:** Each user is assigned to an own color-coded toolbar that is placed on each edge of the tabletop display. They can be moved freely. All toolbars feature their own individual virtual keyboard that can be used to enter query terms or URLS. The desired URL or a search engine page is opened in an individual browser view.
- **Browser:** Each browser view can be moved, rotated and scaled using touch manipulations. Additionally, three buttons allow further usage like panning or scrolling search result lists, following links on websites or clipping content by holding the certain button while interacting.
- **Clips:** Clips represent snippets or parts of websites. They can be interacted in the same way like browser elements but they add the opportunity to be tagged by multiple users for later sensemaking activities.
- **Containers:** Containers are used to organize clips via lists, grids and free-form positioning. Clips can easily be added or removed via drag-and-drop gestures and users can name containers by virtual keyboards.

Figure 10 shows a session of WeSearch with four users, their toolbars and multiple browsers, clips and containers.

Figure 10: WeSearch (Morris et al. 2010)

3.1.2 CoSearch

Amershi and Morris (Amershi & Morris 2008) present results from interviews with teachers, librarians and developing world researchers concerning the aspects of collaborative search habits in shared-computer scenarios. They choose those interview partners as schools, libraries and public facilities in developing regions often share the nature of a high student-to-computer ratio, which results in multiple individuals gathering around single desktop computers. Their formative study shows several limitations of that practice:

- **Difficulties Contributing:** As the nature of desktop computers implies, there can only be one individual controlling input devices like mouse or keyboard. This is known as the classical driver-observer problem, where drivers control input devices and observers can make suggestions verbally or trough gestures. Controlling drivers may thereby ignore suggestions and ideas by observers as well as demanding observers may make it difficult for drivers to make contributions.
- **Lack of Awareness:** Group members can show a more dominant behavior, which leads to reduced awareness of other's skills and ideas.
- **Lack of Hands-On Learning:** The driver-observer problem leads to situations where group members have no chance to interact with search technologies or technologies in general.
- **Pacing Problems:** As each group member can have an own way of reading documents, there can be pacing problems as drivers scroll too fast or too slow or navigate to new documents too quickly or too slowly.
- **Referential Difficulties:** Multiple individuals sitting around a desktop computer can lead to referential difficulties as group members situated further from the display cannot read documents or point to interesting parts of documents.
- **Single-Track Strategies**: As desktop computers only allow one active user, there are no possibilities to divide tasks into subtasks.
- **Information Loss:** The nature of public schools and libraries imply that it's mostly not possible to keep information persistent.

Based on these findings they developed *CoSearch*, a system that supports explicitly groups of co-located individuals searching the Web while sharing a single computer. To do so, they provide cheap enhancements to existing desktop computers by connecting one mouse per user. Thus, multiple cursors with different user-specific colors are used to enable each group members e.g. to point to interesting aspects or select and open documents of search results. Search queries are entered by a shared keyboard. By clicking on a hit, a new tab with the document in the specific user-color is opened in the background. By this, each group member can still read the currently opened document but the new page is queued for later reading. Group members can take notes for each search document and send results of the search to their email addresses.

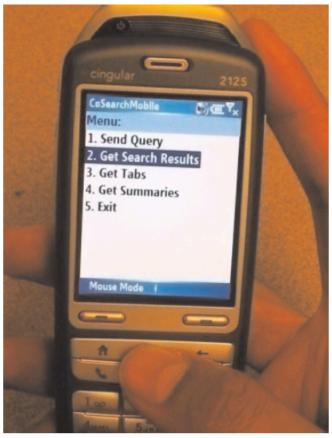


Figure 11: CoSearchMobile (Amershi & Morris 2008)

In addition to their so-called *CoSearchPC* system with user-specific cursors controlled by multiple mice, they provide the opportunity to connect to the CoSearchPC via mobile phones running *CoSearchMobile* (see Figure 11). Mobile phones can be used to either control the user-specific cursor on the CoSearchPC system, to trigger searches or to read search results, documents or summaries on personal devices. That helps to avoid driver-observer problems and each individual can contribute equally to the task while having the opportunity to read documents on personal devices.

3.1.3 SearchTogether & CoSense

CoSense is a system that supports sensemaking for collaborative search tasks using *SearchTogether*. Thus, this chapter first describes aspects of SearchTogether and proceeds with describing Cosense.

SearchTogether (Morris & Horvitz 2007) is a tool that enables groups of remote users to synchronously or asynchronously search information on topics collaboratively. Morris and Horvitz emphasize the ability to divide labor, provide persistence and support individual awareness. The SearchTogether system is used by individuals on their personal computer. Users can log in into the system, create a new search topic and invite friends or peers. Each invited individual can see the newly added topic the next time he or she logs into the system. The actual SearchTogether client is divided into multiple aspects of search to enable division of labor, provide persistence and support awareness. It is shown in Figure 12.

Figure 12: SearchTogether (Morris & Horvitz 2007)

On the left hand side, the area *a* can be used to chat with other group members while area *b* provides awareness about each individual queries. Each group member can use the toolbar *i* to recommend search results to others. These recommendations appear in area *d*. Area *c* shows current results of each user while area *j* is the actual browser that is used by each individual. Thus, each individual can search on his or her own, send results and recommendations to others while being able to discuss with them and see their current activities.

CoSense builds on SearchTogether and uses its features to enable collaborative search activities. CoSense is a tool for collaborative sensemaking. It takes information about each group members' search process as well as products and provides different visualizations using different views to support sensemaking activities (Paul & Morris 2009). Each user can log in into CoSense and by this, his or her data like search queries, comments or chats are imported into CoSense from the SearchTogether environment. Once a connection is established, the different views are updated in real time from any SearchTogether or CoSense instance. CoSense uses four different views with different representations of the information to enable sensemaking:

- **Search strategies view:** This view provides information about roles and skills of group members. It shows aggregated data about the group's and each individual's search process. A query history shows a visualization with queries by users and multiple clouds show group's and individual's keywords. A browsing history uses the same approach and shows domains visited by the group and each individual. An additional query history color-coded timeline shows each individuals query history to see how search activities evolved over time and users (Paul & Morris 2009).

- **Timeline view:** This view provides a chronological representation of all activities of all group members during one session. It combines visited websites and their associated comments and chat messages. Again the color-coding is used to distinguish multiple users and their activities. The timeline can be filtered and used to trigger new searches (Paul & Morris 2009).
- **Chat-centric view:** This view shows all chat messages by all users and allows to see active websites at the moment each message was sent (Paul & Morris 2009).
- **Workspace view:** The workspace allows to structure the collections of links and comments. Additionally, areas for creating to-do lists and notes are provided. Users can tag and rate the different links. The system enables to add external files like documents or photos, too (Paul & Morris 2009).

3.1.4 Cambiera

Cambiera is a tabletop visual analytics tool that enables individual and collaborative information seeking and foraging activities. It allows users to collaboratively search through documents while maintaining awareness of each other's progress and building on group partners' findings (Isenberg & Fisher 2009).

Each user sits at an edge of the tabletop display and is assigned to a user-specific color. The Cambiera system allows for up to four group members. Individuals start their search activities by pressing a colored search button on their side of the table, which triggers a virtual keyboard to appear. Active searches result in a colored *search box* with all search results concerning this keyword (Isenberg & Fisher 2009). Figure 13 shows two search boxes.

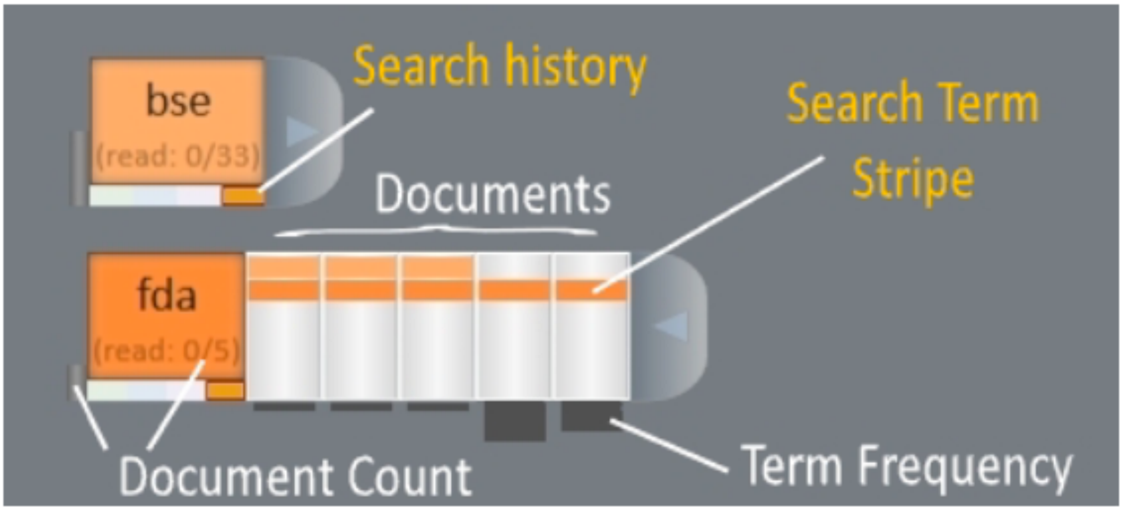


Figure 13: Cambiera's search boxes (Isenberg & Fisher 2009)

Each search box contains the keyword (in this examples “bse” and “fda”) and additional information: The background-color of the search box is connected to the user who initiated the specific search keyword. Each keyword search box receives the user-color in a different hue. This allows to distinguish between different search boxes and to relate them to their user. The number of found documents is shown underneath the keyword with information about documents that are already read by any user. Additionally, a bar shows the count, too. A search history highlights who has already triggered searches for this keyword and allows for discussion. Search boxes can be collapsed (compare search box on top) and expanded. Expanded search boxes show all found documents. Each document is visualized by a gray bar that gets darker every time a document is read. Stripes are used to show which keyword can be found in the document and gray bars at the bottom show the frequency of the keyword in the document (Isenberg & Fisher 2009).

By sliding a finger over the search results details-on-demand like the documents timestamp, title and sentences that include the search term as well as other keywords that also found this document are shown. Interesting documents can be pulled out of the search box. They can be enlarged and a full document reader is shown with the entire text. This document reader also features icons to show which user already has read this text. Selecting interesting words or snippets from documents allow to trigger searches. Figure 14 shows an overview of the Cambiera workspace with two individuals arranging search boxes and documents.

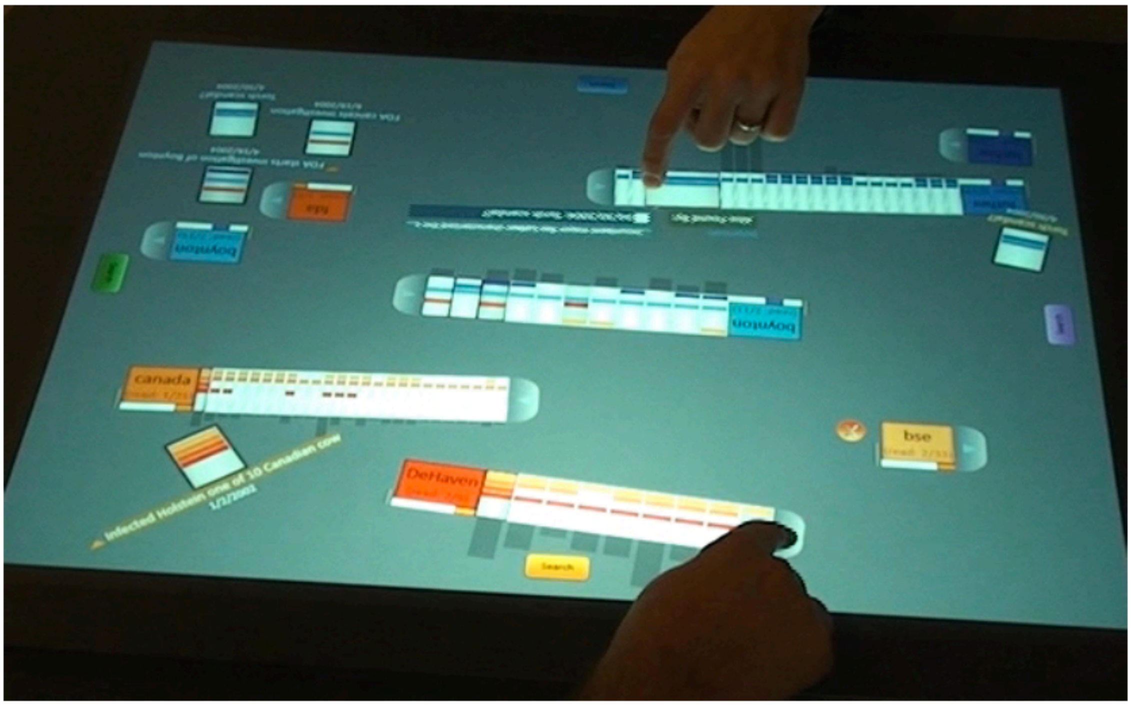


Figure 14: Overview of the Cambiera workspace (Isenberg & Fisher 2009)

Isenberg et al. (Isenberg et al. 2010) used their Cambiera system to study co-located collaboration of groups of two. They were able to identify eight collaboration styles that describe how closely or loosely coupled individuals interact with each other while working on an analytical task. The found collaboration styles expand the shown coupling styles in Chapter 2.2:

- **View Engaged:** One individual is actively working, the other one watches or engages in conversation.
- **Sharing of the same view:** Both individuals look at the same document reader or same search box together at the same time.
- **Sharing of the same information with different views:** Both individuals read the same document but use their own view.
- **Same specific problem:** Labor is divided and individuals read e.g. each one half of the documents.
- **Same general problem:** Individuals start different searches but want to find information on the same topic.
- **Different problems:** Individuals search for different problems.
- **Disengaged:** One person is actively working, the other is passively or disengaged.

Isenberg et al. grouped their coupling styles into the two groups *closely* and *loosely coupled* collaboration. The first five coupling styles belong to closely, and the last three to loosely coupled collaboration. These coupling styles form a basis to our *evaluation* shown in Chapter 5.

3.2 Research on Table and Tabletop Sizes

This chapter describes two research projects that deal with the influence and effects of table and tabletop sizes. This is important as the size of tabletops is a basis of our evaluation in Chapter 5.

Ryall et al. (Ryall et al. 2004) are interested in the effects of group size and table size in the domain of co-located group applications. They investigate how work strategies are affected by group size, how social interaction and thus communication and collaboration vary with respect to table size and how the speed of task performance is influenced by group size. Therefore, they identified different size considerations for tabletop design:

- **Resource Management:** Resources can either be physical or digital. Larger tables could make it difficult to share copies of resource with others (Ryall et al. 2004).
- **Work Strategy:** The table size could affect the distribution of work, the different roles that group members may assume and the problem-solving strategies (Ryall et al. 2004).
- **Social Interactions:** Small table sizes could lead to overlapping personal spaces and larger sizes could lead to less awareness about everyone else's activities (Ryall et al. 2004).
- **Display Resolution:** Increasing the size of a tabletop display does not mean increasing the display resolution as well. Higher resolutions provide more space to store information (Ryall et al. 2004).
- **Physical Reach:** "If you cannot reach something, you cannot interact with it." (Ryall et al. 2004)
- **Visibility**: As documents can be placed on the far side of the tabletop, they can be hard to reach and to read, this can limit the shared context of the group (Ryall et al. 2004).

- **Task:** Tasks that encourage users to use a certain area of the shared table, as their personal and individual work area might be better suited for larger table sizes and vice versa (Ryall et al. 2004).

Ryall et al. conducted an experiment to observe groups of varying sizes working on two different table sizes on a collaborative task. They used a within-subjects design. Thus each group of two, three or four persons had to participate in two trials. The experimental task was to search for words on the table with the goal of reproducing a target poem. The task is based on the game *Magnetic Poetry*[4].

They used two DiamondTouch devices with touch surfaces measuring 80cm and 107cm diagonally. Groups with two members were placed on the opposite shorter edges of the table, groups with three members added the third person to one long side of the table and groups with four members were equally distributed to each side.

As we are interested in the influences and effects of tabletop sizes, we only describe their findings concerning this topic:

- **The size of the table had no significant effect on the speed with which groups were able to assemble poems:** The application recorded the duration each group needed to build a poem. This time was divided by the number of words in the created poem. This results in a word-per-minute score for speed. No main effect for table size was found. (Ryall et al. 2004)
- **Subjects' agreement with certain statements about the task was affected by the size of the table, but never by the size of the group:** Subjects agreed more strongly with the statement "Overall, I felt the table was good for this task." for the large table (Ryall et al. 2004)
- **Table size did not affect the distribution of work:** Any of the five touch types (picking up, dropping, placing items into a container, taking items out of a container, menu operation) or the total touches were affected (Ryall et al. 2004).

Sepinood Hajizadehgashti (Hajizadehgashti 2012) studied physical features like the size of traditional tables on collaborative tasks. Communication and collaboration behaviors of groups with two members are observed while doing problem-solving tasks like travel planning or storytelling on two traditional tables with different sizes. Hajizadehgashti is interested in awareness and external cognition during taskwork. Therefore, two observa-

----

[4] Magnetic Poetry: www.magneticpoetry.com

tional experiments have been conducted with different table sizes, two tasks and various seating arrangements.

The experiment was designed as a mixed within-subjects design with two different tables sizes, two tasks and two seating arrangements (face-to-face and corner arrangement). All participants were assigned to both table sizes and tasks but only one seating arrangement was used for each group. The two tables are sized 1.24cm x 0.77cm and 154.5cm x 1.24cm. The two tasks are constructed as creative problem-solving tasks and include a story telling and a travel planning part with printed materials on each table (Hajizadehgashti 2012).

Interesting findings of the experiment are:

- Individuals tend to **distribute task materials** on the larger table (Hajizadehgashti 2012).
- Thus, more materials are visible at the same time, which facilitates **parallel searching** (Hajizadehgashti 2012).
- Individuals use the additional space on **the large table to build categories**. The smaller table forces users to **pile and overlap** artifacts (Hajizadehgashti 2012).
- The smaller table leads participants **to extend their personal space** on the table by holding artifacts in their hands (Hajizadehgashti 2012).
- The **seating arrangement** affects how the table space is used as there might be areas that are out of reach or overlaps of personal and shared territories (Hajizadehgashti 2012).

Figure 15 shows situations from the experiments with the smaller table (left) and the larger table (right). The smaller table leads to pilled-up artifacts whereas the larger table supports a better distribution of artifacts.

Figure 15: Different usages of space (Hajizadehgashti 2012)

4 System & Application

We implemented a prototype that supports analytical group work and allows for both coupling styles (Isenberg et al. 2010): loosely coupled parallel work and closely coupled collaboration. We opted for a reimplementation of TwisterSearch (Rädle et al. 2013) with modern web technologies like the open source platform Meteor[5] and HTML5, CSS3, and JavaScript. TwisterSearch is a distributed user interface for collaborative Web search. It was already tested at a German high school and its interaction concepts were conceived valuable by both students and teachers.

Figure 16: A group of two persons working with TwisterSearch

----

[5] Open source platform for building web applications: www.meteor.com

TwisterSearch allows multiple persons to search at the same time on their personal devices and contribute to group activities on a shared space such as an interactive surface (Figure 16). Thereby all search phases described by (Evans & Chi 2008) are supported.

Instead of providing Google as search engine as implemented in a previous version (Rädle et al. 2013), we replace it with an Elasticsearch[6] engine. Therefore, we can control indexed contents. As content we use the VAST 2006 contest dataset (Grinstein et al. 2006). This enables group members to search individually in a constant heap of articles, sheets and images.

The following chapters describe all important aspects of the different interfaces. At first, all interfaces that are used on personal devices are shown and described, then the focus changes to the shared group work space. Additional interfaces that serve administrative purposes like the login or creation of new user accounts as well as the entire user management are available but not shown in this thesis. Aspects regarding the development can be found in the technical report (Zagermann 2014). A video showing all aspects of the prototype can be found on the attached USB flash drive.

4.1 Personal Device Interfaces

TwisterSearch can be characterized as a distributed interface system according to (Elmqvist 2011): it uses multiple devices for input and output and thus, there are multiple interfaces. This chapter presents the different interfaces that are used on personal devices in various sub-chapters. Personal devices can be any devices that feature the size of phones or tablets. The application does not require an installation process due to its web-based architecture. The only requirements are a modern web browser like Google Chrome[7] or Apple Safari[8] and an active Wi-Fi connection. Even if the personal device interfaces are working on a variety of devices and web browsers, we chose to optimize it for the usage with Apple iPads 2. They represent a common size for tablets and their web browser provides all needed functionalities. Thus, the visual appearance might slightly differ if used with other devices.

----

[6] Real-time search and analytics engine: www.elasticsearch.org

[7] Google Chrome: www.google.com/chrome/browser/desktop/index.html

[8] Apple Safari: www.apple.com/de/safari/

4.1.1 Keyword Input and Personal Keyword Overview

Each user has its own personal mobile device that features a user-specific background color and can be used to add search keywords individually. They appear in a keyword list and each of them can be edited, deleted or shared with other group members on the shared tabletop (see Figure 17). Later, keywords serve as materialized thoughts and ideas. At first, this interface shows a single input field in the middle of the device. Adding a keyword moves this input field smoothly upwards. This allows for a better keyword overview and the ability to add further keywords. Deleting all keywords results in a reverse animation to focus on the keyword input.

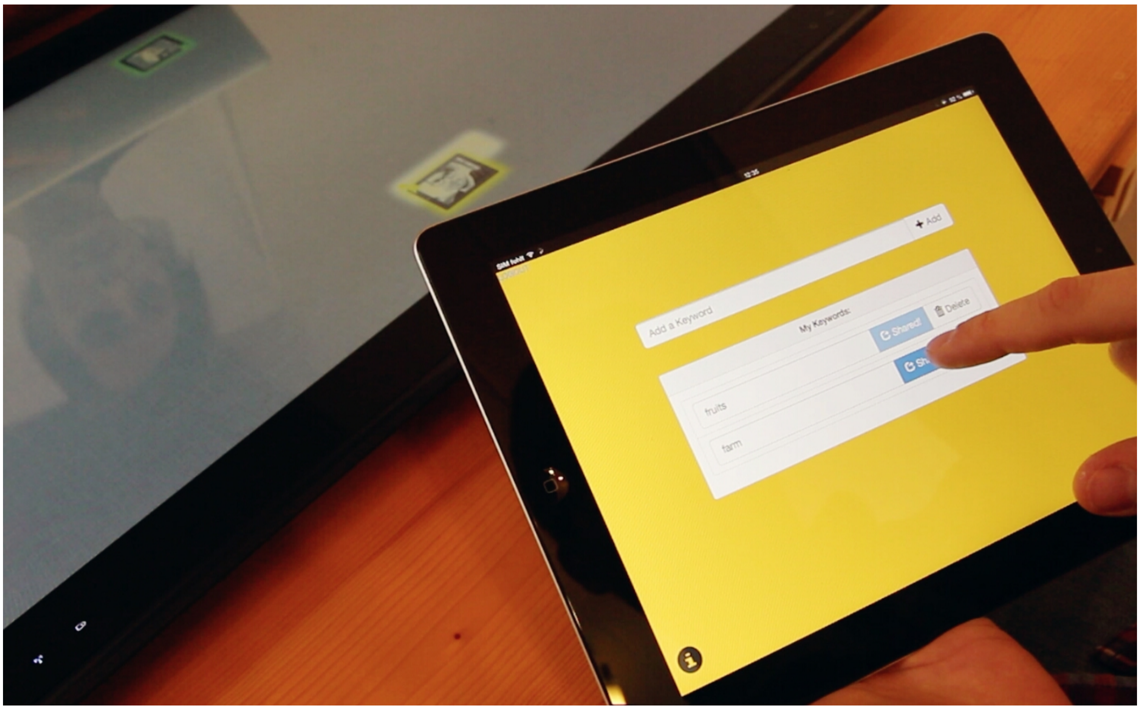


Figure 17: Keyword input and personal keyword overview

4.1.2 Individual Search

Placing the user-specific digital token (see Figure 23) that features an image of the user onto a cluster (see Figure 27) changes the interface of the user's personal device. This change is animated: All elements move smoothly from left to right. It allows users to better recognize that the interface changes when placing their token onto a cluster. The now shown individual search interface (see Figure 18) is divided into two major parts. The left-placed sidebar contains all keywords of the chosen cluster on the shared group space. Its background color is the same as the color of the cluster to have a visual connection between the personal device and the shared group space. Single or multiple keywords can be selected individually to trigger full text searches. Matching documents appear in a ranked search results overview list on the right hand side. This main part of the individual search interface features the already mentioned user-specific background color. The upper part consists of a navigation bar including buttons to navigate back and forward between the search results overview and different documents. Active keywords are additionally placed in a search box on top to allow a better overview. Users can easily deactivate them by tapping on the keyword. Since the search box has the look and feel of an input field as known from the Web, users can add further keywords by typing them into the search box with the device-own digital keyboard. Keywords added like this appear also on the sidebar of the personal device and on the shared group space in the center of the specific cluster.

The search results overview shows the number of found documents, the elapsed search time and the current results page. If the results list contains more than 10 hits, pagination at the bottom can be used to jump to different pages. The main part of the search results overview is the result list with multiple hits. Each hit features a document title and a document-specific overview with previews of sentences that contain all active keywords (see Figure 18). These previews help to gain a better overview of the specific document and users can easily scan and filter their results without reading entire documents. Icons are used to highlight documents that are already read by any group member and to favor search results.

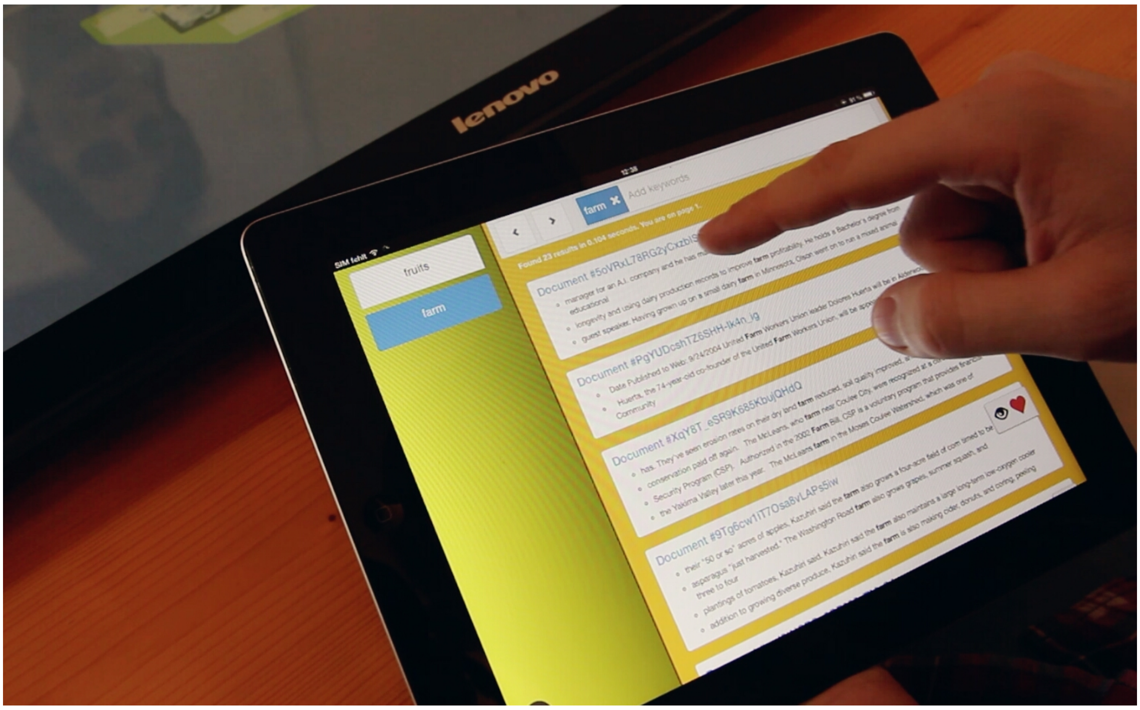


Figure 18: Individual search - search results overview

Documents can be selected and read by taping on the document title in the search results list. Text files, images and even spreadsheets are supported in the moment. The interface adapts to the document's type, size and length. Active keywords in active documents are additionally highlighted with a bold font-type to grant a better visual connection between keywords and documents and to get a better overview as there is no limitation to text's length (see Figure 19).

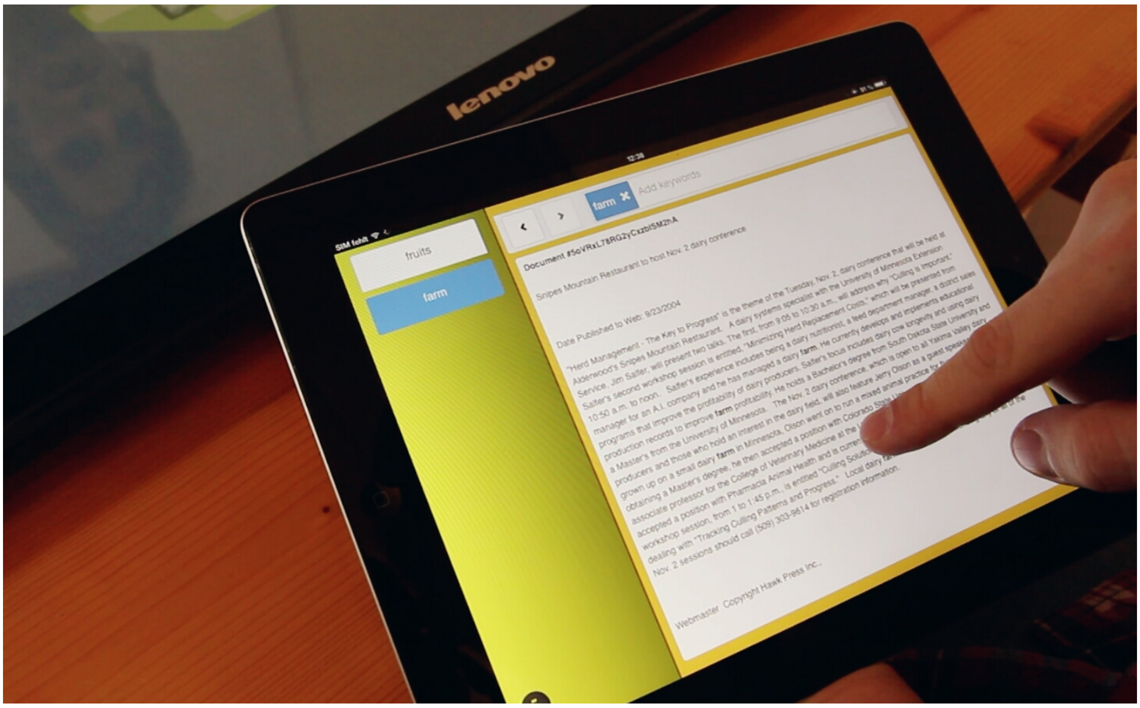


Figure 19: Individual search - reading a document

Interesting and important parts of documents can be copied and shared for further usage. Sentences or images can easily be selected with known selection mechanisms and further tapping on the "Copy"-button. A popup appears and the cut content can be complemented with additional comments and notes at the bottom (see Figure 20). Snippets like these can be shared with the group and appear on the shared group space in the center of the specific cluster.

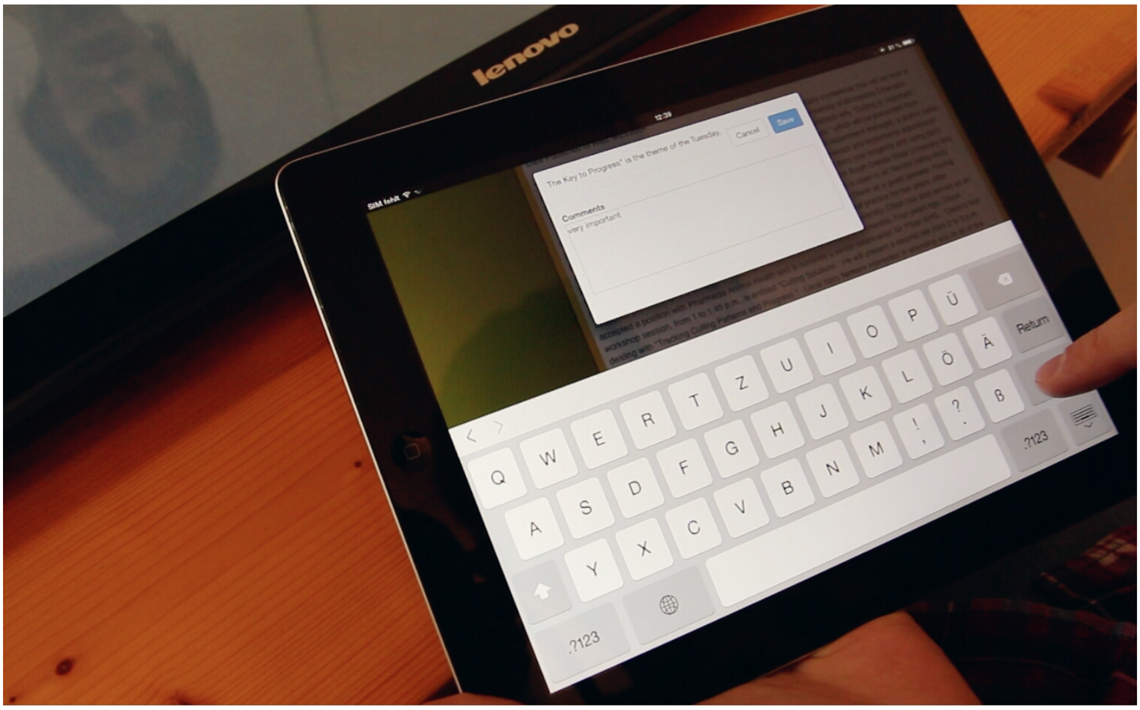


Figure 20: Individual search - adding a comment to an interesting snippet

4.2 Shared Group Space Interface

While the mentioned keyword input and individual search interfaces run on personal devices, the shared group space interface is available on an interactive tabletop. This interface is used to spatially arrange keywords, create clusters of them and to support group activities like communication and collaboration in general. The web-based development allows for a great variety of devices. The device only needs a modern web browser and an active Wi-Fi connection. We optimized the interface for usage with Google Chrome as this web browser supports all desired functionalities. Saying this, the TwisterSearch system combines a wide variety of devices with its several interfaces distributed to different devices by various manufacturers. Figure 21 shows the shared group space interface running on a 27″ Lenovo Horizon IdeaCentre Windows 8 tabletop system. Additionally, two users are actively working with their personal devices.

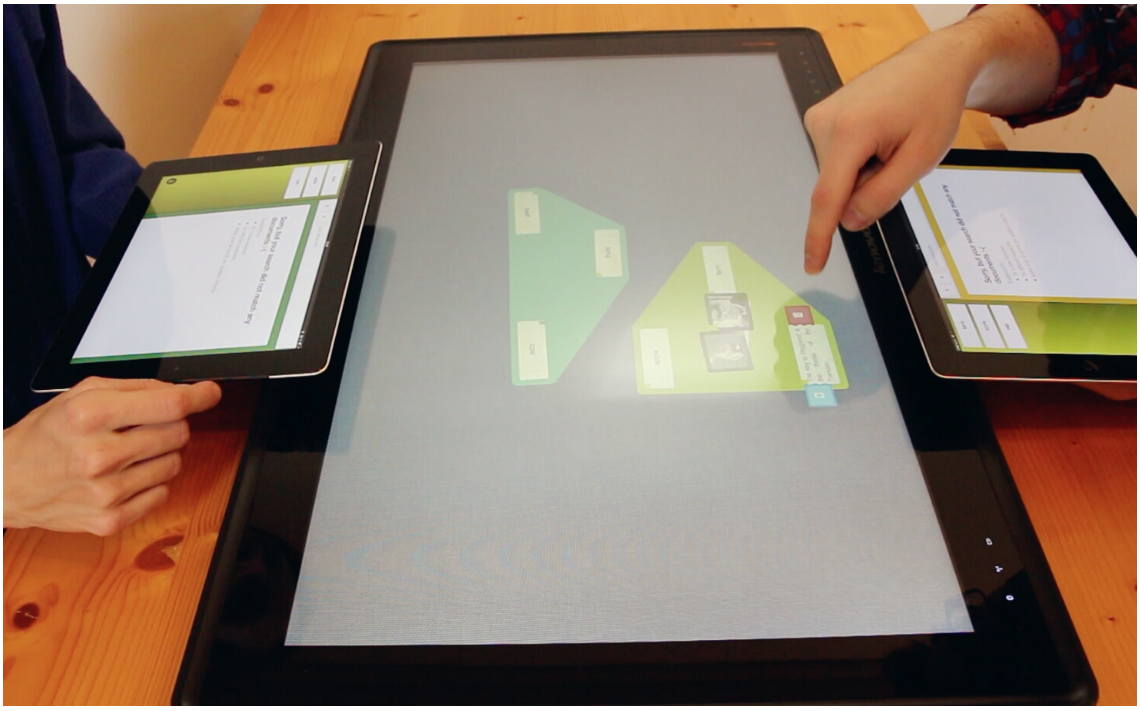


Figure 21: Shared group space with cluster, keywords and digital user tokens

The shared group space interface consists of a variety of different elements and users can individually interact with all of them equally.

4.2.1 Digital Elements of the Shared Group Space

The shared group space features several digital elements that are described in the following.

Keywords

The shared group space can be used to organize, arrange and spatially cluster keywords added by group members on their personal device. Newly added keywords appear next to the user-specific digital token and share its rotation. Thus, users can control the position and rotation by placing their digital token at a certain area of interest. The size of each keyword depends on the length of the text content, which is centered and wrapped if a predefined width is exceeded. The background-color is white with a paper texture to gain a look and feel of paper snippets. Each keyword has a colored badge on the top-left corner that visualizes the user who created it. Keywords that are not connected to a cluster have a white glow (see Figure 22).

Figure 22: Keyword on shared group space

Keywords can be moved and aligned freely. Unnecessary keywords can easily be deleted by tapping on them and pressing the red delete-button that appears.

Digital User Token

Each user has an own digital user token. That token can be used to define an area where newly shared keywords appear and to connect a user to a specific cluster of keywords. Thus, users can be aware about own assignments and other group members can easily comprehend who is working on which topic. The digital user token comes with the shape of a personal user device and features a profile picture of each user to have a direct connection between token and user. The token has the same colored badge as keywords and has a user-specific colored glow (see Figure 23). All tokens can be moved and aligned freely on the shared group space.

Figure 23: Digital user token

Clusters - Convex Hulls

Semantically coherent keywords can be clustered by aligning them spatially. In addition to this, drawing a line around them with a finger highlights them visually as a convex hull (see Figure 24 and Figure 25). Clustered keywords have no white glow. That visual feedback is used to recognize affiliation of keywords, as they also can be located on a cluster without belonging to it. Then they show their visual white glow. Keywords can be clustered and thereby can belong to multiple clusters at the same time to allow intersections and subsets of topics.

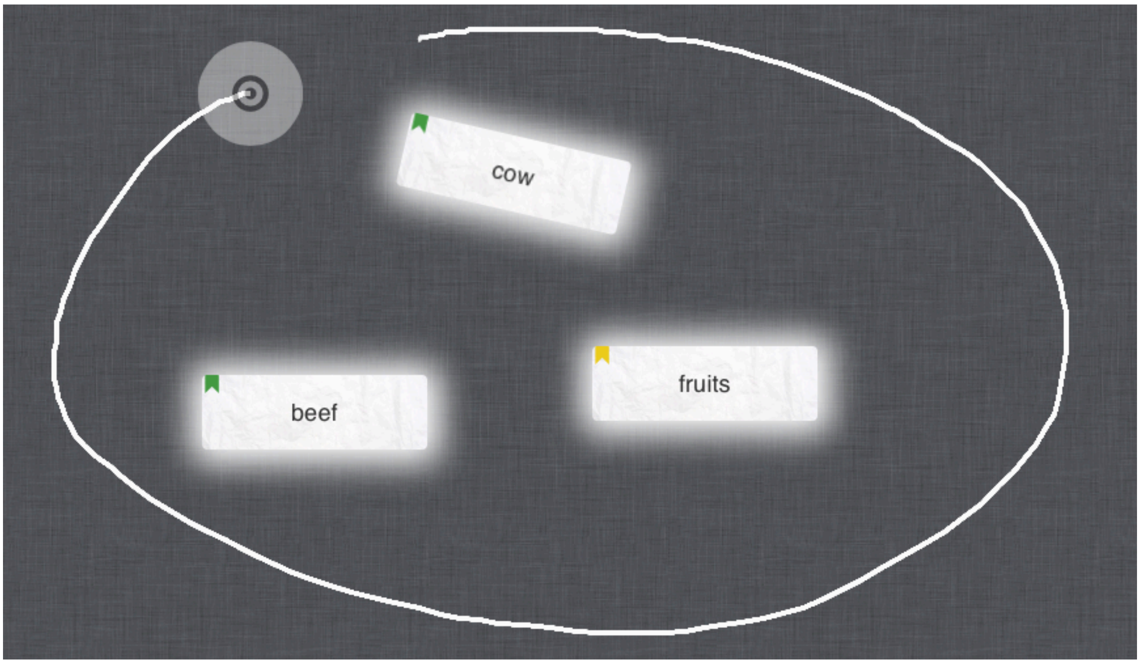


Figure 24: Encircle keywords to draw a cluster

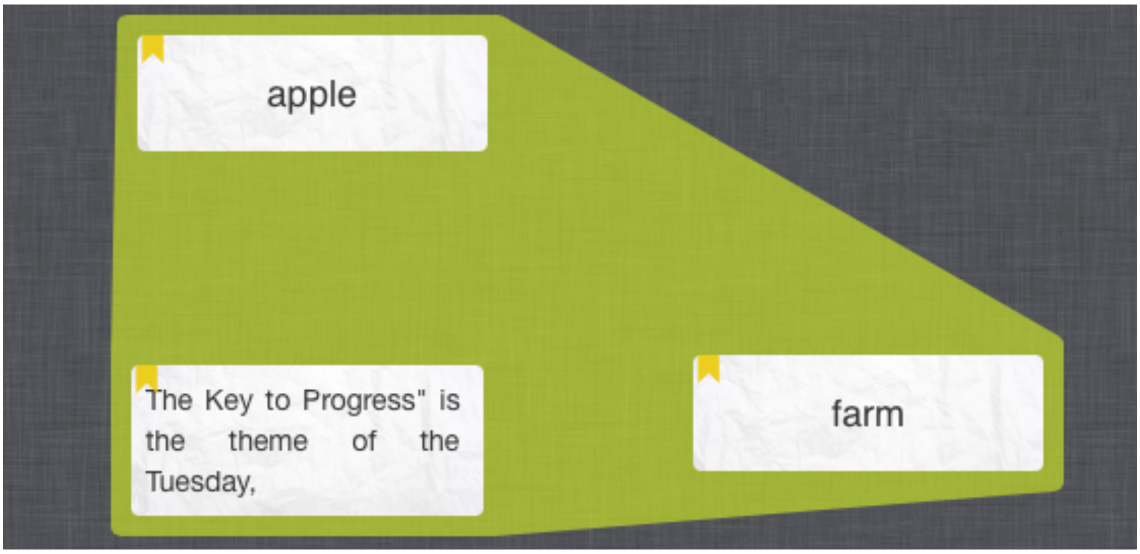


Figure 25: Cluster with keywords and a snippet

Figure 25 shows a cluster with two keywords and one snippet. Clusters are based on several ideas of the Gestalt laws (Goldstein 2014). Their spatial proximity on the one hand helps to identify the relationships whether keywords belong to a group or not. As each cluster has its own randomized background-color, keywords surrounded by the same color get perceived as a group due to the laws of similarity and closed form. This also allows to distinguish multiple clusters easily. Entire clusters can be aligned and moved freely. The shape of the cluster is updated and adjusted as content changes. As demands change, keywords can be lifted into or out of existing clusters. In contrast to this, snippets cannot be lifted and reassigned to other clusters to provide a connection, which cluster of keywords led to which snippet. Unimportant snippets can be deleted in the same way as keywords. Additionally, a second button on the left-hand side provides the opportunity to reread the snippet on personal devices (see Figure 26), as snippets on the shared group space are limited to a predefined size. Therefore, users have to place their digital user token on the specific cluster (see Figure 27). Digital user tokens hide their colored glow the same way keywords do when attached to a cluster.

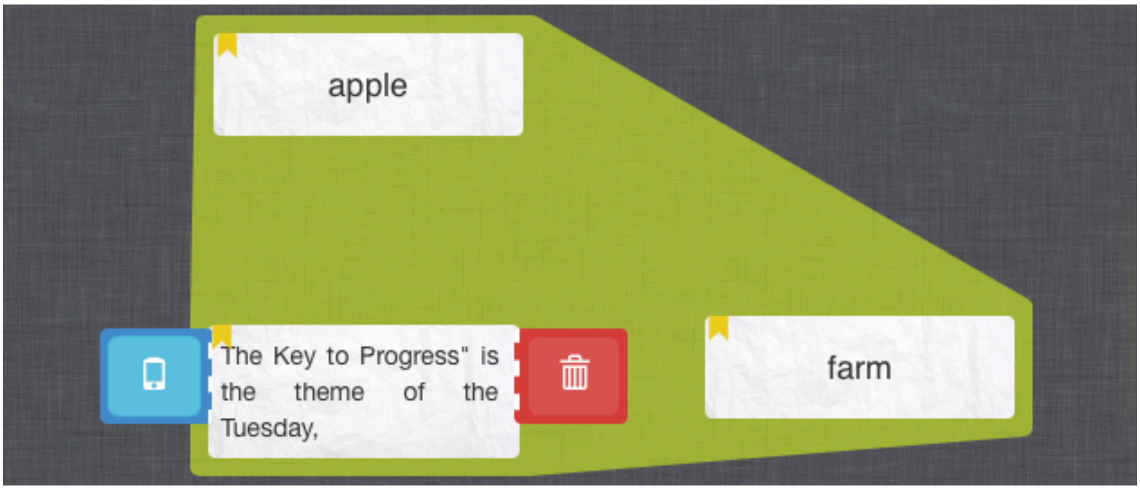


Figure 26: Share or delete snippet

Figure 27: Cluster of keywords with digital user token

Entire clusters can be collapsed and moved freely if the task's focus changes or due to spatial restrictions. Therefore, users have to perform a well-known pinch gesture to shrink the cluster to the size of its largest element. Keywords and snippets get piled up and a paper clip icon at the top right positions indicates the collapsed status of the cluster (see Figure 28). Collapsed clusters can easily be expanded again by a zoom-gesture. The paper clip icon disappears and all containing elements are shown again. An algorithm calculates the new positions as all items are smoothly aligned circularly around the element the gesture was performed on.

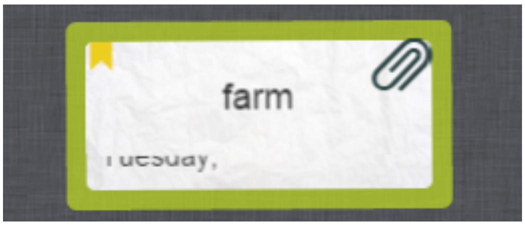


Figure 28: Collapsed cluster

5 Evaluation

The main research goal of this work was to understand the effects and other influences of tabletop display size during collaborative search. For this purpose, we want to compare different interactive tabletop display sizes in a predefined group work scenario, namely analytical group search. Conceivable effects, for example, are that larger display sizes encourage group members to take on equal roles and thus equally contribute to a task solution. Whereas smaller display sizes and due to their spatial restrictions could lead groups to evolve different roles (Heilig et al. 2011). Or perhaps, larger display sizes might help users to better externalize their thoughts and communicate them to others (Kirsh 1995; Kirsh 2010), which in turn could lead to a better shared mental model. Related to it, different sizes of a shared display might influence how group members interact, collaborate, and communicate with each other. For instance, a display size could dominate whether a group might work either loosely coupled or closely coupled (Isenberg et al. 2010).

These and other influences and effects of shared space sizes for group work scenarios are researched. This chapter shows different aspects of the evaluation regarding research question, variables & hypotheses as well as the study design, sources of data acquisition and results as well as the discussion of results regarding the research goal and question.

5.1 Research Question & Hypotheses

The following research question has been formulated out of the named research goal:

RQ: Do tabletop display sizes influence groups' communication and collaboration during collaborative search?

Therefore, the independent variable is *tabletop display size* and the dependent variables are *communication* and *collaboration* with various manifestations.

The research question is based on several hypotheses:

H1: Larger display sizes encourage group members to equally contribute to a task solution.

H2: Smaller display sizes lead groups to evolve different roles.

H3: Larger display sizes lead groups to a better shared mental model.

H4: Different display sizes influence how group members interact, collaborate and communicate with each other.

All hypotheses are tested with the help of different sources of data acquisition. They are described in Chapter 5.3 Sources of Data Acquisition.

5.2 Study Design

In the following sub-chapters the different aspects of the study design are shown in detail. These include the actual task and prerequisites for participants. In addition, the apparatus and procedure are presented.

Our study was designed as a between-subjects study with the shared display size as independent variable with three conditions. It was also designed for a group of two people. We chose dyads similar to Isenberg et al. (Isenberg et al. 2010) who were able to find significant coupling styles with just two group members. Five groups were recruited for each of the three conditions. The time to process the given task was limited to 90 minutes. In this time span participants had enough time to feel comfortable with the system, the given task, and their group member. School classes and lectures in universities have the same duration, as they are potential candidates to apply our research insights.

5.2.1 Task

Studying the influences and effects of different shared space sizes for group work activities implies the availability of a task that matches multiple requirements:

- The task should be solvable as a group, which means it should be possible to distribute results to other group members.
- The cognitive demand should be almost constant during the task to avoid confounding factors.

- The task's topic should be interesting and entertaining to keep participants interested for a period of about 90 minutes to avoid early termination of the task.

Therefore, we decided to use the "Stegosaurus" (Grinstein et al. 2006) scenario data set from VAST 2006 Challenge. This scenario was developed by the National Visual Analytics Center at Pacific National Labs and can be seen as a standard visual analytics task. The scenario describes a fictional story with multiple articles, images and data sheets. Groups have to find relevant information, filter non-important articles and connect important facts that lead to new insights and suppositions.

5.2.2 Participants

There were no special requirements or restrictions for participants. Age, field of study or experiences with computers were not of interest. As our task and thus all information was in English, we asked possible participants if they dare to read and comprehend articles written in English and in general if they mind to work in a group of two. If these prerequisites were passed, they were asked to send a profile picture of themselves that was used for their digital user token during the course of the study by email. If they agreed to participate but did not want to share their personal pictures (4 of 30), they were assigned to a character of the American animated sitcom *The Simpsons*[9]. Thus, each participant was able to use the provided digital user token. Providing a picture of each participant allowed to prepare and set up the environment by the examiner before the actual session. Alternatively, the TwisterSearch system allows to create new user accounts dynamically including contact data, user-specific colors and the ability to take a picture on the fly.

To recruit the participants we used postings and flyers looking for "spotters" (see Appendix C). That helped us to find participants who were interested in solving puzzles, mysteries or in this case the VAST 2006 contest task. The goal was to have 5 groups of 2 persons each for each of our three conditions. Isenberg et al. (Isenberg et al. 2010) also used dyads to study and find significant coupling styles. Thus, there were 30 participants in total.

5.2.3 Apparatus

Participants had separated working desks to fill out the questionnaires in the beginning and in the end of the study. During the actual task, they were sitting face-to-face at a conventional office table (1.4x0.8m) on which we placed the different interactive surfaces

---

[9] The Simpsons: www.fox.com/the-simpsons/full-episodes

that served as a shared group space. We chose the three sizes: 10.6″, 27″, and 55″ (see Figure 29) as each of them constitutes a common size for personal mobile device, personal workspace and multi-user workspace.

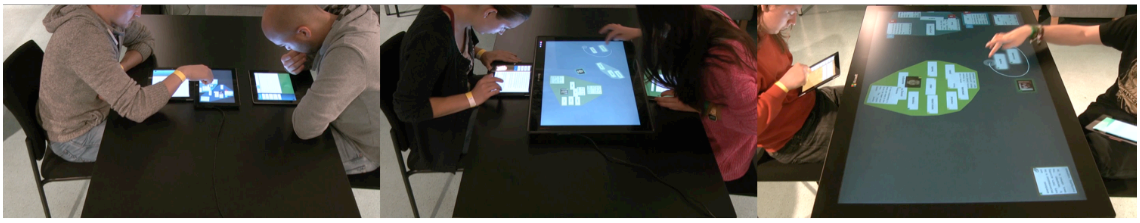


Figure 29: The three different interactive surfaces

All three interactive surfaces have the same display resolution of 1920x1080 pixels, which guarantees for a high internal validity of the study. Each participant was provided with an Apple iPad 2 (9.7") as a personal device. The setting is shown in Figure 30.

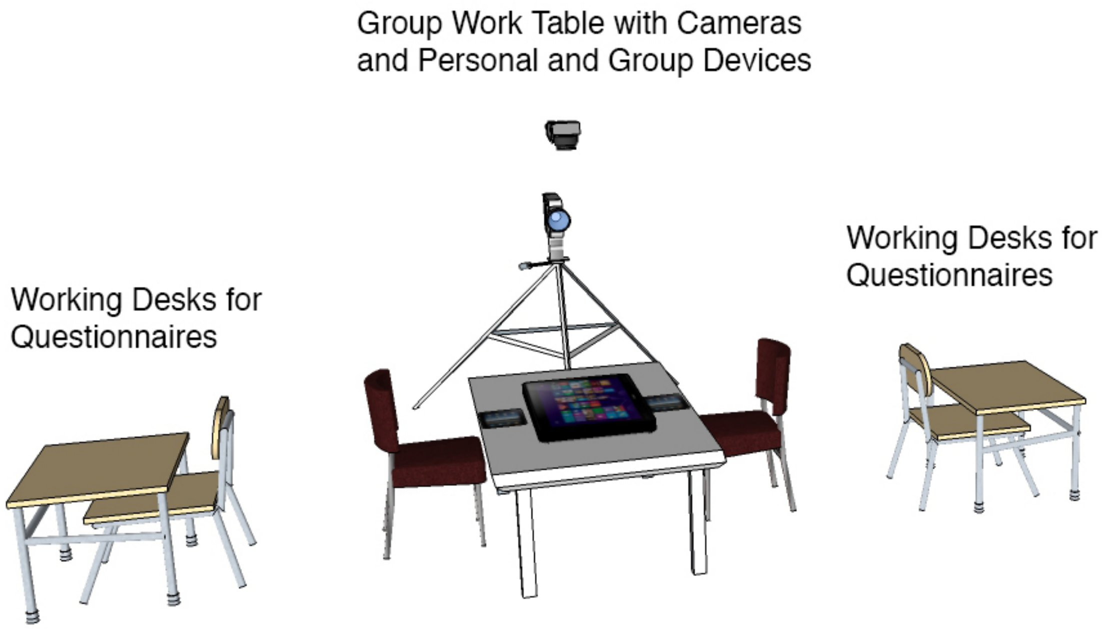


Figure 30: Setting of the study

Group members were assigned to individual colors (green or yellow). This personal color-coding was used for the background-color of their personal device, the outline-color of their digital proxy on the shared space and as badge-color of each item they create. That provided a visual connection between the participants, their personal devices, the shared space, and of course their group partner. Additionally, each user was also provided with two colored wristbands to distinguish user interactions in video recordings (helpful for later video coding).

We used two cameras to record the study. One camera was installed in birds-eye view above the table and a second camera recorded the scene from a side position to be able to see gestures, postures, and interactions of both participants. In addition, the examiner noted important interactions and statements during each session.

5.2.4 Procedure

Participants were seated on two separated desks at the beginning of the study. They were asked to fill out a questionnaire to gather information about their demographics and experience with technology. After the participants finished their questionnaires, they were seated to both long sides of the table, opposite each other. They received a tutorial on all features of TwisterSearch and were given time to explore the system with a sample dataset until they felt comfortable using it. The examiner then introduced them to the VAST 2006 task. During the study, they were provided with the introductions on their personal device if necessary. They were given a maximum duration of 90 minutes to solve the task. During the task they were not allowed to ask questions about their performance or how to solve different aspects, but general question about the procedure were answered. The participants were informed 10 minutes before the end of the time span that their procedure is going to be stopped in 10 minutes. The group work was stopped after the maximum duration and both group members were seated to their initial desk without visual sight to each other and to the interactive tabletop. They were asked to individually fill out a second questionnaire. Each session lasted in total about 2 hours and participants were compensated for their time. As the correct solution of the task officially takes up to three hours, the time was fixed on 90 minutes. If the group was not finished after 90 minutes, the conductor interrupted the group and asked them to continue with the post-questionnaire. The following table shows the exact course of actions.

Action	**Time**
Welcoming	2 min
Pre-Questionnaire	3 min
Tutorial	5 min
Exploration	3 min
Task	90 min
Post-Questionnaire	15 min
Farewell	2 min
Total time	120 min

Table 2: Procedure of the study

5.3 Sources of Data Acquisition

Different types of data sources were used in this study: observations, questionnaires, video recording, and data logging. The combination of multiple sources strengthens a finding, when found in two or more of the named sources (data triangulation). Different sources were used to collect qualitative as well as quantitative data to explore influences and effects of different sizes of interactive shared group work spaces and thereby to answer the research question. The following sub-chapters explain each data source in detail.

5.3.1 Observations

The conductor of the study observed the participants while they were working on their specific task. Traditional group work encourages communication between group members, so does TwisterSearch. Communicating participants allowed the conductor to observe and understand their behavior in a better way. To log the observations, the conductor used a protocol to define time and specific event. The observations were focusing on several points:

- How do participants use the shared group work space?
- How do participants communicate and interact with each other?
- Do they work on same or on different topics?
- Do they evolve different roles?
- How do they use the different features of TwisterSearch?
- Are there any unexpected behaviors?

Additionally each session was video recorded. The observation protocol helps to understand the videos that are analyzed in Chapter 5.3.3.

5.3.2 Questionnaires

Two questionnaires were used in the context of this study: a pre-questionnaire and a post-questionnaire. The pre-questionnaire had to be answered before the participants started with their task and the post-questionnaire finalized each session.

<u>Pre-Questionnaire</u>

The pre-questionnaire was used to collect demographic data such as sex, age, handedness or color blindness. Information about the daily usage of computers, the amount of years of experience with them and a self-assessment of their skills on a five point Likert scale helped to classify participants. Participants needed to explain their experiences with touchable devices, their usage and frequency of use for a better understanding. The pre-questionnaire can be found in Appendix A

Post-Questionnaire

The post-questionnaire had to be filled out individually after completing the task. It consisted of several sub-questionnaires and covered several aspects of group work. At first, participants needed to fill out a NASA-TLX questionnaire (Hart & Staveland 1988). The NASA-TLX uses six subscales like Mental Demands, Physical Demands, Temporal Demands, Own Performance, Effort and Frustration that had to be rated by participants. It was used to gain information about the overall subjective workload of each participant.

The second sub-questionnaire was a teamwork questionnaire adapted from (Heilig et al. 2011). 16 different statements about the group work, the communication and the support by the TwisterSearch application had to be answered with the help of a seven point Likert scale. It ranged from *strongly disagree* to *strongly agree*. This teamwork questionnaire provided information about the subjective experience of each participant concerning group work processes.

The third sub-questionnaire asked about the distribution of roles during the task. Participants were asked to describe their own and their group partner's role. If the group did not evolve roles, the participants needed to explain the reason for it.

The next sub-questionnaire dealed with the procedure during the task. Participants had to describe their own and their partner's procedure. They were allowed to use a step-by-step overview for this.

The next two sub-questionnaires were used to get information about the shared mental model. First, participants had to write down all their insights and results of their task. In a second step, they had to use the provided pens to draw the final structure of the shared group work space.

The last sub-questionnaire asked for the usage of the shared group work space for their group processes. All sub-questionnaires can be found in Appendix B .

5.3.3 Video Recording and Video Coding

Each session was video recorded with two high definition cameras. One of them was placed in bird's-eye view above the table and the other was recording the entire scene from the side. The latter perspective allowed to see both participants, their personal devices, the shared group work space and all interactions as well as communication in between. The setting is shown in Figure 30. As each of the 15 sessions lasted about 90 minutes there were 1350 minutes (22,5 hours) in total for each camera. As the second camera from the side point of view showed all relevant aspects of group work, we only analyzed this recording material.

The video recording was used to find possible patterns, influences and effects of the size of the shared group work space regarding collaboration and communication of groups. Therefore, the recorded video material needed to be coded. Coding videos means to attach predefined behaviors to specific sequences. This allows quantifying qualitative aspects like groups' communication or interaction with several devices. The raw outcome of this procedure is a spreadsheet per video file with all behaviors, times and duration that can be used for further processing and analyzing. Before starting to code all videos, one has to define all codes. They are described in the following.

Video Coding

All used codes are based on literature and a qualitative pre-analysis of the material. The latter is used to find several criteria in the recordings, as there might be differences to the codes based on scientific sources. Pre-analyses are part of the content analysis, an empirical method to describe contentual and formal characteristics of communication (Früh 2011). A detailed qualitative pre-analysis as described by Früh (Früh 2011) was conducted first. This pre-analysis consisted of watching a representative amount of video material and protocoling interesting aspects alongside.

In total, 240 minutes of the existing video material were used for the qualitative pre-analysis. It contained watching two videos of each of the three conditions for the duration of 25 minutes. The chosen sequences took place at various stages in the middle of each session. Thus, phases like getting to know the group partner and the TwisterSearch system were excluded as they might differ from the actual collaboration. Additionally, another video was protocolled entirely (90 minutes) to find criteria concerning the beginning and the end of sessions.

In addition, criteria from literature were used. As described in Chapter 3 Isenberg et al. (Isenberg et al. 2010) were able to find eight different coupling styles concerning groups communication and collaboration. Heilig et al. (Heilig et al. 2011) were able to find different search and work strategies in a collaborative information seeking task. All used codes are shown in Table 3. During the qualitative pre-analysis of the recordings all relevant coupling styles from (Isenberg et al. 2010) were found. As the systems (Cambiera and TwisterSearch) both were used as tools to investigate collaboration and communication, the found coupling styles using Cambiera were adapted to the specifications of Twister-

Search. Additionally, relevant communication and collaboration behaviors found during the qualitative pre-analysis were added.

Code	**Description**	**Origin**
Active Discussion	Both participants talk about the task at the same time, they discuss several aspects	Isenberg et al.
View Engaged	One participant is reading, the other one watches or engages in e.g. conversation	Isenberg et al.
Sharing of the same view	Both participants look at one single personal device or one participant shows his device	Isenberg et al.
Sharing of the same information / Different Views	Both participants read for example the same document on each of their personal devices	Isenberg et al.
Same specific problem	Both participants are working with the same cluster	Isenberg et al.
Same general problem	Participants work on different clusters but have a similar topic	Isenberg et al.
Different problems	Participants work on different clusters with different topics	Isenberg et al.
Disengaged	One participant is actively working, the other one is passive or fully disengaged	Isenberg et al.
Combination	Participants combine information on different personal devices (e.g. combining a map with an article)	Qualitative pre-analysis
Explaining	One participant explains articles or answers questions	Qualitative pre-analysis
Delegation	One participant delegates the other one (e.g. verbally)	Qualitative pre-analysis

Info	The conductor of the study provides information	Qualitative pre-analysis
Procedure	Participants talk about their procedure	Qualitative pre-analysis
Question	One participant asks the other one a question	Qualitative pre-analysis
Recap	One or both participants recap the found information	Qualitative pre-analysis
Structure Group	Both participants structure the shared group work space	Qualitative pre-analysis
Structure Single	One participant structures the shared group work space while the other one is reading	Qualitative pre-analysis
Irrelevant Task	Statements or actions that are not relevant for the task	Heilig et al.
Irrelevant Procedure	Statements or actions that are not relevant for the procedure	Heilig et al.

Table 3: Codes derived from literature and a detailed qualitative pre-analysis

The actual coding of all sessions was done with Noldus Observer XT 11.5[10], an application for the collection, analysis and presentation of observational data such as video material. The Noldus Observer XT 11.5 needs a schema that has to be set up before the actual coding. Therefore, *Observer* distinguishes *Subjects*, *Behaviors* and *Modifiers*.

Subjects can be participants, the entire group or devices. They build the basis of the observations. Each *subject* can have several *behaviors*. *Behaviors* are the actual activities of participants. *Behaviors* can be specified by *modifiers*. An example is used to clarify this relationship:

Participant A explains an article to participant B.

In this case Participant A is a *subject*, explaining is the *behavior*, which is specified by the *modifier* article. As *Observer* uses the technique to map subjects, behaviors and modifiers

---

[10] Noldus Observer XT: www.noldus.com/human-behavior-research/products/the-observer-xt

to predefined keys on a keyboard that have to be pressed whenever one of the activities occur, the named example needs three different keys to be coded. *Observer* is case-sensitive. That allows for a great variety of codes, but the more codes are used, the more errors are possible as coding video material is a very demanding task due to rapid changes in communication and interaction of several group members.

The flexible structure of *Observer* allows to adapt the relationship between subjects, behaviors and modifiers to specific needs. Therefore, another approach was used for the coding of the actual video material. The self-developed approach was on the one hand adapted to the predefined codes (see Table 3) and on the other hand to ensure a better workflow and a lower error-proneness.

As shown in Table 3 most codes describe the behavior, communication and interaction based on the entire group. The behaviors of single participants therefore were not focused as they can be classified to group behaviors. But in addition to this, there are behaviors that describe the action of single participants like *explaining* or *questioning*. It's important to gain knowledge about the owner of these specific behaviors as they might lead to insights concerning user roles and group behavior.

There was no need to distinguish several *subjects* when focusing on the entire group. *Behaviors* were attached to the group activities and specified with *modifiers* if relevant. Thus, the shown example can be coded with only two instead of three codes to accelerate and simplify coding. No *subject* is defined, while the *behavior* is explaining and the modifier is A to B. As nearly all provided documents in this task were articles, the *article* modifier from before could be ignored.

Observer allows to relate all *behaviors* with mandatory *modifiers*. That lowers the error-proneness, too. At first, all codes had to be assigned to keys on the keyboard. The shown codes of Table 3 were translated to *behaviors* and *modifiers* were added if needed. Table 4 shows the mapping of all codes:

Behavior	**Key**	**Modifiers**	**Key**
Active Discussion	Q	--	--
View Engaged	W	left to right / right to left	L / R
Sharing of the same view	E	left to right / right to left	L / R
Sharing of the same information / Different Views	R	--	--

Same specific problem	T	--	--
Same general problem	Z	--	--
Different problems	U	--	--
Disengaged	I	left to right / right to left	L / R
Combination	K	--	--
Explaining	A	left to right / right to left	L / R
Delegation	S	left to right / right to left	L / R
Info	C	--	--
Procedure	J	--	--
Question	H	left to right / right to left	L / R
Recap	G	--	--
Structure Group	D	--	--
Structure Single	F	left to right / right to left	L / R
Irrelevant Task	Y	--	--
Irrelevant Procedure	X	--	--

Table 4: Behaviors and modifiers mapped on keys

The shown modifiers in Table 4 only have two states: Left to right and right to left. This helped to be aware about the direction of communication or interaction. As participants were seated face-to-face, one participant was sitting on the left and the other one on the right hand side of the video. The first value of the modifiers (e.g. left) described the active participant whereas the second value (e.g. right) marked the passive participant. Two examples are used to illustrate this:

Participant A (left) asks participant B (right) a question.

First, the key for behavior *Question* needed to be pressed. Then *Observer* prompted an info and the modifier *left to right* was chosen.

Participant A (left) reads an article while participant B (right) is sleeping.

In this case, participant B was *Disengaged* in group work activities. As participant A was reading an article, the modifier *left to right* was chosen.

The coding scheme could easily be populated to *Observer*. Thus, the environment was set up and videos could be loaded individually into the system to start coding. *Observer* is restricted to certain video codecs and file sizes. Too large files resulted in juddering videos or even separation of video and audio materials. That inhibited proper video coding. Therefore, a two-way rendering procedure was used to adapt the video material to the *Observer* system:

The recorded video material differed in its length since the recording was started before and stopped after the actual group task to ensure to capture all relevant information. The raw material therefore varied in its length between 100 and 120 minutes, which resulted in file sizes of 6 to 6.5 GB per recording. The format of the files was *mov.* First, each video was converted with Handbrake[11] to a MP4 file with a H.264 video codec. This shrinked each video to half its file size without quality loss. This first step took about 45 minutes per video file. As a second step, each video was loaded into *Adobe Premiere Pro*[12]. Premiere Pro allows to edit and cut video sequences. All videos were cut to their specific task length of about 90 minutes and were rendered with a resolution of 640x360 pixels (Origin: 1920x1080 pixel). This procedure took about 110 minutes per video. This resulted in a total rendering length of 2325 minutes or 38,75 hours. The rendered video files could then be used with the *Observer* system.

Each video was loaded individually into the coding environment. This enabled to render and code videos synchronously. The setting of the video coding environment is shown in Figure 31. In total three different computers were used at the same time to render and code videos. As *Observer* only runs on Microsoft Windows computers, a high-end Windows laptop was used to code all video files. In addition to this, a high-end Mac Pro was used to render the raw videos into the specific format and length synchronously. A MacBook Pro was also used to show all behaviors with their specific keyboard codes for a better overview.

---

[11] Video converting tool Handbrake: www.handbrake.fr

[12] Video editing environment Adobe Premiere Pro: www.adobe.com/de/products/premiere.html

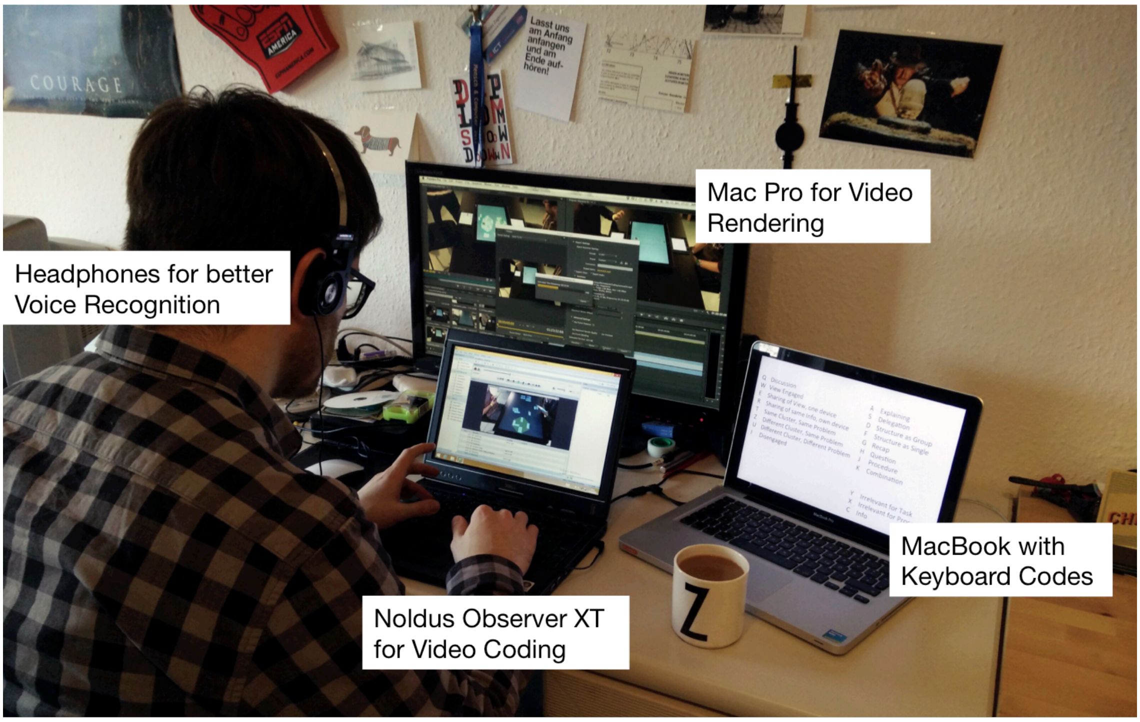


Figure 31: Setting of video coding

The combination of three different computers saved time and ensured high quality videos for coding. The support of an overview screen as well as headphones for better voice and sound recognition enabled an efficient coding workflow with low error-proneness. If errors occurred, the recording had to be stopped or rewound to correct mistakes. *Observer* provided these possibilities. They were also used to watch sequences again in doubt as communication and interaction could change rapidly. Each coding session lasted about 140 minutes including time for corrections and breaks. This resulted in a total coding time of about 35 hours for all 15 videos. This fast coding time can be based on the self-developed approach. Outsourcing different parts of the coding procedure such as video rendering, overview and the actual coding on different computers allowed parallel work-flows to save time, ensure high quality and low error-proneness. Observer allowed to ex-port the coded material e.g. as CSV or XLS files for further usage for example via Microsoft Excel or statistical tools like IBM SPSS Statistics[13].

5.3.4 Data Logging

Events concerning the interaction with the prototype were logged and stored in a data-base. This quantitative data was not used for statistical comparison, but as additional

---

[13] IBM SPSS Statistics: www-01.ibm.com/software/de/analytics/spss/

triangulation tool. It could help to reason findings emerged from qualitative data. These events are logged:

- Adding keywords using the keyword input interface
- Sharing of keywords using the keyword input interface
- Changing text of keywords using the keyword input interface
- Adding keywords during individual search
- Activating keywords for individual search
- Deactivating keywords for individual search
- Adding snippets
- Adding comments to snippets
- Deleting keywords and snippets
- Connecting user to cluster via digital user token
- Disconnecting user from cluster via digital user token
- Adding items to cluster
- Removing items from cluster
- Collapsing of clusters
- Expanding of clusters

5.4 Results

This chapter shows results from the various sources of data acquisition. All statistical analysis is done with the help of IBM SPSS Statistics.

5.4.1 Demographic Data and Experience

30 participants (15 female, 15 male) were recruited to take part in the experiment. The mean age was 22.2 years (SD = 2.6, aged 19 - 29). 27 participants were right-handed, two participants were left-handed and one participant did not answer this question. None of the participants had color vision deficiency and thus no problems with the employed color-coding of the digital user token and user devices. 28 participants were students from non-technical subjects such as psychology, sociology or politics. One participant had a computer science background and one participant was an employee of the university. For each of the three different display sizes we assigned five groups of two participants each – two groups were both female, two both male and one mixed group.

Participants have a mean experience with computers of 11.9 years (SD = 3.7, 7 - 20 years) and see themselves neither as experts nor as beginners (median 3 on a five point Likert scale). 25 of 30 have used touchable devices before and use them often (median 4 on a five point Likert scale).

Number of Participants	30
Sex	15 female, 15 male
Mean age	22.2 years (SD = 2.6)
Job	29 students, 1 employee
Handedness	27 right-hander, 2 left-hander, 1 no answer
Color-Blindness	0 participants
Mean computer experience	11.9 years (SD = 2.7)
Median computer skills	3 - scale: 1 (“beginner”) to 5 (“expert)
Experience with touch devices	25 Participants
Frequency of usage	4 - scale: 1 (“very seldom”) to 5 (“very often”)

Table 5: Demographic data

5.4.2 Video Analysis

The video analysis reveals insights about the way participants collaborate and communicate with each other. The shown coupling styles in Table 3 and Table 4 are aggregated to different groups to gain various levels of insights in group work activities. Isenberg et al. (Isenberg et al. 2010) grouped their eight coupling styles into the groups *closely* and *loosely* coupled. In this case, the various coupling styles are additionally grouped into five groups that are shown in Table 6.

Talking	**Reading**	**Structuring**	**Show and Tell**	**Other**
Delegation	Same Cluster, Same Problem	Structure as Group	Combination	Disengaged
Discussion	Different Cluster, Same Problem	Structure as Single	Sharing of same Information	Irrelevant Topic
Explaining	Different Cluster, Different Problem		Sharing of View	Irrelevant Procedure
Procedure			View Engaged	
Question				
Recap				
Other				

Table 6: Grouped coupling styles

The groups represent several activities that happen during the task's solution. *Talking* combines all actions where at least one group member is talking. *Reading* stands for all activities where group members read articles on their personal devices. They can either read articles from the same cluster and thus facing the same problem, they can read articles from different clusters with the same problem and they can read articles that can be found in different clusters focusing on different problems. The latter is highlighted as it can be described as *loosely* coupled whereas the non-highlighted items in this table focus on *closely* coupled activities. *Structuring* consists of activities where one or both group members structure content on the shared work space. *Show and Tell* describes all actions where group members e.g. take a look at the exact same information or show found articles on their personal device to their group member. The last group is *Other.* Irrelevant behaviors can for example be flirting participants whereas *disengaged* shows all activities where one group members does not participate in group work. This is therefore also marked as *loosely* coupled.

The task completion time is limited to a maximum duration of 90 minutes; therefore each session has a related length and absolute times of all sessions can be statistically compared.

The observation of the medians of the five groups shows obvious differences for the three different tabletop sizes regarding the groups *Talking* and *Reading.*

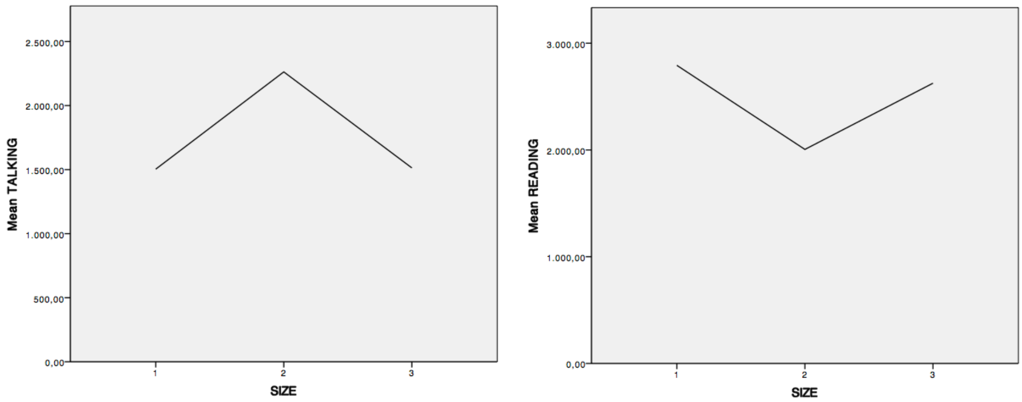


Figure 32: Means for the groups TALKING and READING (y-axis: mean duration in seconds)

Figure 32 shows visualizations for the means of the groups *Talking* and *Reading*. On the x-axis the three different tabletop sizes (1: 10.6″, 2: 27″, 3: 55″) are listed. On the y-axis, the mean duration in seconds is shown. There is an obvious difference: The mean *Talking* time of tabletop size 1 is 1503,38s (SD = 204,33), size 2 shows a *Talking* time of 2262,87s (SD = 365,73) and size 3 1513,38s (SD = 898). In contrast to this, the *Reading* times are 2792,83s (SD = 399,28) for tabletop size 1, 2005,05s (SD = 423,28) for size 2 and 2625,44s (SD = 936,71) for size 3. As a first step, a Median test is used to verify if these obvious differences are significant. The results are shown in Table 7.

	TALKING	STRUCTURING	READING	SHOW AND TELL	INFO	OTHER
N	15	15	15	15	15	15
Median	1689,3600	427,7100	2684,1700	216,3600	48,5300	,0000
Chi-Square	8,571[b]	,536[b]	6,964[b]	,536[b]	,536[b]	2,500[c]
df	2	2	2	2	2	2
Asymp. Sig.	**,014**	,765	**,031**	,765	,765	,287
Exact Sig.	**,021**	1,000	**,068**	1,000	1,000	,725
Point Probability	,012	,466	,047	,466	,466	,659

Table 7: Median test for the grouped coupling styles on a basis of absolute values

The Median test shows a significant effect for the groups *Talking* and *Reading.* The significance level of 0,05 is matched by *Talking* with an asymptotic significance of *0,014* and an exact significance of *0,021*. The asymptotic significance of *Reading* shows a value of *0,031*, which matches the significance level, whereas the exact significance is higher (*0,068*). The Median test has a low power rate, which means, that its calculations are not as strict as other tests. Therefore, a second test is used. The non-parametric Kruskal-

Wallis analysis allows to compare independent samples with different sample sizes with more than two groups. This test ranks the different groups and shows in general if a significant effect occurs, but not for which condition.

The Kruskal-Wallis analysis ranks the groups in Table 8.

	SIZE	N	Mean Rank
TALKING	1	5	6,40
	2	5	11,80
	3	5	5,80
	Total	15	
READING	1	5	9,80
	2	5	4,40
	3	5	9,80
	Total	15	

Table 8: Kruskal-Wallis ranking for TALKING and READING

The ranking shows, that the mean rank of the group *Talking* for tabletop size 2 is nearly twice as large as the mean ranks for tabletop sizes 1 and 3. The opposite effect can be seen for the group *Reading.* Other groups don't show this effect and are therefore not shown in Table 8. As further step, it has to be checked whether this effect is significant. The calculation can be seen in Table 9.

	TALKING	STRUC-TURING	READING	SHOW AND TELL	INFO	OTHER
Chi-Square	5,460	,420	4,860	,540	,140	2,034
df	2	2	2	2	2	2
Asymp. Sig.	**,065**	,811	**,088**	,763	,932	,362
Exact Sig.	**,060**	,832	**,085**	,794	,954	,725
Point Probability	,005	,025	,002	,012	,028	,220

Table 9: Kruskal-Wallis analysis for all groups on a basis of absolute values

The table shows no significant values for the groups *Structuring, Show and Tell, Info and Other.* As this analysis can be considered more strict in comparison to the Median Test, the significance values for *Talking* ($H(2) = 5,46$, $p > 0,05$) and *Reading* ($H(2) = 4,86$, $p > 0,05$) are higher and slightly above the significance level of 0,05. Thus, **H4** *cannot be statistically confirmed* by this comparison. But there is a tendency that tabletop size 2 groups spend more time *Talking* than both other sizes. In contrast to this, they spend less time *Reading* than both other sizes.

Talking in Detail

As a second step, each group is split up in order to investigate and compare the different coupling styles in detail. First, the group *Talking* is analyzed with the help of another Kruskal-Wallis analysis. The calculation is now based on relative values, as the mean *Talking* times of each condition and group differed. This allows to find differences in *Talking* behaviors and e.g. to see if any roles evolved.

	DELEGATION	DISCUSSION	EXPLAINING	PROCEDURE	QUESTION	RECAP	OTHER
Chi-Square	6,492	,080	,140	,215	,560	5,400	1,086
df	2	2	2	2	2	2	2
Asymp. Sig.	**,039**	,961	,932	,898	,756	,067	,581
Exact Sig.	**,032**	,968	,954	,909	,783	,057	1,000
Point Probability	,001	,014	,028	,008	,023	,002	,714
Mean Size 1	4,11%	48,76%	41,09%	1,55%	3,64%	0,68%	0,17%
Mean Size 2	1%	49,92%	40,72%	1,83%	3,04%	3,34%	0,15%
Mean Size 3	0,27%	49,37%	42,21%	1,92%	4,89%	1,35%	0,00%

Table 10: Kruskal-Wallis Analysis for talking coupling styles on a basis of relative values

Table 10 shows a significant value for H(2) = 6,492 with $p < 0,05$ for *Delegation* activities on a basis of relative values. The absolute mean values for delegation activities are 59,28s (SD = 78,57, 4,11% of average total *Talking* time) for size 1, 25,53s (SD = 34,64, 1% of average total *Talking* time) for size 2 and 6,18s (SD = 8,85, 0,27% of average total *Talking* time) for size 3. As the Kruskal-Wallis analysis does not show to which condition this significance is assigned to, a post-hoc test has to be done. Therefore, a Mann-Whitney Test is used to compare the three conditions pairwise. This pairwise testing can result in an inflation of the error rate (Field 2005). To avoid this, a Bonferroni correction is applied, which changes the level of significance to 0.0167 (0,05 / number of pairwise tests (3)).

The Mann-Whitney Test does not show any significant values for the pairwise comparison of sizes 1 and 2 ($p > 0,05$) nor of sizes 2 and 3 ($p > 0,05$). The pairwise comparison of sizes 1 and 3 shows a significant difference with **p = 0,008**. This enables to calculate an effect size, which is **r = -0,84**. Groups that worked with tabletop size 1 did significantly delegate more than groups with size 3. Tendentially – but not significantly – size 2 ranges in the middle of them. This speaks *in favor of* **H4**. As delegation can be seen as a way to express role behavior, this *supports* **H2**. Other coupling styles concerning the group of *Talking* show no significant and mostly just minor differences. The mean relative values for *Discussion* are 48,76% for size 1, 49,92% for size 2 and 49,37% for size 3. The mean relative values for *Explaining* are 41,09%, 40,72% and 42,21%. The mean relative values for *Procedure* are 1,55%, 1,83% and 1,92%. The mean relative values for *Question* are 3,64%, 3,04% and 4,89%. The main difference can be found for *Recap:* The mean relative values are 0,68%, 3,34% and 1,35%. A Kruskal-Wallis analysis shows no significant values for *Recap,* but the tendency lets assume that groups working on size 2 tend to recap more than groups working on both other sizes. This can also be seen as all five groups working on size 2 recapitulated whereas only two groups working on size 1 and one group working on size 3 did.

<u>Structuring in Detail</u>

Structuring can happen as a group or individual. The absolute mean values for group-based structuring are: 70,86s (SD = 39,66; 17,15% of average structuring time) for size 1, 121,55s (SD = 93,12; 22,33% of average structuring time) for size 2 and 116,22s (SD = 123,96; 16,91% of average structuring time) for size 3. Individual structuring happened more often (comparing mean values): 356,23s (SD = 113,1; 82,85% of average structuring time) for size 1, 434,99s (SD = 222,48; 77,67% of average structuring time) and 447,27s (SD = 238,64; 83,09% of average structuring time) for size 3. It can be seen, that *Structuring* tendentially happened more on sizes 2 and 3, while size 2 shows more group-structuring activities. Figure 33 shows the relative proportions of group vs. individual structuring activities.

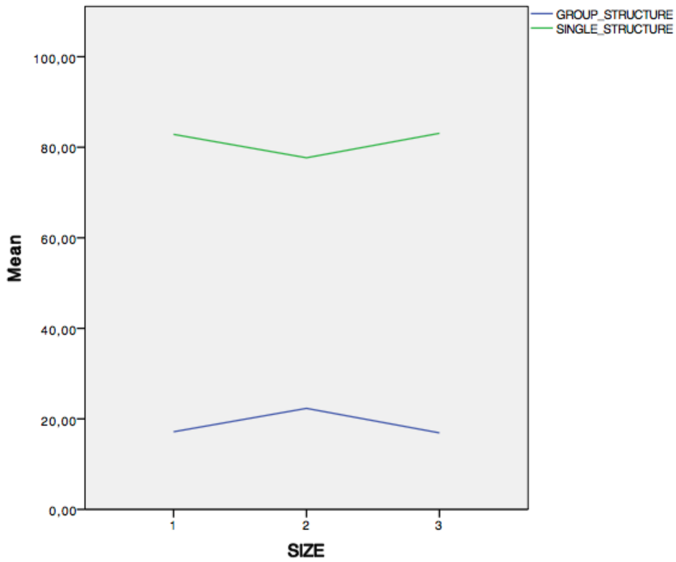


Figure 33: Group vs. individual structuring (y-axis: relative mean values)

A Kruskal-Wallis analysis has shown no significant values neither for group nor for individual structuring activities on a basis of relative values (both: $H(2) = 0{,}366$, $p > 0{,}05$). Thus, **H1, H2 and H4** *cannot be statistically confirmed* by the group *Structuring.* Nevertheless, the tendency shows more structuring activities for tabletop sizes 2 and 3 with more group-based structuring activities for size 2.

Reading in Detail

Reading covers all activities where both group members read articles on their personal devices. These articles can either be from the same or from different clusters. The latter can be split up to different clusters with the same problem and different clusters with different problems. For size 1, participants spend the mean time reading 1273,84s (SD = 1011,11; 48,25% of average reading time) articles from the same cluster, 153,96s (SD = 243,2; 4,81% of average reading time) articles from different clusters with the same problem and 1365,02s (SD = 1039,33; 46,94% of average reading time) articles from different clusters with different problems. For size 2, participants spend the mean time reading 465,46s (SD = 741,91; 25,70% of average reading time) articles from the same cluster, 56,02s (SD = 66,22; 2,79% of average reading time) articles from different clusters with the same problem and 1483,58s (SD = 955,06; 71,51% of average reading time) articles from different clusters with different problems. For size 3, participants spend the mean time reading 759,4s (SD = 671.86; 26,97% of average reading time) articles from the same cluster, 112,47 (SD = 251.49; 3,96% of average reading time) articles from different clusters with the same problem and 1753,57s (SD = 945, 69,07% of average reading time) articles from different clusters with different problems. As shown before, groups working on tabletop sizes 1 and 3 spend more time reading articles than on size 2. Figure

34 shows the relative reading behaviors. It can be seen, though the absolute times from all sizes differ, that the values for sizes 2 and 3 are relatively similar.

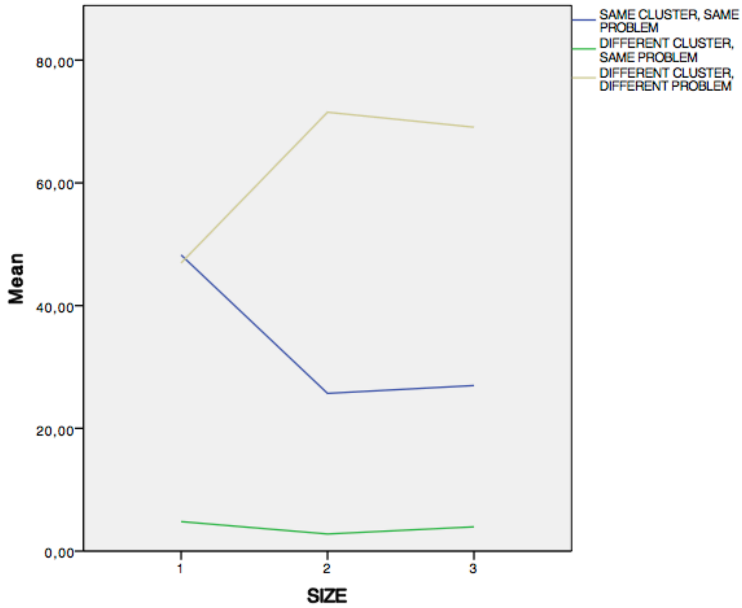


Figure 34: Relative reading behavior (y-axis: relative mean values)

A Kruskal-Wallis analysis shows no significant values: $H(2) = 2{,}560$, $p > 0{,}05$ for different cluster with different problems; $H(2) = 1{,}410$, $p > 0{,}05$ for same cluster with same problem and $H(2) = 0{,}484$, $p > 0{,}05$ for different clusters with same problem. Thus, **H4** *cannot be statistically confirmed* by this. The reason for the similar values for size 1 regarding two different cluster types can be found in the data logging: groups working with tabletop size 1 usually worked on the basis of 1 or 2 clusters in total, whereas groups working on the other sizes used up to 10 clusters.

Show and Tell in Detail

Show and Tell describes all activities where both group members e.g. look at the same result or show an article on their personal device to their group member. Participants spend 284,62s (SD = 143,89) in mean for size 1 in *Show and Tell*. Whereas in size 2 they spend 314,11s (SD = 199,23) and in size 3 219,24s (SD = 122,27). The main differences can be seen for activities like *Sharing of View* and *View Engaged.* Participants spend on average 11,2s (SD = 23,45) with size 1 in sharing their view and by this their personal device, 36,72s (SD = 62,91) with size 2 and 12,79s (SD = 18,21) with size 3. For *View Engaged*, they spend 63,77s (SD = 61,81) for size 1, 38,95s (SD = 49,34) for size 2 and 27,96s (SD = 45,31) for size 3.

A Kruskal-Wallis analysis reveals no significance:

- Combination (H(2) = 2, p > 0,05)
- Sharing of Same Information (H(2) = 0,27, p > 0,05)
- Sharing of View (H(2) = 0,867, p > 0,05)
- View Engaged (H(2) = 1,765, p > 0,05)

Thus, **H4** *cannot be statistically confirmed* by this. But groups working on size 1 spend tendentially more time looking at their group member's activity while groups working on size 2 spend more time showing their personal device to their group member.

As the groups *Info* and *Other* do not show relevant aspects, they are excluded in this analysis. As Isenberg et al. (Isenberg et al. 2010) compared group activities based on the idea of *closely* and *loosely* coupled collaboration styles, the different coupling styles are grouped into these two aspects and analyzed:

Closely and Loosely Coupled

The different coupling styles are grouped into closely and loosely coupled as shown in Table 6. As coded events like *Irrelevant Procedure* or *Info* cannot be assigned to one of those aspects, they are listed as *Other.* A Kruskal-Wallis analysis is performed and shows no significant values:

- Closely coupled (H(2) = 1,5, p > 0,05)
- Loosely coupled (H(2) = 0,62, p > 0,05)
- Other (H(2) = 0,14, p > 0,05)

Thus, **H4** *cannot be statistically confirmed* by this.

The mean times for working in closely and loosely coupled collaboration are 3284,58s (SD = 849,26) closely and 1721,25s (SD = 960,47) loosely coupled collaboration for size 1; 3194,72s (SD = 1153,22) closely and 1941,09s (SD = 1104,86) loosely coupled collaboration for size 2 and 2720,72s (SD = 1239,1) closely and 2200,84s (SD = 1056,4) loosely coupled collaboration for size 3. This results in closely-loosely coupled ratios of about 2:1 for size 1, 5:3 for size 2 and 5:4 for size 3.

Timeline

The previous analysis has focused on time-based activities and compared absolute and relative times to get insights on how participants communicate and collaborate. In addition to this, factors like the chronology of events or alternating intervals of coupling styles can be analyzed. This allows to gain insights e.g. on how different groups of coupling styles are spread along the duration of the task. A traditional way to visualize such diagrams are Gantt charts. As neither SPSS, nor R[14], nor Microsoft Excel are able to create Gantt charts with reoccurring event types, the visualization was chosen to be self-implemented with the help of HTML, JavaScript, CSS and D3.js[15]. The goal was to have all 15 groups in one diagram to be able to compare them easily. To achieve this, the data had to be structured in a scheme that was pre-defined by the used D3-Gantt-Chart library[16]. The library needs a start and end time, a task name and a status. Task names are assigned to the y-axis while each status is mapped on the x-axis with a certain width that symbolizes its duration calculated by its start and end time and color that shows its activity. The different coupling styles logged via Noldus Observer represent the category *status*, the *task names* are the names of each group and the *start time* is the certain time each behavior starts. The *end time* can be calculated by summing up the start time and given duration. As Noldus Observer logs two lines per event symbolizing its start with start time and duration and its end with no duration and end time, the so-called *stop events* had to be filtered. Then all events were restructured to the new scheme; all separate logs of all groups were combined to one single Excel sheet (> 4300 events in total) and saved as a CSV file.

The given D3 library had to be edited to enable the processing of the external CSV file. Additionally, the various coupling styles had to be set up as *status* and the different groups were assigned as *task names*. Each status had its own CSS class name, which allowed to define e.g. background-colors for each type of coupling style like *Discussion* or *Explaining* or to each group like *Talking* or *Reading* to get various insights. The result was the **TimeVis** web app that can be found on the attached USB flash drive of the print version of this thesis. It is best-viewed in Apple's Safari browser.

The entire visualization in larger dimensions can be found folded as a map attached to the last page of the print version. In this chapter, the visualization is adapted to find differences and features of the named groups of coupling styles in Table 6. Therefore, multiple views of the visualization are generated to be able to focus on single aspects of group work activities.

---

[14] R is a tool for Statistical Computing: www.r-project.org

[15] D3.js is a JavaScript library for manipulating documents based on data: www.d3js.org

[16] A Gantt-Chart library based on D3.js: www.github.com/dk8996/Gantt-Chart

Timeline – Talking

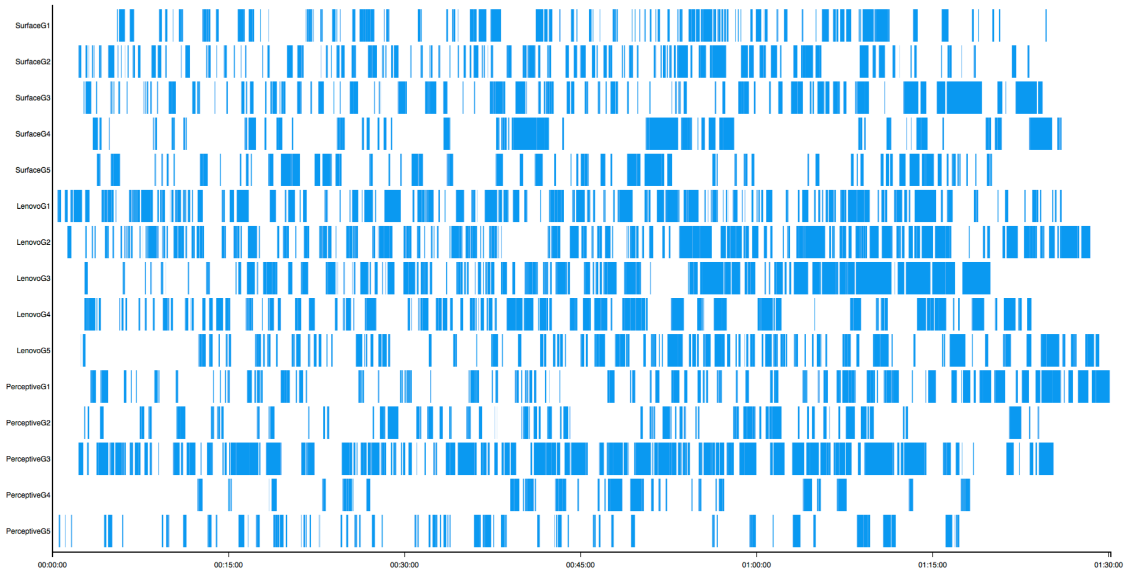


Figure 35: Timeline – Talking (y-axis: group / x-axis: time)

Figure 35 shows the *Talking* sequences of all 15 groups. The top five rows stand for tabletop size 1, the next five rows for size 2 and the last for size 3. The sequences are mapped to their duration on the x-axis. The analysis of *Talking* times has shown that groups working with size 2 tend to talk more than groups of both other conditions. This impression can be confirmed by Figure 35. The *Talking* sequences of size 2 show a higher density and also longer intervals compared to sizes 1 and 3. The only non-size 2 group that can be compared to the *Talking* behavior of size 2, is group 3 of size 3. This group shows constant *Talking* activities, while other groups of this size do not. Therefore, it can be seen as an outlier. In accordance with the observation protocol, groups working with size 2 had various durations of getting to know their group member and thus to start talking with each other. The first two groups and group 4 of size 2 instantly started *Talking* while groups 3 and 5 at first needed some time to get to know each other before they really started to talk to each other for longer durations. This effect cannot be seen for both other conditions, as their conversations in general are shorter (except group 3 of size 3). It can be seen that the density and duration of *Talking* activities increases in the second half of the task completion time for all three conditions. Especially groups working on size 2 started talking for intervals of several minutes at once. The same effect can be seen for most of the other groups, although less obvious. This *supports* **H4**.

Timeline – Reading

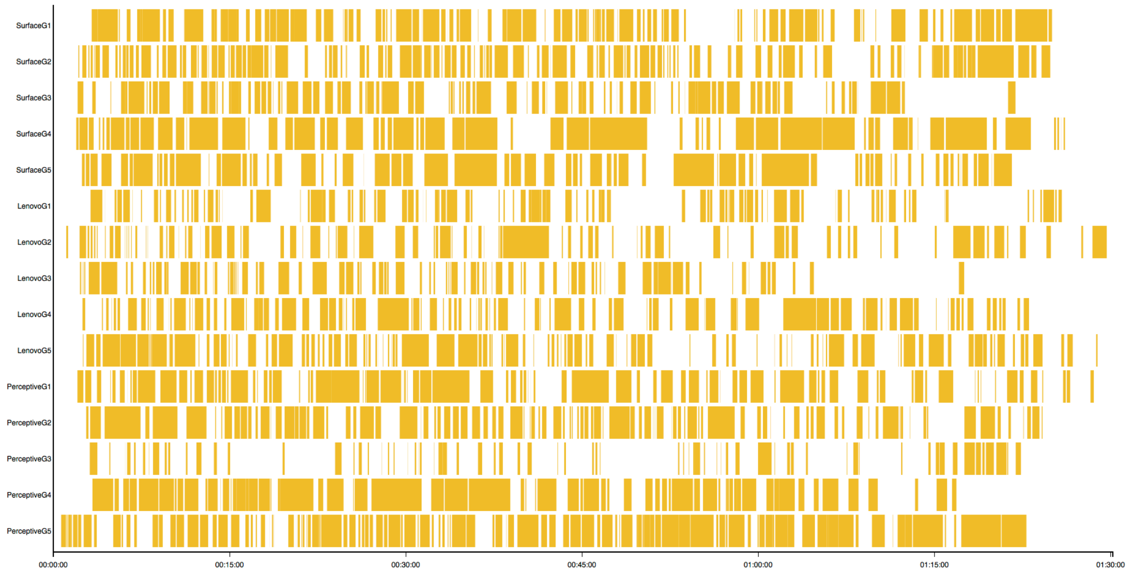


Figure 36: Timeline – Reading (y-axis: group / x-axis: time)

Figure 36 shows the *Reading* sequences of all 15 groups. The analysis of *Reading* times has shown that groups working with size 1 and 3 tend to spend more time reading than groups working on size 2. This impression can be confirmed by Figure 36. The *Reading* intervals of groups working on sizes 1 and 3 show a higher density and longer sequences of reading articles. Group 3 of size 3 is an outlier in this case, too. It can be seen that the density of *Reading* activities decreases in the last third for four groups of size 2. This *supports* **H4**.

Timeline – Show and Tell

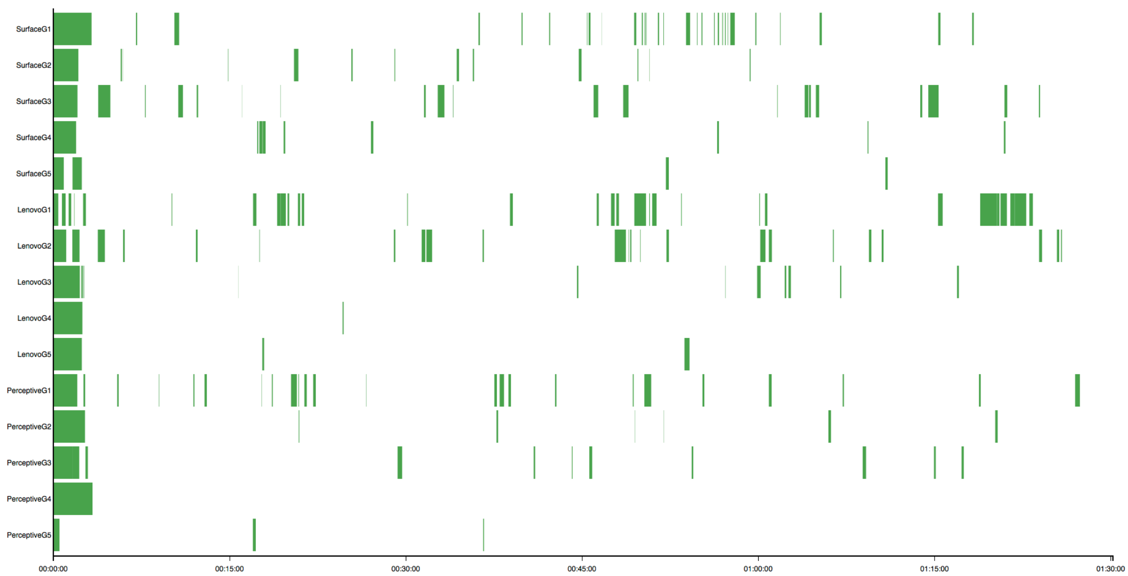


Figure 37: Timeline - Show and Tell (y-axis: group / x-axis: time)

Figure 37 shows the *Show and Tell* sequences of all 15 groups. All groups start their task by reading the introductory letter on their personal devices. As both group members focus on the exact same information, comparable to looking at the same result, this coupling style is assigned to this group and therefore the beginnings of all groups are marked by this. Most groups spend similar times on reading the introductory letter. Groups working on sizes 1 and 2 show a slightly higher tendency for this group of coupling styles. This *supports* **H4**. Especially groups working on size 2 show several phases with longer durations. Group 1 of size 2 shows a long duration of *Show and Tell* activities in the last quarter. The observation protocol references this as the coupling style *Combination* where one participant opened a city map on the personal device while the other participant opened an article describing various localities. The group combined the information shown on each personal device to get new insights. No other group collaborated this way.

Timeline – Structuring

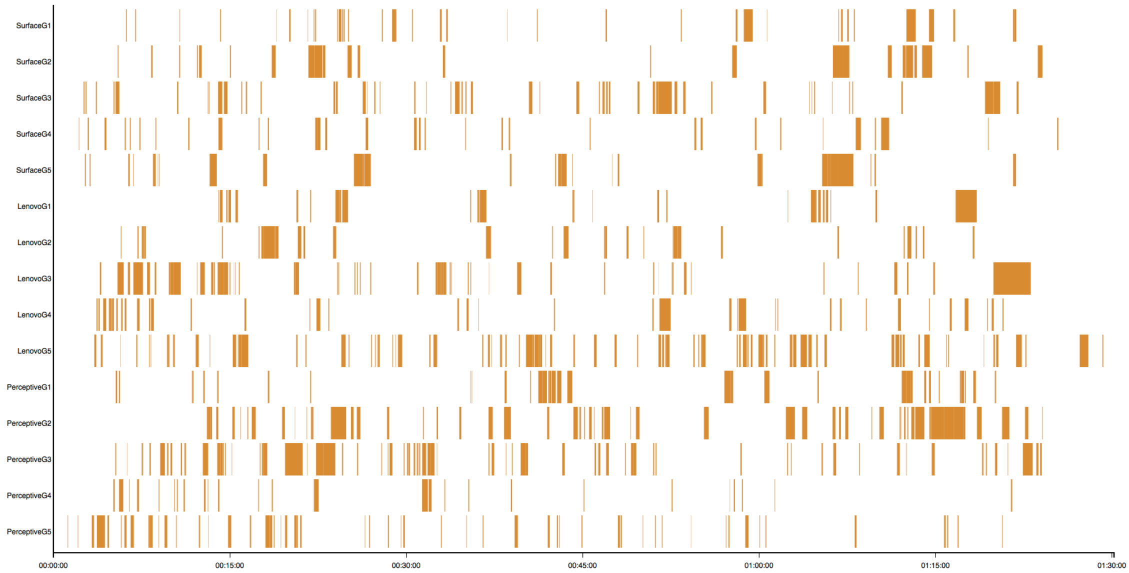


Figure 38: Timeline – Structuring (y-axis: group / x-axis: time)

Figure 38 shows the *Structuring* activities for all 15 groups. The visualization shows no patterns or unequal distributions of *Structuring* times. Groups working on sizes 2 and 3 show a higher tendency for *Structuring* activities. It can be seen that these activities increase during the last third or last quarter of the task completion time. This is recognizable for all conditions. As groups working on size 1 tend to structure less of their time, this speaks *in favor* of **H4**.

Timeline – Closely & Loosely Coupled Collaboration

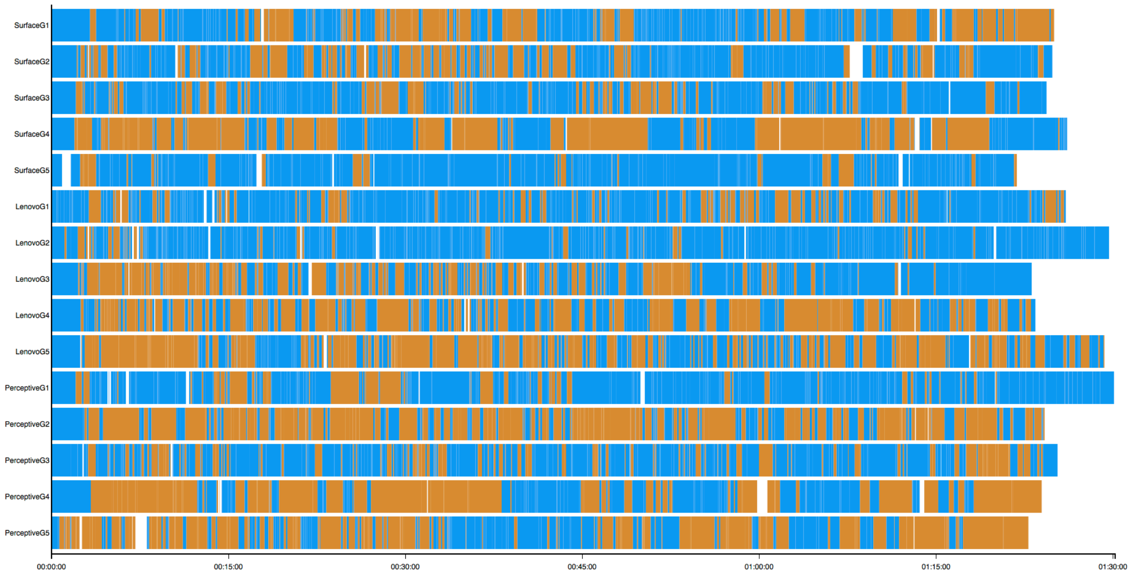


Figure 39: Timeline - Closely & Loosely coupled collaboration (y-axis: group / x-axis: time)

Figure 39 shows the visualization of closely and loosely coupled collaboration activities for all 15 groups. Closely coupled activities are mapped to blue, whereas loosely coupled activities are mapped to orange. The visualization shows no obvious patterns and for all three sizes there are groups dominated by either closely or loosely coupled collaboration. Anyway, it shows that groups evolve a group-specific mixture of both ways of collaboration and stick to it. For example the first two groups working on size 2 collaborate closely most of their time with evenly distributed phases of loosely coupled collaboration. Groups working on size 1 (except group 5) have a similar ratio of closely and loosely coupled collaboration whereas groups working on size 2 and 3 evolve either a closely-dominant mixture (2 groups each) or a mixture with long sequences of loosely coupled collaboration.

Timeline – General

The visualization of all activities (see folded map in the back of the print version) shows that the structures of collaboration and communication behaviors for sizes 1 and 2 are similar within their condition. This means, that all groups working on size 1 have similar structures with an emphasis on *Reading* while groups working on size 2 have similar structures with an emphasis on *Talking.* Groups working on size 3 differ. The first two groups can be compared to groups working on size 1 regarding *Reading* and *Talking* behaviors and to groups working on size 2 regarding *Structuring* activities. Group 3 attracts attention due to its very long *Talking* intervals and a total *Talking* time of nearly 50 minutes (no other group talked more). Group 5 can be seen as a total contrast with *Reading* times of more than 57 minutes in total. Group 4 shows also a long period of *Reading* with more than 50 minutes. Additionally, this group shows intervals with a total duration of more than 12 minutes where one group member was disengaged from group work activities. This *supports* **H4**.

5.4.3 Participants load

The mean results with their standard deviation of the NASA TLX (Hart & Staveland 1988) are shown in Table 11.

	Size 1	**Size 2**	**Size 3**
Mental Demands	79 (SD = 15,71)	71,5 (SD = 15,47)	69,5 (SD = 15,71)
Physical Demands	30,5 (SD = 18,77)	27,5 (SD = 26,8)	18,5 (SD = 18,5)
Temporal Demands	67 (SD = 16,87)	55,5 (SD = 21,14)	55 (SD = 19,72)
Own Performance	63,5 (SD = 23,34)	64,5 (SD = 15,54)	59,5 (SD = 22,42)
Effort	68,5 (SD = 21,74)	69 (SD = 13,7)	62 (SD = 13,78)
Frustration	63,5 (SD = 27,49)	54,5 (SD = 23,97)	48 (SD = 24,74)
Total	*372 (SD = 60,75)*	*342,5 (SD = 55,49)*	*312,5 (SD = 22,14)*

Table 11: Results of NASA TLX

A Kruskal-Wallis analysis shows no significant values for none of the six individual subscales (Mental demands: H(2) = 2,075, p = 0,354; Physical Demands: H(2) = 2,138, p = 0,343; Temporal Demands: H(2) = 2,624, p = 0,269; Own Performance: H(2) = 0,205, p = 0,902; Effort: H(2) = 1,14, p = 0,566; Frustration: H(2) = 2,243, p = 0,326). Using the Kruskal-Wallis analysis for the sums shows no significance, too (H(2) = 5,512, p = 0,064). In this study, the NASA TLX is only used for complementary inquiries; therefore no hypothesis is tested by it. Anyway, there is no significant difference between the three table sizes. The tendency shows that the larger the size, the smaller are the individual values of the subscales and in total.

5.4.4 Teamwork Questionnaire

To analyze the data of a Likert scale as it is used in the teamwork questionnaire the mode or median value has to be checked. The mean value could falsify the results as its ordinal structure shows an inherent order but one cannot assume that each value has the same distance.

A Kruskal-Wallis analysis shows no significant difference between the three sizes: $H(2) = 2{,}0$, $p > 0{,}05$ for all statements. Thus, **H1** *cannot be statistically confirmed* by this. The teamwork questionnaire can be found in the appendix. The following table shows the different modal values for overview reasons (1: strongly agree – 7: strongly disagree):

Statement #	Size 1	Size 2	Size 3
1	2	1	1
2	3	1	3
3	7	5	5
4	1	1	1
5	2	1	1
6	1	1	1
7	2	1	1
8	1	1	1
9	6	2	2
10	2	1	1
11	1	2	3
12	4	7	7
13	2	3	2
14	6	6	6
15	2	2	2
16	1	1	1

Table 12: Mode results of the teamwork questionnaire

Table 12 shows that there are only minor differences between statements regarding the three sizes. Tendentially, the main differences can be seen for statements 9 and 12 ("I'm satisfied with our results" / "Coordination was hard"). Participants working on size 1 *disagree* while participants working on sizes 2 and 3 *agree* to first statement. The latter *totally disagree* to the coordination statement while participants of size 1 are undecided. Participants of all three conditions *strongly* agree that the size of the tabletop is adequate. Another difference can be seen for statement 2 ("The group was working very collaborative"). Participants working on size 2 *totally agree* whereas groups working on both other sizes *rather agree.*

5.4.5 Roles

Participants described their subjective estimation of roles in the post questionnaire. 9 of 10 participants working on size 3 answered the question for roles with *No*. They name reasons like “we searched for information at the same time” (Original statement to bypass possible translation errors: “Wir haben gleichzeitig nach Informationen gesucht”), “we searched equally for solutions” (“Wir haben gleichberechtigt nach Lösungen gesucht”) or “We both had the same role to search, to categorize and to investigate, etc. This allowed us to have two perspectives on one problem” (“Wir hatten beide die Rolle zu suchen, zu kategorisieren und zu untersuchen, etc. Das ermöglichte es uns zwei Sichten auf ein Problem haben zu können.”). The one participant who answered the question for roles with *Yes* described the different topical foci of the group members.

7 of 10 participants working on size 2 answered the question for roles with *No.* They name reasons like “Both searched equally for solutions” (“Beide haben gleichermaßen nach Lösungen gesucht”) or “we divided the labor and worked with the same role” (“Wir haben die Arbeit aufgeteilt und mit der gleichen Rolle gearbeitet”). The three participants who answered the question for roles with *Yes* describe their own roles as “Distribution of tasks” (“Verteilung von Aufgaben”), “Connecting facts” (“Verbinden von Fakten”) or “I felt that I rather took the initiative” (“Hatte das Gefühl mehr Inititative an mich gerissen zu haben”) and their partners’ roles as “rather executive” (“eher ausführend”), “gain new insights” (“neue Erkenntnisse gewinnen”) or “Wing-Man” (“Wing-Man”).

6 of 10 participants working on size 1 answered the question for roles with *No.* They name reasons like “We worked together” (“Wir haben zusammen gearbeitet”), “We had equal rights” (“Wir waren gleichberechtigt”) or “We found the solutions at the same time” (“Wir haben die Lösungen gleichzeitig gefunden”). Participants who answered the question for roles with *Yes* describe their own roles as “rather dominant, rather active” (“eher dominant, eher aktiv”), “searching and reading” (“suchen und lesen”) or “boss” (“Chef”) and their partners’ respective roles as “rather passive” (“eher passiv”), “interpreting articles and looking for solutions” (“Artikel interpretieren und nach Lösungen suchen”) and “assistant” (“Assistent”).

A Kruskal-Wallis analysis is performed to see if the subjective estimations are significant. The test shows no significance with $H(2) = 3{,}683$, $p = 0{,}159$ for *Yes* and $H(2) = 2{,}762$, $p = 0{,}251$ for *No.* Therefore, **H2** *cannot be significantly confirmed* by this. Anyway, the tendency is that the smaller the tabletop, the more estimate participants to evolve different roles.

5.4.6 Procedure

The description of each participant's subjective estimation of the procedure is used to find abstract patterns and differences how they proceed during the task's solution. Therefore, their description is analyzed for buzzwords that allow to assign an abstract type of procedure. For example "we used a spiral way to find solutions" points to the abstract type "Spiral". The different types of procedure and their concerning groups can be found in Table 13.

Size 1	**Size 2**	**Size 3**
Divide and Conquer	Spiral	Divide and Conquer
Divide and Conquer	Spiral	Divide and Conquer
Repetitive	Divide and Conquer	Divide and Conquer
Repetitive	Spiral	Divide and Conquer
Spiral	Chaotic	Divide and Conquer
Repetitive	Divide and Conquer	Spiral
Spiral	Spiral	Divide and Conquer
Divide and Conquer	Divide and Conquer	Repetitive
Spiral	Divide and Conquer	Spiral
Chaotic	Divide and Conquer	Divide and Conquer

Table 13: Abstract types of procedure

In total there are four different types of procedure found in the analysis of procedure types:

- **Chaotic:** No structure
- **Divide and Conquer:** Dividing the task into subtasks
- **Repetitive:** Repetitive searching for keywords concerning a topic
- **Spiral:** Successive narrowing of the problem space

A Kruskal-Wallis analysis shows no significance ($H(2) = 2{,}686$, $p = 0{,}261$). Thus, **H4** *cannot be statistically confirmed* by this. Though, participants tend to use the procedure of *Divide and Conquer* preferable on larger sizes.

5.4.7 Participants' findings

The solution of the Stegosaurus VAST Challenge cannot be distributed. Therefore, it is not allowed to show findings or the sample solution. The sample solution is divided into 20 relevant aspects that are checked for each participant's solution. The relevant aspects consist of names of significant persons, their possible relations and important events that occurred. Thus, each participant can get a maximum score of 20 points. The more the solutions of each group member are similar to the solutions of the group partner, the more likely it is that both group members have the same (shared) mental model of their task and solution. Therefore, a second score is used to show the number of relevant aspects found by both group members. Thus, the group score cannot be higher as the lower score of both group members.

A comparison of the mean values of the three conditions shows individual scores of 2 (SD = 1,89) and group scores of 0,8 (SD = 0,79) for size 1, individual scores of 3,3 (SD = 2,16) and group scores of 2,6 (SD = 1,71) for size 2 and individual scores of 3,4 (SD = 2,76) and group scores of 2,6 (SD = 2,27) for size 3. There is an obvious difference between the findings. Groups working on sizes 2 and 3 found a similar amount of aspects individually and group-based whereas groups working on size 1 found less for both scores. A Kruskal-Wallis analysis is performed to check if this difference is significant. The analysis shows no significance for the individual score (H(2) = 2,224, p = 0,329) but for the group score (H(2) = 6,4, p = 0,041). Three Mann-Whitney tests (with Bonferroni correction to change the significance level to 0,0167) are used to find significant values for the pairwise comparison of the three sizes regarding the group score. The comparison of sizes 1 and 2 shows a significant value with **p = 0,013**. This allows to calculate an effect size, which is **r = -0,56**. The comparisons of sizes 1 and 3 as well as 2 and 3 do not show significant values (p = 0,061 and p = 0,878). Thus, groups working on size 2 have a significant higher group score than groups working on size 1. As the group score is the equivalent to the task solution and thus to the shared mental model of group activities, **H3** can be *statistically confirmed* for those two sizes.

5.4.8 Usage of the shared work space

Participants are asked to describe their usage of the shared work space in their own words. Representative statements for each condition are shown in Table 14.

Size 1	„Rather less, because I had all keywords on the iPad. But it was useful to create a story of the results at the end." (Original: "Eher weniger, weil ich auf dem iPad alle Stichwörter hatte. Zum Schluss um die Resultate zusammen zu basteln zu einer Geschichte war es aber ziemlich nützlich") „I hardly used the interactive table. I just used it to save articles to share them with my partner. My partner used it more often (for sharing, too)" („Den interaktiven Tisch habe ich wenig genutzt. Ich habe diese nur genutzt, um Artikel zu speichern, damit mein Partner diese Artikel auch sehen kann. Von meinem Partner wurde dieser häufiger (auch zum Teilen) genutzt.")
Size 2	„Arranging of words and articles, connecting coherent keywords, saving of notes, exchange with patner, sorting, fast re-finding of articles, fast connecting and new distribution" („Anordnen der Wörter und Artikel, Zusammenfügen von zusammengehörigen Stichwörtern, Abspeichern von Notizen, Austausch mit dem Partner, Sortieren, schnelles wiederfinden von Artikeln, schnelles zusammenfügen und neu verteilen ") „As a kind of whiteboard, which i used to try sorting relevant keywords. As a platform to present solutions to the team partner or to allow him to start searches. As starting point for investigations." („Als eine Art Whiteboard, auf dem ich versucht hatte, alle relevanten Begriffe zu sortieren, Als Plattform, dem Teampartner die Lösungen zu präsentieren bzw ihm seine Recherche zugänglich zu machen, Als Ausgangspunkt für Nachforschungen.")
Size 3	„The table was used to save keywords and to create cluster (Mindmaps) -> assigning keywords, too. Additionally, it was used as an overview to see what the group partner was searching and collecting -> better overview -> no coincident search / intersection. Snippets of articles could be accessed faster." („Der Tisch wurde eingesetzt zum Speichern der Stichwörter und Cluster (Mindmaps) anzufertigen. -> auch zuordnen der Begriffe. Zudem bot er Übersicht, was der Partner für Stichworte sammelte und suchte. -> bessere Übersicht -> keine gleichzeitge Suche / Überschneidung. Außerdem konnte dort auch schnell auf einen Artikelausschnitt zugegriffen werden.") "The table enabled to sort and arrange articles and keywords. The navigation, which article has to be read next happened on the table. Coherent structures could be easily combined. Most of the time all keywords and articles were easy to find, even better than on a real table." ("Auf dem Tisch konnte man gut Artikel und Keywords ordnen und ausrichten. Die Navigation, welchen Artikel man als nächstes lesen möchte, erfolgte auch auf dem Tisch. Inhaltlich sinnvoll zusammenpassende Strukturen konnte man auch gut mit dem Tische zusammenfassen. Meistens waren alle Keywords und Artikel auch gut wiederzufinden, was bei einem echten Tische manchmal nicht so leicht ist."

Table 14: Statements on the usage of the shared work space

It can be seen that the impression of the usage of the shared work space differed. It cannot be statistically evaluated since the comments differ too much. Representative screenshots of the final outcome of each tabletop size are shown in the following to gain insights on the different usages.

Figure 40: Outcome of a group working on size 1

Figure 40 shows that in this case size 1 was rather used for deposition than for structuring activities. There are many (double) and overlapping keywords, clusters and snippets. This group shows a higher number of clusters than other groups working on size 1. Most of these groups used one or two clusters. These clusters show less structure as groups working on both other sizes.

Figure 41: Outcome of a group working on size 2

Figure 41 shows that in this case size 2 was used to structure keywords and snippets. The clusters partially overlap and thereby show connections and intersections. Groups working on size 2 structured their content more actively and divided the task thereby into subtasks.

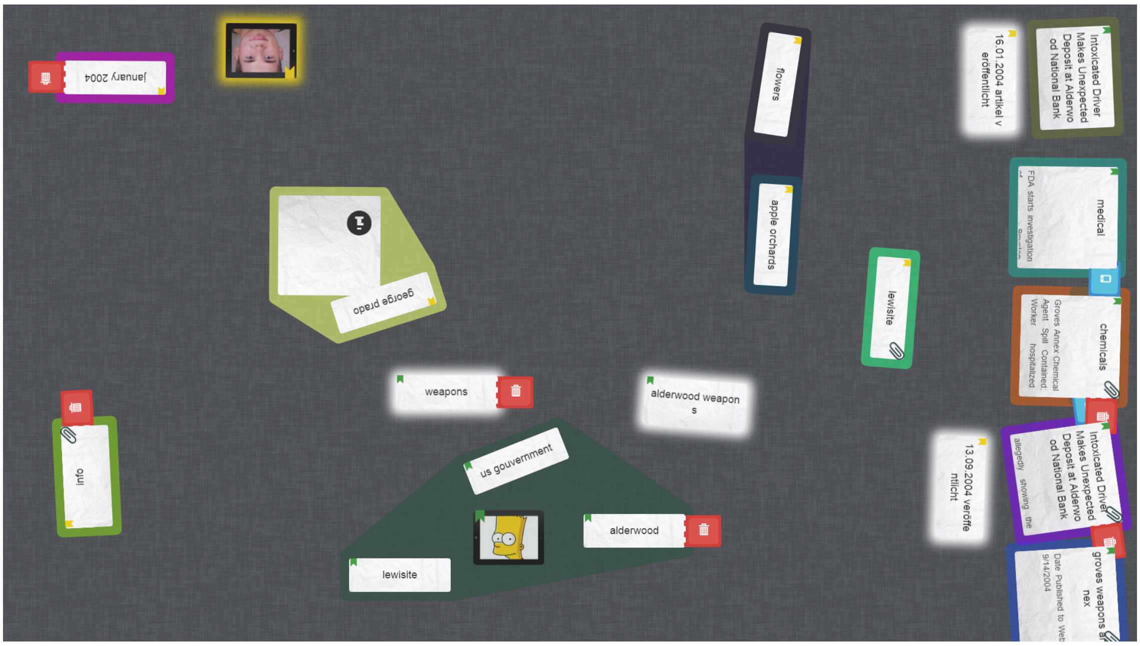


Figure 42: Outcome of a group working on size 3

Figure 42 shows that in this case the group used the table to structure their content in various ways: on the right hand side, they aligned different events in a timeline structure. On the left hand side, they placed general information and the middle part of the tabletop was used as their “chaos” and working space. Important snippets and keywords are rotated to allow both group members to read them. Groups working on size 3 structured their content mostly comparable to groups working on size 2.

5.4.9 Participants’ general feedback

During and after completing the task multiple participants stated ideas and gave feedback on different functionalities of the TwisterSearch system. These statements were noted in the observation protocol and can mainly be seen as improvements and further refinements of the TwisterSearch system. Thereby they are not size-specific.

“I would like to have the functionality to share an entire article to the shared group space by pressing just one button.”

“Can I send the article directly to my group partner’s iPad?”

“It would be cool to change the ranking of search results when searching for events.”

“Too bad the notes I added to the snippet are not shown on the table.”

“It might be useful to have a handwritten bullet list to check different topics.”

“Let’s define different zones on the table!”

The statements mainly focus on the personal device interface. Participants wish for an easy exchange of entire articles in addition to snippets of them with their group partner. For time-specific searches they want to have the ability to change the representation of search results e.g. to find oldest or newest articles. Multiple participants mentioned that it might be beneficial to have the possibility to note important findings or to check assumptions. These ideas and statements are valuable for possible future refinements of the TwisterSearch system.

5.5 Discussion

The shown results of the study are summarized, critically examined and discussed in this chapter. The goal is to comment on new insights with respect to the research question in Chapter 5.

The study results show that tabletop sizes influence groups’ communication and collaboration during collaborative search activities in various ways. Video analyses allow to get deep insights in communication and collaboration behaviors. The results of the grouped coupling styles show differences regarding *Talking* and *Reading* times for the three different sizes. While these differences are not significant, there is still a tendential difference. **Groups working on tabletop size 2 talk more** than groups working on the other two sizes. In contrast to this, the times for *Reading* activities are contrary to this. **Groups of sizes 1 and 3 spend more time reading**. There is a connection between these two groups of communication and collaboration. As the task completion time is limited, groups who talk more have less time to use for reading activities and vice versa. Thus, the question is: does the size of the tabletop rather affect *Talking* or *Reading* activities?

The interactive tabletop can be seen as the visual foundation of groups' communication, shared source of information and thus a connection between the two group members during their task. That suggests that tabletop size affects *Talking* activities. The different tabletops are placed on the same office desk to have the same physical distance between participants of all three conditions. In contrast to this, the physical distance between participants and their assigned interactive tabletop differs. Thus, size 1 is further away from the participants than sizes 2 and 3. The distance to the latter is the smallest. Additionally, the item-landscape ratio is constant for all three sizes. As the physical sizes differ, so do the digital sizes. Digital objects on size 1 are physically smaller than on size 2 and on size 3. The perceived sizes of all items support the physical distances as elements of smaller sizes tend to be perceived with a greater distance (Goldstein 2014). Thus, the smallest size has physically and perceptive the furthest distance to the participants. Whereas size 3 has the smallest distance and size 2 resides in the middle. Edward T. Hall (Hall 1990) shows different proxemic zones in which individuals communicate and collaborate with each other. They reach from an intimate to a public level and categorize thereby with who and how individuals talk. As the interactive tabletop can be seen as the basis of groups' communication like described in the beginning of this paragraph, this effect could possibly be transferable. In this case, the distance to the tabletop symbolizes the perceived distance to the group partner. Thus, group partners working on size 1 perceive each other in a greater distance; get thereby assigned to a different proxemic zone and talk less to each other than groups working on size 2 that possibly perceive each other in a known social zone, where they are used to communicate with others. Groups of size 3 could possibly perceive each other at a very intimate level – a level at which they are not used to talk to group partners especially if they don't know each other. The results of the teamwork questionnaire support this theory for example as participants working on size *totally agree* that their group worked very collaboratively whereas both other rate this with *rather agree*. These ideas are speculative and need to be verified. Therefore, the setting of a future study can be varied either towards same physical sizes of digital elements on all three sizes, which results in differences in the usable space, the variation of the office desk where the tabletop is placed, which results in smaller physical distances between participants or with a remote scenario where group partners are placed in different rooms and work on synchronized tabletops. In addition to this, groups of participants could be selected by their relationships, as strangers, co-workers and friends each might work differently.

A more pragmatic view on this relation could be that participants perceive tabletop size 3 as a very massive device that takes up a lot space. Thereby it resides between the two participants and literally blocks their communication. The communication of groups working on size 1 could be inhibited by its size as it provides to less space to externalize each other's ideas that form the basis of communication.

The difference in *Talking* times between groups working on all three sizes might be connected to the size of the digital artifacts. As the item-landscape ratio is the same, the artifacts used on size 3 are physically larger than on size 2. Thus, the artifacts and thereby the user interface might be easier to perceive and read. In other words, groups working on size 3 might be better supported by the interface, work more effective and need to talk less to their group partners, as the cognitive load for both group partners is lower. Regarding this theory, groups working on size 2 compensate this shortcoming of efficiency by talking to their group partners to gain similar scores of results. Groups working on size 1 might talk less because the sizes of the artifacts might be too small to work effectively and perceive and thus they focus on *Reading* articles on their personal devices because they feel too less supported by the interface.

Structuring keywords and snippets on the interactive tabletop **support individuals and groups to find more relevant information**. Groups working on sizes 2 and 3 spend more time to *Structure* the content of the interactive tabletop. Both of them also show a higher score concerning their individual, and a significant higher score concerning their group findings. These two findings might be related. While keywords are used to externalize each participant's thoughts and ideas and thereby to explain them, structuring and clustering of them can build upon this to find relations, interconnections or differences. Kirsh (Kirsh 1995) names clustering as an activity that categorizes and thereby simplifies tasks. This allows participants to get better insights by themselves, and their partners can comprehend their ideas and solutions more easily. This allows groups to better understand the problem space of their task as it is more structured on the one hand and on the other hand they can comprehend and explain their ideas at a more abstract level of detail that allows to find e.g. relations between findings of both participants. The better ratio of the individual and group score of groups working on size 2 can be linked to more and longer recap activities, where both group members talked about their findings and relations between them.

The more information is found, the more groups structure and discuss them. The TimeVis app shows an increase in *Structuring* and *Talking* activities during the second half and last quarter of the task completion time especially for groups working on size 2. There might be two reasons for this effect. During the beginning and the middle parts of their task completion time participants gained information and insights on the topic by reading articles and talking in general to their group partners. At a certain point, they possibly got the feeling of being enough prepared and having enough information gathered. Then they started to discuss the topic with their group partners for longer durations – supported by structuring activities that helped them to focus on important aspects. This might lead to the assumption that groups need a certain amount of time to feel comfortable enough to start longer conversations about a specific topic. The second reason for this effect could be that groups spent time to forage a lot of information about the topic and at a certain point they had to tidy their cluttered shared work space, which resulted in structuring and talking activities, where they discussed which pieces of information are relevant or not. The information about the remaining time also forced both of the activities for some groups. They might have felt some pressure that led them to discuss and structure important aspects and filter irrelevant information.

Structuring simplifies perception. The benefits of *Structuring* activities can be found in other results, too. The results of the NASA TLX show higher values for the mental demand of groups working on size 1. They spend less time *Structuring* and use the interactive tabletop more for deposition reasons. This makes it harder to comprehend each other's progress, solution and ideas. Thus, the mental demand increases.

Tabletop size has an effect on groups' procedure. A connection can be found regarding the procedures of participants. Most of them work more in a *Divide and Conquer* manner on larger tabletops. This procedure is characterized by a division of labor on the one hand and on the other hand it assumes that different categories and subtasks can be build. This fits to their structuring-based approach.

Awareness and the shared mental model affect the behavior of group members. The significant lower group score and thus the weaker shared mental model of groups working on size 1 take effect in other results. The tendency shows more activities regarding the coupling style of *View Engaged.* It can be characterized as one actively reading participant and a participant who takes a look at the partner's activity with optional communication. As they are less aware about the activities of the group partner, they have to actively look at their partner's activity. The less awareness and knowledge about the partner's work and progress is also link to the *Delegation* coupling style, which is significant higher for groups working on size 1. They are less aware and there is less structure on the tabletop, which might result in participants that are more confident to evolve a more dominant behavior as they are aware of their own progress but not of their partner's.

A non-interactive edge of the table has influence on collaboration. In general there is a tendency of groups working on sizes 1 and 2 to spend more time on *Show and Tell* activities. This result can be explained by a pragmatic approach. Most of these activities involve physical efforts like leaning onto the table while showing the personal device to a group partner. In general, this is possible for all three tabletops but groups working on sizes 1 and 2 can lean onto the office desk underneath the interactive tabletop while as groups of size 3 need to lean onto the interactive tabletop itself. This might either be uncomfortable or participants might be afraid of damaging the device. Future studies could evaluate this idea. In combination with the shown idea of adapting the size of the office desk, where the interactive tabletop is placed on, there could be a predefined non-interactive edge with a fixed size for all tabletop sizes to allow participants to lean onto in the same way.

The mixture of closely and loosely coupled collaboration has an effect on groups' procedure and their outcome. The results show different ratios of closely-loosely coupled collaboration. Whereas groups working on size 1 tend to work more closely, the ratio of sizes 2 and 3 are more equal with an emphasis on closely coupled collaboration. Groups working on size 1 tend to work more closely but one reason for this might be that they prefer to have a strategy with only one or two clusters and thereby often search for same or similar keywords and articles. Whereas groups working on both other sizes show strategies of using multiple different clusters for various aspects and thereby search for different subtopics supported by a division of labor (compare the procedure of *Divide and Conquer*). Thus, the TwisterSearch application allows for different working styles without restricting the group members to a given procedure. The benefits of a more equal mixture of closely and loosely coupled collaboration can be found in the outcome of groups and individuals. Loosely coupled collaboration can e.g. be described by group members reading articles found by keywords in different clusters with different topics. Therefore, they search for information in different directions. They gain knowledge about their subtopic and switch to closely collaboration to discuss and explain their findings to their group partner, they connect to the same cluster to find similar information or they structure the content on the shared group space. This back-and-forth behavior in combination with their *Divide and Conquer* procedure stretches a web as each new piece of information could be possibly connected to others. Thus, each new insight leads to new assumptions or connections. Johnson (Johnson 2010) calls this the *adjacent possible*. As a result, they are able to find more information individually and collaboratively than groups working too closely or too loosely. A too closely approach might lead groups to search for information in only one direction. Therefore, they cannot find and connect that many aspects, as both concentrate on the same problem. If the groups pursue a wrong path, then both group members search for irrelevant information. Groups working too loosely might be able to find a lot of information individually, but they are not able to benefit from their partners' findings and therefore the group score is lower.

The goal of this evaluation was to find influences and effects of tabletop sizes regarding groups' communication and collaboration in collaborative search. The results have shown partially significant, partially tendential differences in communication and collaboration behaviors. Not all of these differences are important for group work in general and thereby for the process of it in fields of applications as schools or universities.

One of the **main purposes of group work is to gain knowledge** about a given task or a topic as a group. Thereby it is important that each individual group member can work independently and gain information as well as knowledge about the topic. Additionally, group activities need to be supported and the final outcome should be a democratic solution with all members contributing, understanding and knowing equally (Evans & Chi 2008). Therefore, the individual and groups scores are of main interest. The results show that groups working on sizes 2 and 3 have higher individual as well as group scores, which can be seen as the outcome and result of their group work activities. Their scores are similar, which allows to recommend the setting of both of them for a possible usage in future group work scenarios. Regarding financial factors, that are not included in this evaluation, size 2 might be preferable, as its price is lower than size 3.

6 Conclusion

This chapter concludes the thesis. It summarizes the different chapters as well as results and provides future perspectives regarding scientific and technological aspects.

6.1 Summary

Goal of this thesis was the development and evaluation of the TwisterSearch system that supports collaborative search activities. The theoretical background shows insights in individual and collaborative search behaviors as well as different collaboration styles. Frameworks like Reality-based Interaction or Blended Interaction help to develop novel interactive systems that can for example evolve as distributed user interfaces.

Research projects regarding collaborative search and analytical activities are presented to get familiar with the application domain. Additional table-based research helps to understand the influence of tables for group work activities.

The TwisterSearch system allows groups to search and work collaboratively with the help of an interactive shared tabletop and the usage of personal devices. Thereby the shared tabletop is used for group activities like categorizing, arranging and clustering of search terms, which helps to communicate ideas with other group members and benefit from an around-the-table situation of multiple group members. All of them are equipped with an own personal device that can be used to trigger full text searches and read articles regarding the task of the group. Interesting findings can be shared with other group members on the shared work space.

The evaluation of the TwisterSearch system was used to find effects and influences of three different tabletop sizes on groups' collaboration and communication. Therefore, participants needed to find a solution to a given task in an analytical group work scenario.

The results of the experiment with three different tabletop sizes (namely 10.6″, 27″, and 55″) have shown tendential and significant differences regarding collaboration and communication of groups. Groups working on larger tabletops work more collaboratively, tend to use the shared work space more frequently in a more structured way. In contrast to this, the smallest tabletop lets groups evolve different roles. These trends might influence the outcome of each group, as groups working more collaboratively and with same rights show higher individual and group scores.

6.2 Future Work

Based on the evaluation of the TwisterSearch application and the subsequent results and findings there are different possibilities of future work. On one hand there are possible scientific follow-up studies to gain knowledge about the way groups communicate and collaborate with each other from various perspectives. On the other hand there are aspects that aim on facets that concern the technology and furniture.

The web-based nature and thereby the versatility of the TwisterSearch application allowed to conduct the shown experiment as there are no special requirements or installation routines for it. The different interfaces literally work on any newer device featuring a modern web browser. Reverse said, this enables to study group work scenarios from a wide variety of perspectives. In addition to the mentioned ideas of conducting studies with varying physical desk sizes and a fixed non-interactive edge, there are some more possibilities: The value and influence of the given size of an interactive group space could be studied by adapting the digital elements to have the same physical size on all three tabletops. Thereby, the given space on the tabletop would vary for each condition. Possible assumptions are different working strategies as shown by (Hajizadehgashti 2012) where the smaller work space was e.g. used to pile elements.

The value of the physical presence or absence of group partners could be studied in a remote scenario. Group members could be separated in different rooms, whereas each of them is equipped with a personal device and an own interactive group work space that is synchronized with the others. Thus, all group members could contribute equally to a task's solution and see each other's activities on the shared work space. Participants thereby could communicate via an included chat or by talking to each other via microphones or via video conference systems. This would additionally allow to investigate the value and influence of verbal, non-verbal and face-to-face communication.

Another interesting study could deal with varying sizes of personal devices. In the shown experiment, most of the participants are used to smartphones. Using them as a personal device could possibly affect the ratio of *Reading* and *Talking.* It could influence they way group members *Show and Tell* their results and findings, as they are more used to the size of smartphones.

The number and seating arrangements of participants might be interesting, too. As the two participants were seated at the two long sides of each tabletop, they also could be seated to the shorter sides. Additional participants could be seated in various ways to support different types of face-to-face communications. This could also influence the way the shared tabletop is used as a higher number of participants reduces the available space per head. Therefore, there could also be different shapes of shared work spaces fitting the needs for different numbers of group members and seating arrangements.

As this experiment was conducted as a between-subjects design, it might be interesting to have a within-subjects design where each group uses all different tabletop sizes for their group work activities. Therefore, three different tasks and participants willing to contribute for the duration of three times 90 minutes would be needed. Realistically, the duration of each task would have to be shortened to decrease mental demand, frustration and fatigue. It might be possible to rate the different sizes based on the participants' subjective rating.

Studying collaboration and communication behaviors in general is a complex topic with many possibilities and perspectives. There are many facets that could be included in future research: Whereas the communication behaviors in this experiment focused on verbal communication, future studies could focus on non-verbal communication as facial expressions and gestures are used to show feelings, beliefs and to enhance communication in general. Another interesting aspect could be to find influences regarding the position and orientation of group partners towards each other.

The combination of the TwisterSearch application with current research projects on multi and cross-device environments like Connichiwa (Schreiner et al. 2015) or HuddleLamp (Rädle et al. 2014a; Rädle et al. 2014b; Rädle et al. 2015) seams promising. Both allow to combine and assemble multiple devices like tablets to create e.g. larger landscapes that are spread but connected on several devices. This could lead to a modular approach of the shared work space. Multiple group members could place their personal tablets on an office desk and *stitch* them together to create the shared work space. During the solution of their task they could dynamically add or remove tablets to the shared work space whenever they are in need of more or less space. In addition to this, different arrangements could create various shapes that fit the purpose of the task or help to simplify the solutions as the shape could help to express the understanding of the problem (e.g. compare circular vs. arrow shaped work spaces to express aspects of time). As the shared work space would be spread among multiple devices like tablets, clusters of keywords could be assigned to specific tablets and group members could take them to start personal search activities out of the shared work space according to the keywords in the specific cluster. This could raise the awareness of all group members on the current activities of everyone. Rearranging clusters and keywords could either be done by moving them on the shared work space or by moving entire tablets to new positions or rotations. This physical activity could be supportive for the shared mental model. A mockup of a combination of the TwisterSearch system with Connichiwa or HuddleLamp can be seen in Figure 43.

Figure 43: TwisterSearch using Connichiwa or HuddleLamp

The modular approach could also lead to different purposes for the variety of used devices. They could be dynamically assigned to various views on found information. Some tablets could show events in a timeline view whereas others could focus on relations between important persons. Additionally other devices could be used for handwritten notes. The different points of view could be interrelated by spatial arrangements of tablets to gain new insights and find more information. The combination of TwisterSearch with HuddleLamp or Connichiwa might be interesting regarding possible future studies. The interactive surface of the shared work space and its usage or influence could be compared by providing different numbers or designs of tablets or phones. As an example, there could be a large number of mobile phones and a smaller number of tablets, both representing the same amount of space. It might be interesting if there are differences regarding the usage of the given devices or if a variety of devices might lead groups to more fine-grained structures and solutions.

Mobility is an important factor for German high schools. This can be seen for various aspects: TVs are placed in shelves on wheels to allow teachers to move them freely and use them for their teaching wherever they want to. Moveable computer supported working stations including projectors allow for presentations in random classrooms and materials for creative group work tasks can be found in suitcases (so-called “Methodenkoffer”) to be able to work flexibly. The reason for this mobility has financial and organizational origins. Assigning rooms to specific activities or subjects on the one hand demands more organizational structure as only one school class at the same time can usually use a specific room. Timetables are needed to overcome this. The disadvantage of the timetable approach is a less flexible structure for teachers and school classes. On the other hand many schools cannot effort to endow entire rooms with specific technologies. These special rooms could also be not usable for other purposes and thereby important space could be wasted.

As high schools are a potential field of application for the TwisterSearch system, it might be beneficial to have an all-in-one mobile solution (see Figure 44 for a sketch of the solution) to support group work. The mobile solution could combine the advantages of suitcases and mobile desks. Therefore, the solution has to be collapsible for easy transportation and storage. The grasp of the collapsed desk allows the system on the one hand to be portable and on the other hand, it can be hung up on walls either for storage or to serve as an additional wall display. This allows for further scenarios of usage, too. The personal devices are stored in drawers on the left and right hand side. A centralized power supply allows to charge them there. Thus, the system is ready to use at any time. Setting up the table-based style of the system is facilitated by pressing a button next to the grasp. Thereby, the collapsed legs get folded out. The height of the table can be adjusted to the size of the students and their chairs by adapting the length of the table legs. This allows for a variety of possible scenarios and target students. The table itself consists of an embedded interactive part in the middle and a non-interactive edge around it that allows to lean onto and to place personal devices as well as additional materials on it. The edge of the table is extendable. The additional provided space transforms the table to an oval shape that invites groups to communicate more easily. Thus, this all-in-one solution would be flexible as well as mobile.

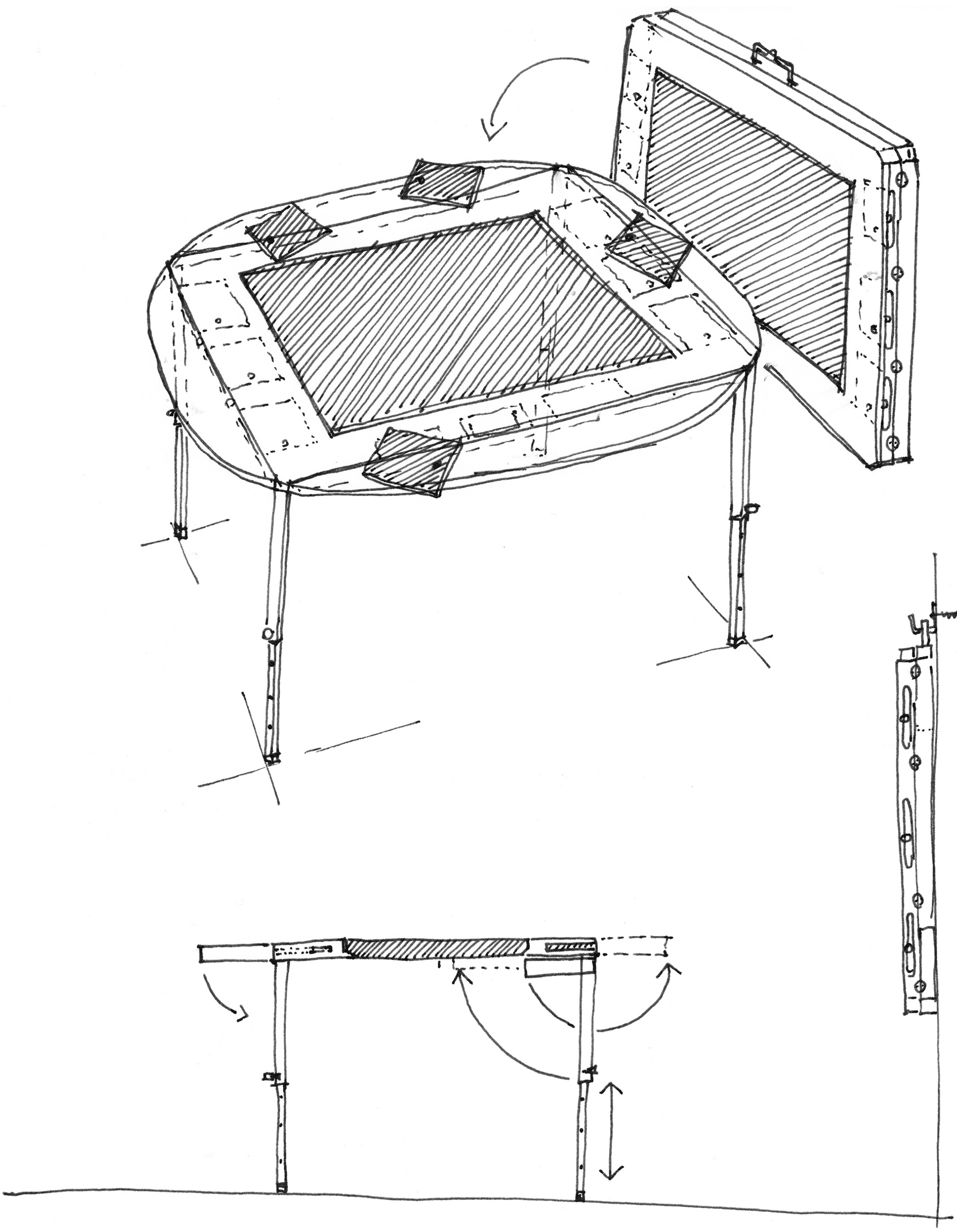


Figure 44: Sketch of the TwisterSearch all-in-one mobile solution[17]

---

[17] The ideas of the design and this sketch were developed in collaboration with architect and furniture designer Stefan Oeschger (www.jom.ch / www.bobmoebel.ch)

Bibliography

Amershi, S. & Morris, M.R., 2008. CoSearch: A System for Co-located Collaborative Web Search. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '08. New York, NY, USA: ACM, pp. 1647–1656.

Buxton, B., 2007. Sketching User Experiences: Getting the Design Right and the Right Design, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Van Dam, A., 1997. Post-WIMP User Interfaces. Commun. ACM, 40(2), pp.63–67.

Elmqvist, N., 2011. Distributed User Interfaces J. A. Gallud, R. Tesoriero, & V. M. R. Penichet, eds. , pp.1–12.

Evans, B.M. & Chi, E.H., 2010. An Elaborated Model of Social Search. Inf. Process. Manage., 46(6), pp.656–678.

Evans, B.M. & Chi, E.H., 2008. Towards a Model of Understanding Social Search. In Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work. CSCW '08. New York, NY, USA: ACM, pp. 485–494.

Fauconnier, G. & Turner, M., 2002. The Way We Think: Conceptual Blending and the Mind's Hidden Complexities, Basic Books.

Ferguson, K., 2012. Everything is a Remix. Available at: http://everythingisaremix.info/watch-the-series/ [Accessed January 26, 2015].

Field, A., 2005. Discovering Statistics Using SPSS, SAGE Publications.

Fisher, E.R., Badam, S.K. & Elmqvist, N., 2014. Designing peer-to-peer distributed user interfaces: Case studies on building distributed applications. International Journal of Human-Computer Studies, 72(1), pp.100–110.

Früh, W., 2011. Inhaltsanalyse - Theorie und Praxis 7. Auflage., Stuttgart: UTB Gmbh.

Geyer, F. et al., 2011. Designing reality-based interfaces for creative group work. In Proc. of C&C '11, p.165.

Goldstein, E.B., 2014. Sensation and Perception, Wadsworth.

Greenberg, S. et al., 2012. Sketching User Experiences, Elsevier/Morgan Kaufmann.

Grinstein, G. et al., 2006. Vast 2006 Contest - A Tale of Alderwood. Proc. of the IEEE Symposium on Visual Analytics Science and Technology, (February), pp.215–216.

Hajizadehgashti, S., 2012. Investigating the Impact of Table Size on External Cognition in Collaborative Problem-Solving Tabletop Activities.

Hall, E.T., 1990. The Hidden Dimension, Anchor Books.

Hart, S.G. & Staveland, L., 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In Human mental workload. P.A. Hancock and N. Meshkati (Eds.), Amsterdam: Elsevier, pp. 139–183.

Heilig, M. et al., 2011. Hidden details of negotiation: the mechanics of reality-based collaboration in information seeking. In In Proc. of INTERACT '11. Springer-Verlag, Berlin, Heidelberg, pp. 622–639.

Isenberg, P. et al., 2010. An Exploratory Study of Co-located Collaborative Visual Analytics Around a Tabletop Display. 2010 IEEE Symposium on Visual Analytics Science and Technology, pp.179–186.

Isenberg, P. & Fisher, D., 2009. Collaborative Brushing and Linking for Co-located Visual Analytics of Document Collections. Computer Graphics Forum, 28(3), pp.1031–1038.

Jacob, R.J.K. et al., 2008. Reality-Based Interaction: A Framework for Post-WIMP Interfaces.

Jakobsen, M.R., 2014. Up Close and Personal: Collaborative Work on a High-Resolution. ACM Trans. Comput.-Hum. Interact., 21(2).

Jetter, H.-C., Reiterer, H. & Geyer, F., 2013. Blended Interaction: understanding natural human–computer interaction in post-WIMP interactive spaces. Personal and Ubiquitous Computing, 18(5), pp.1139–1158.

Johansen, R., 1988. GroupWare: Computer Support for Business Teams, New York, NY, USA: The Free Press.

Johnson, S., 2010. Where Good Ideas Come From, Penguin Publishing Group.

Kirsh, D., 1995. The intelligent use of space. Artificial Intelligence, 73(1-2), pp.31–68.

Kirsh, D., 2010. Thinking with external representations. Ai & Society, 25(4), pp.441–454.

Morris, M., Lombardo, J. & Wigdor, D., 2010. WeSearch: Supporting Collaborative Search and Sensemaking on a Tabletop Display. Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work, pp.401–410.

Morris, M.R. & Horvitz, E., 2007. SearchTogether: An Interface for Collaborative Web Search. Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology, pp.3–12.

Paul, S.A. & Morris, M.R., 2009. CoSense: Enhancing Sensemaking for Collaborative Web Search. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '09. New York, NY, USA: ACM, pp. 1771–1780.

Rädle, R. et al., 2014a. Demonstrating HuddleLamp: Spatially- Aware Mobile Displays for Ad-hoc Around-the-Table Collaboration. In Proceedings of the 2014 ACM international conference on Interactive tabletops and surfaces (ITS '14), Demo Paper. ACM.

Rädle, R. et al., 2014b. HuddleLamp: Spatially-Aware Mobile Displays for Ad-hoc Around-the-Table Collaboration. In Proceedings of the 2014 ACM international conference on Interactive tabletops and surfaces (ITS '14). ACM.

Rädle, R. et al., 2015. Spatially-aware or spatially-agnostic? Elicitation and Evaluation of User-Defined Cross-Device Interactions. To Appear: In Proceedings of the 33rd annual ACM conference on Human factors in computing systems (CHI '15). ACM.

Rädle, R., Jetter, H. & Reiterer, H., 2013. TwisterSearch: A distributed user interface for collaborative Web search. Distributed User Interfaces: Usability and Collaboration, pp.1–17.

Ryall, K. et al., 2004. Exploring the Effects of Group Size and Table Size on Interactions with Tabletop Shared-Display Groupware. In Proc. of CSCW '04, p.284.

Schön, D.A., 1983. The Reflective Practitioner: How Professionals Think in Action, Basic Books.

Schreiner, M. et al., 2015. Connichiwa - A Framework for Local Cross-Device Web Applications. To Appear: In Proceedings of the 33rd annual ACM conference extended abstracts on Human factors in computing systems (CHI EA'15). ACM.

Shneiderman, B., 1998. Codex, Memex, Genex: The Pursuit of Transformational Technologies. In CHI 98 Cconference Summary on Human Factors in Computing Systems. CHI '98. New York, NY, USA: ACM, pp. 98–99.

Shneiderman, B., 2000. Creating Creativity: User Interfaces for Supporting Innovation. ACM Trans. Comput.-Hum. Interact., 7(1), pp.114–138.

Shneiderman, B., Byrd, D. & Croft, W.B., 1997. Clarifying Search: A User-Interface Framework for Text Searches, Corporation for National Research Initiatives.

Tang, A. et al., 2006. Collaborative Coupling over Tabletop Displays. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '06. New York, NY, USA: ACM, pp. 1181–1190.

Wallace, J.R. et al., 2011. Investigating the Role of a Large, Shared Display in Multi-Display Environments. Computer-Supported Cooperative Work (CSCW), pp.529–561.

Weiser, M., 1991. The Computer for the 21st Century. Scientific American, pp.18–25.

Zagermann, J., 2014. TwisterSearch (Technical Report).

Appendix A Pre-Questionnaire

ID:___

Pretest-Fragebogen

Herzlichen Dank, dass Sie sich bereit erklärt haben an dieser Untersuchung teilzunehmen. Bevor wir anfangen, benötigen wir von Ihnen noch einige Angaben zu Ihrer Person. Wir möchten Ihnen hiermit noch einmal mitteilen, dass alle Daten vertraulich behandelt werden.

Demografie

Alter: ________

Geschlecht: ☐ weiblich ☐ männlich

Ihre momentane Tätigkeit

____

Leiden Sie an einer Rot-Grün Schwäche?

☐ nein ☐ ja

Sind Sie…?

☐ Linkshänder ☐ Rechtshänder ☐ Beides ☐ Weiß nicht

Computeraffinität

Seit wie vielen Jahren benutzen Sie einen Computer? _________

Wie viele Stunden verbringen Sie pro Tag an einem Computer?

☐ bis zu 1 Stunde
☐ mehr als 1, bis zu 2 Stunden
☐ mehr als 2, bis zu 3 Stunden
☐ mehr als 3 Stunden

Wie würden Sie sich selbst als Computernutzer einschätzen?

Anfänger Experte

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

1

ID:___

Interaktionsgeräte

Haben Sie bereits Erfahrung im Umgang mit berührungsempfindlichen Eingabegeräten (z. B. Apple iPad oder Samsung Galaxy Tab)?

☐ ja ☐ nein

Falls ja, welche Art von berührungsempfindlichen Eingabegeräten haben Sie bereits selbst genutzt?

Nennen Sie die Arten von berührungsempfindlichen Eingabegeräten die Sie nutzen oder bereits genutzt haben (z.B. Handy, Fahrkartenautomat)?

____

____

Bei welchen Gelegenheiten benutzten Sie diese?

____

____

Wie häufig nutzen Sie diese Art von Geräten?

sehr häufig ☐ ☐ ☐ ☐ ☐ sehr selten

Appendix B Post-Questionnaire

ID:___

Post-Fragebogen

Wir möchten Ihnen hiermit noch einmal mitteilen, dass alle Daten vertraulich behandelt werden.

ID:___

Beanspruchungsstruktur

Geben Sie jetzt bitte an, wie hoch die Beanspruchung in den einzelnen Aspekten war. Markieren Sie dazu bitte auf den folgenden Skalen, in welchem Maße Sie sich in den sechs genannten Aspekten von der Aufgabe beansprucht oder gefordert gefühlt haben:

Beispiel:

gering hoch

Geistige Anforderungen

Wie viel geistige Anstrengung war bei der Informationsaufnahme und bei der Informationsverarbeitung erforderlich (z.B. Denken, Entscheiden, Rechnen, Erinnern, Hinsehen, Suchen, ...)? War die Aufgabe leicht oder anspruchsvoll, einfach oder komplex, erfordert Sie hohe Genauigkeit oder ist sie fehlertolerant?

gering hoch

Körperliche Anforderungen

Wie viel körperliche Aktivität war erforderlich (z.B. ziehen, drücken, drehen, steuern, aktivieren, ...)? War die Aufgabe leicht oder schwer, einfach oder anstrengend, erholsam oder mühselig?

gering hoch

Zeitliche Anforderungen

Wie viel Zeitdruck empfanden Sie hinsichtlich der Häufigkeit oder dem Takt, mit dem Aufgaben oder Aufgabenelemente auftraten? War die Abfolge langsam und geruhsam oder schnell und hektisch?

gering hoch

Ausführung der Aufgaben

Wie erfolgreich haben Sie Ihrer Meinung nach die vom Versuchsleiter oder Ihnen selbst gesetzten Ziele erreicht? Wie zufrieden waren Sie mit Ihrer Leistung bei der Verfolgung dieser Ziele?

gut schlecht

Anstrengung

Wie hart mussten Sie arbeiten, um Ihren Grad an Aufgabenerfüllung zu erreichen?

gering hoch

Frustration

Wie unsicher, entmutigt, irritiert, gestresst und verärgert fühlten Sie sich während der Aufgabe?

gering hoch

2

ID:___

Gruppenarbeit

Sie werden auf der folgenden Seite eine Reihe von Aussagen finden, die Sie nach dem Ausmaß Ihrer Zustimmung beurteilen sollen. Machen Sie bitte ein Kreuz in dem Kästchen, das Sie am zutreffendsten finden.

Beachten Sie bitte folgendes:

- Es gibt keine richtigen oder falschen Antworten. Es geht um Ihre persönliche Einschätzung.
- Überlegen Sie bitte nicht erst, welche Antwort den „besten Eindruck" machen könnte, sondern antworten Sie so, wie es für Sie persönlich gilt.
- Lassen Sie bitte keine Aussage unbeantwortet.
- Denken Sie nicht lange über den Satz nach, sondern geben Sie die Antwort, die Ihnen unmittelbar in den Sinn kommt.

ID:___

	Trifft nicht zu	Trifft weitgehend nicht zu	Trifft eher nicht zu	Teils teils	Trifft eher zu	Trifft weit-gehend zu	Trifft zu
Ich konnte mich gut in die Gruppenentscheidung mit einbringen.							
Die Gruppe hat sehr gemeinschaftlich an den Aufgaben gearbeitet.							
Eine Person hat die Gruppenentscheidungen dominiert.							
Mein Gruppenpartner hat meine persönlichen Präferenzen berücksichtigt.							
Ich hatte das Gefühl, dass sich alle gleichermaßen eingebracht haben.							
Ich habe gerne mit meinem Gruppenpartner zusammengearbeitet.							
Wir haben beide gleich stark zur Lösungsfindung beigetragen.							
Mein Gruppenpartner war aufmerksam, wenn ich etwas beigetragen habe.							
Ich bin zufrieden mit den Ergebnissen, die wir erzielt haben.							
Es hat mir Spaß gemacht, mit dem System zu arbeiten.							
Die Ergebnisfindung verlief chaotisch.							
Ich fand die Koordination in der Gruppe schwierig.							
Das System hat uns gut bei der Bearbeitung der Aufgaben unterstützt.							
Das System ist nicht für die gemeinschaftliche Suche geeignet.							
Wir haben uns bei der Lösung der Aufgabe gut ergänzt.							
Die Größe des interaktiven Tisches war ausreichend.							

ID:___

Rollen

Hatten Sie und Ihr Gruppenpartner unterschiedliche Rollen bei der Lösung der Aufgabe?

☐ Ja

Welche Rollen hatten Sie?

____

____

Welche Rolle hatte Ihr Gruppenpartner?

____

____

☐ Nein

Warum nicht?

____

____

ID:___

Vorgehensweise

Wie sind Sie bei der Lösung der Aufgabe vorgegangen? Zählen Sie wesentliche Arbeitsschritte auf.

Wie ist Ihr Gruppenpartner vorgegangen? Zählen Sie wesentliche Arbeitsschritte auf.

ID:___

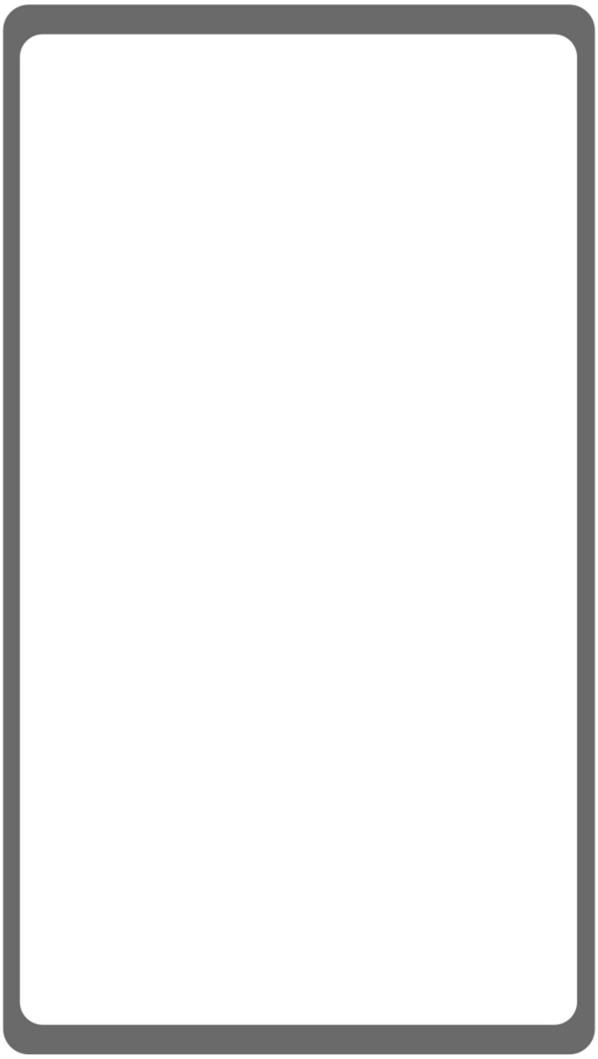
Ergebnisse - inhaltlich

Beschreiben Sie in Ihren eigenen Worten, was Sie herausgefunden haben.

ID:___

Ergebnisse - visuell

Zeichnen Sie in das nachfolgende Bild die finale Struktur ihres Ergebnisses, wie es auf dem interaktiven Tisch zu sehen war. Nutzen Sie die bereitgestellten Farbstifte um die verschiedenen Elemente, wie z. B. Cluster, Stichwörter und Resultate hervorzuheben.



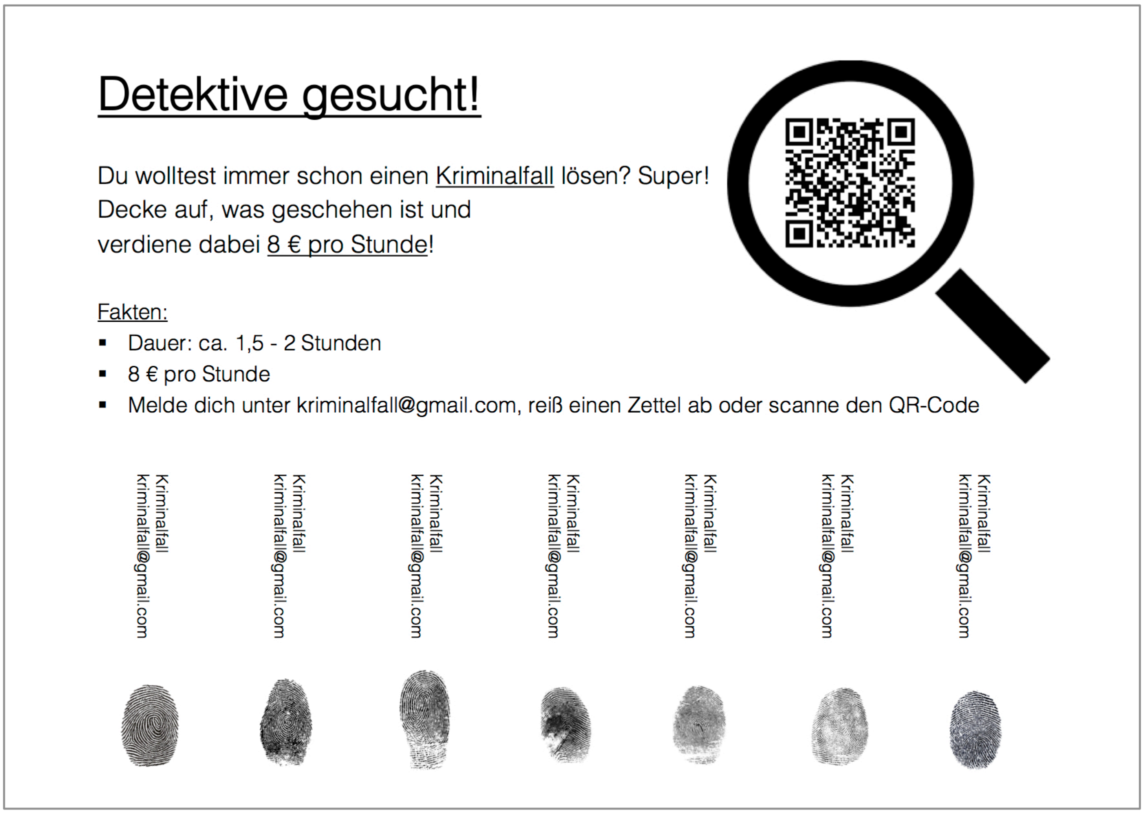
8

ID:___

Nutzung des interaktiven Tisches

Wie haben Sie bei der Lösung der Aufgabe den interaktiven Tisch eingesetzt?

Appendix C Flyer



Detektive gesucht!

Du wolltest immer schon einen Kriminalfall lösen? Super!
Decke auf, was geschehen ist und
verdiene dabei 8 € pro Stunde!

Fakten:

- Dauer: ca. 1,5 - 2 Stunden
- 8 € pro Stunde
- Melde dich unter kriminalfall@gmail.com, reiß einen Zettel ab oder scanne den QR-Code

Kriminalfall kriminalfall@gmail.com
Kriminalfall kriminalfall@gmail.com
Kriminalfall kriminalfall@gmail.com
Kriminalfall kriminalfall@gmail.com
Kriminalfall kriminalfall@gmail.com
Kriminalfall kriminalfall@gmail.com
Kriminalfall kriminalfall@gmail.com

Appendix D Content of the USB Flash Drive

The USB flash drive contains following folders and files:

- **Noldus Observer Data:** The complete video coding outcome.
- **Screenshots:** All screenshots of the final structure of all groups.
- **Study Documents:** All questionnaires used during the study.
- **Thesis:** The thesis document as a PDF file.
- **TimeVis:** Web app to visualize collaboration and communication.
- **Video:** A video showing all features of TwisterSearch.