**Making e-collaboration usable with Digital Affinity Diagram System (DADS)**

William Widjaja

Department of Management of Science and technology Human interface lab, Tohoku University Aobayama 6-6-01-2, Sendai, Japan william.widjaja@most.tohoku.ac.jp

# ABSTRACT

Collaborative affinity diagrams are a popular method for structuring idea exchanges. Creating an affinity diagram involves each member of a team writing individual points on sticky notes and making groups and links between these opinions with others by moving and arranging the sticky notes on a common wall or whiteboard. Many teams continue to use pen and paper to exchange ideas despite problems with usability. E-collaboration solutions so far have tried to provide more usable environments for affinity diagrams, but they have not been widely adopted. The Digital Affinity Diagram System (DADS) is groupware for digital sticky note creation and organization that integrates private input and common interactive spaces in one system. Using an extended dualmonitor setup to separate the two spaces, individuals can build stronger points by attaching supporting documents and media in a private space, and collaboratively organizing these points in a common interactive space synchronized across terminals. Our study compares the usability of the analog penand-paper method to the usability of the digital system we designed. We measure usability in a post-experimental survey, behavioral measures of system actions, and the System Usability Scale.

# Author Keywords

CSCW;Collaboration; Affinity-Diagram; Brainstorming;

Group work; Sticky-note; Groupware

# INTRODUCTION

Group problem solving involves exchanging and organizing ideas in order to come up with a satisfactory solution. Affinity diagrams, also known as the KJ method, are a popular discussion organizing tool [Kawakita, 1991]. Affinity Diagrams are often created by writing individual points on sticky notes or cards and, together with other participants, sorting the cards to uncover the points’ similarities. Collaborative affinity diagrams can be useful tools for problem solving and organizing meetings where a concrete outcome is desired. The affinity

Diagram gives discussions structure and explicit goals, making them ideal for new groups working together or working with ill-defined topics. Collaborative affinity diagrams give groups a tried and tested method to generate and refine ideas in search of the best. The KJ method has even been used within the field of e-collaboration to explore ideas about digital collaboration systems: Keary and Redfern [Keary et al, 2012] used collaborative affinity diagramming to sort and organize points made in interviews with thought leaders and innovators in conferencing and collaboration.

Pen and paper implementations of affinity diagrams present some usability problems, although they are still popular [Holtzblatt et al., 2005]. Analog idea-exchange methods create social and psychological problems that are unrelated to the exchange of content, and may interfere with the efficiency of idea generation. Studies of face-to-face discussions in groups reveal that production blocking [Diehl and Stroebe, 1987], fear of judgment by others [Gallupe et al., 1992], and social loafing [Williams et al., 1989] make pen-and-paper affinity diagrams less efficient than they could be. Social loafing can be a response to the perception that with a group, the effort needed to complete a task is the same whether it is being done by a group or an individual, and thus, their required effort is reduced. Collaborators may also believe that others are not as capable as they are, and reduce their contributions in order to meet this ‘least common denominator.’ Organizations may be better served by tools that use private idea generation to lower the impact of social loafing or production blocking [Nijstad et al., 2003].

Problems with analog discussions exist not only in the human element, but also in the technology used to create them. Teams that create collaborative affinity diagrams commonly use 3.5” square sticky notes to write their points on [?]. This makes points easy to read and quick to understand by others in the group, but even if the space for expression is expanded, the points are not integrated with supporting material. Methods that encourage more engagement with ideas may ultimately produce more creative solutions. While most uses of the affinity diagram employ sticky notes to represent individual points, the method does not require physical objects to represent users’ ideas, and several digital implementations of the collaborative affinity diagram method have been developed, including our own. Indeed, the recent trend toward a paperless office means that e-collaboration and paperless implementations of collaborative tools deserve further research.

E-collaboration, described by Kock [Kock, 2005] as using electronic technologies among different individuals to accomplish a common task, presents many opportunities to change traditional methods for the better in these respects. There is a long tradition of computer supported collaborative work that we can build on to encourage interaction through technology and proxemics while avoiding face-to-face brainstorming problems by protecting an individual idea-space. Benford et al. offer a definition of computer supported collaborative work (CSCW) that is probably the most appropriate for the current work: it “examines the possibilities and effects of technological support for humans involved in collaborative group communication and work processes” [Benford et al., 1994]. Groupware describes the technology that emerges from CSCW principles [Grudin, 1994]. There are many groupware systems available for collaborative discussion and brainstorming. Not all groupware enables synchronous discussion or collaborative work space, but collaborative affinity diagram creation requires these features.

Collaborative affinity diagram groupware has both costs and benefits that developers and users should consider. It has been suggested that digital diagrams can make sharing the results of collaborative work easier by letting the system generate a report that includes users’ points, comments, and the structure that groups create to organize their ideas [Judge et al., 2008]. Digital systems can make other actions during discussion easier, e.g. commenting on others’ points and making commonalities between points obvious by sorting them into groups, or creating visual links between points. Some digital collaboration tools have focused on the problem of recording the results of the collaborative process [Miura et al., 2011]. Important points shared during discussion are hard to document because transcription by hand takes time away from other important tasks, and even if transcribed records of discussions exist, it can be hard to distill the most relevant points. It is also difficult to know how well ideas are supported, whether speakers are basing their opinions on real facts, or are simply charismatic enough to get consensus from the group. Digital systems that preserve a record of discussions can help teams deal with similar problems consistently, or provide historical comparisons that can be used as a source of change and improvement.

However, digital collaboration tools can also suffer from the problem of unnecessary novelty: users’ initial impression of new technology is that it will be difficult to learn, and technologies that appear more novel are often rated less usable [Mugge et, 2012]. Many examples of groupware solve usability problems by adding uncommon or expensive hardware and software to accomplish the same tasks that can be performed more easily with the old method. Carpus [Ramakers et al., 2012] uses digital vision software to identify users’ hands automatically while they are collaborating on a single digital screen. These add-ons are usually not practical for most organizations, and novelty is no guarantee of increased usability. In our design of DADS we have made usability the top priority.

Another common problem with collaboration technologies is over-design or feature saturation. Enterprise-level tools are feature-rich and versatile but also monolithic and difficult to implement and adopt [Jeners et al., 2013]. A new generation of lightweight e-collaboration systems are emerging that solve this problem, producing environments that are lightweight additions to already familiar environments. Users of enterprise digital collaboration systems do not report good experiences, and are often resistant to their use, preferring face-to-face collaboration and analog methods with a secondary record via email.

Other e-collaboration tools may address one or two of these problems using technology, but without thorough usability testing, it is not clear whether any system has successfully solved all the problems of groupware and affinity diagram creation. The following review puts our system in context and outlines the specific problems we will address. Our main concern in this project is usability: to provide a solution to these problems that is usable by groups that need to improve their discussion context or incorporate digital media. Given the number of usability testing methods that have been specifically developed for collaborative discussion groupware [Brooke, 1996, Gutwin et al , 2000], we conclude that usability is an important issue in the field. Since our system was designed as integrated digital affinity diagram creation groupware, we had to solve two sets of problems: those that apply to groupware generally, and those that apply to collaborative affinity diagram systems specifically. General problems with groupware that we have identified in the literature include unnecessary novelty and feature saturation, although they also provide benefits such as parallel action by all participants, automatic documentation, and easy report generation from the digital record of discussion. We designed our system to retain these benefits from the field of groupware, and attempted to eliminate the major problems that we saw in other digital collaborative affinity diagramming systems: interface usability, personal space access to collaborative areas, and lack of depth in point creation.

Many digital tools for creating collaborative affinity diagrams have been developed since Munemori and colleagues’ GUNGEN system [Munemori and Nagasawa, 1996]. GUNGEN was perhaps the first use of computer systems to help multiple users create, group and document ideas. It implemented the collaborative affinity diagram through a distributed network coordinated by a chat system, and lacked the large group board that later digital collaboration systems have relied on. Testing GUNGEN against a paper-and-pen implementation of the KJ method showed that the digital method was comparable in productivity to the traditional method. Since then, collaborative digital affinity diagram developers have tried to surpass the usability and efficiency of traditional methods with computer-based systems.

Some systems combine analog paper notes with digital pens for recording ideas or actions. One example of this hybrid system is Miura and colleagues’ Group KJ system. Group KJ uses digital pens to encourage as much similarity between the digital and analog process as possible, and focuses on the record-keeping advantages of a digital system [Miura et al., 2011]. Until recently, touch-screen technology was somewhat primitive, so it was an appealing solution to use”hybrid” systems with one digital element in a familiar analog environment.

Several systems have been built around large touch-screen monitors. Tse and colleagues [Tse et al., 2008] developed a system that allowed grouping, organizing, labeling and linking digital sticky notes in a public arena. Similarly, Geyer and colleagues’ AffinityTable[Geyer et al., 2011] allows users to create virtual sticky notes with digital pens and move them around via a multi-touch tabletop. These systems can be grouped together based on their use of hand-written interfaces to create points, which preserves the physical interaction style of affinity diagrams, but unlike typed interfaces, limits the amount of information users can write on one point.

The Firestorm tabletop by Claphyan and company [Clayphan et al., 2011] provides users with a central touch-sensitive tabletop interface at which they can generate points and post them using individual keyboard interfaces. Firestorm’s gesture- and keyboard-based interface hybrid is interesting, but since reading direction cannot automatically orient toward the reader, sometimes target notes are upside down from a user’s perspective. Many current digital affinity diagram implementations contain the same errors in proxemics that the analog version does - individuals compete for space and public-board real estate instead of focusing on creating persuasive and relevant points. Systems that are implemented on large collaborative surfaces may have problems with the overlap of users’ territories and anxiety issues over shared space. However, we can learn from these systems that touchscreen interactions are becoming more common, so incorporating this technology into our system can enhance usability and learnability.

Other systems allow users to interact with a common large multitouch surface through distributed tablet interfaces. Ballendat and colleagues’ [Ballendat et al., 2010] system proposed access to a common wall-board space using distributed tablets, an idea that is also being adapted by [Radle et al.,¨ 2013]. These systems are similar to the original GUNGEN system in their use of individual interfaces for each user to interact with the collaborative affinity diagram. However, since users must still move idea cards around a single digital wall, they are required to negotiate personal space with team members, which may be an inefficient use of time during the creation of an affinity diagram.

Outside the category of affinity diagram creators, other researchers have developed systems that focus on using internet sources in discussion. Paul and Morris’ CoSense tool enables group web-based investigation by making better sense of how individuals reach solutions and look for information to support them.[Paul and Morris, 2009] Recent developments in automatic content generation have allowed users to include supporting information from Wikipedia and have points proactively generated by their Facebook content during digital brainstorming[Gartrell et al., 2010]. Gartrell and colleagues’ solution to generate depth with automated contributions may be helpful when a discussion lacks breadth, but it may sacrifice users’ initiative by generating text without engaging users in its content. These systems enable users to support their opinions with web-based content during discussion. We believe that adding a supporting content system to the traditional affinity diagram method would encourage greater understanding of others’ ideas. The existence of these collaborative investigation tools validates our intuition that digital collaborative discussion can benefit from access to internet sources for support.

# HYPOTHESES AND PROPOSED SOLUTIONS

# 

|  |
| --- |
| Figure 1. (A) System Description of Individual Setup and B) 3-Person Group setup ) |

Traditional affinity diagram creation using paper and pen is not as flexible as digital affinity diagram creation, although it is more widely used. Current implementations of collaborative digital affinity diagrams have problems with usability that we have pointed out in the review above, such as problems distributing space between users and failing to utilize the flexibility of digital technology for storage and annotation. The system we have developed to solve these challenges is called the Distributed Affinity Diagram System, or DADS. It uses technology to make collaborative affinity diagram creation easier, providing synchronized multi-user interaction by combining traditional computer interfaces with auxiliary multitouch monitors. We designed DADS to address five specific areas in which the usability of digital affinity diagram creators can be improved, where exchanges of ideas would commonly occur in group discussion, and where our digital system could improve upon the systems previously described. These five areas are: (1) increasing the efficiency of content creation, editing and navigation; (2) improving capacity for information source attachment; (3) making presentation easier; (4) organizing and documenting comment exchanges; (5) and increasing flexibility in collaborating on the affinity diagram.

1. **Efficient Content Creation and Editing**

While most current usability research does not focus on content creation and editing systems, we hypothesize that digital technology can make content creation and editing easier in the context of collaborative affinity diagrams. We designed our system to provide a dual screen interface, where one screen is dedicated to creating and organizing one’s own points. This design feature is intended to give users the freedom to focus on thinking about and creating contents during the idea generation phase of creating the affinity diagram. There is a delicate balance between providing enough private space to users and encouraging them to use common space as a way to promote idea exchange. The division between public and private space will also promote the efficient creation of points by enabling users to generate points without suffering from production blocking. We also hypothesize that good user interface design can improve navigation and increase information flow between users’ discussion points, leading to improved understanding of other users’ points. Our system provides an intuitive 3-column interface ( see Fig 2 (A)) that enables users to create new items and populate them with content, and easily switch back and forth between viewing their own and others’ points. The point viewing section is divided into two areas (Title and Description) which we believe will help users formulate their points more clearly. Users can toggle between their own and others’ points quickly and without using a menu, which will improve the efficiency of navigation. Users are identified by a unique color associated with their points both in the private interface and on the common board. It was important to us in designing the system that we retain familiar design features that make users comfortable with input methods, while also making improvements that enable users to collaborate with each other smoothly and efficiently.

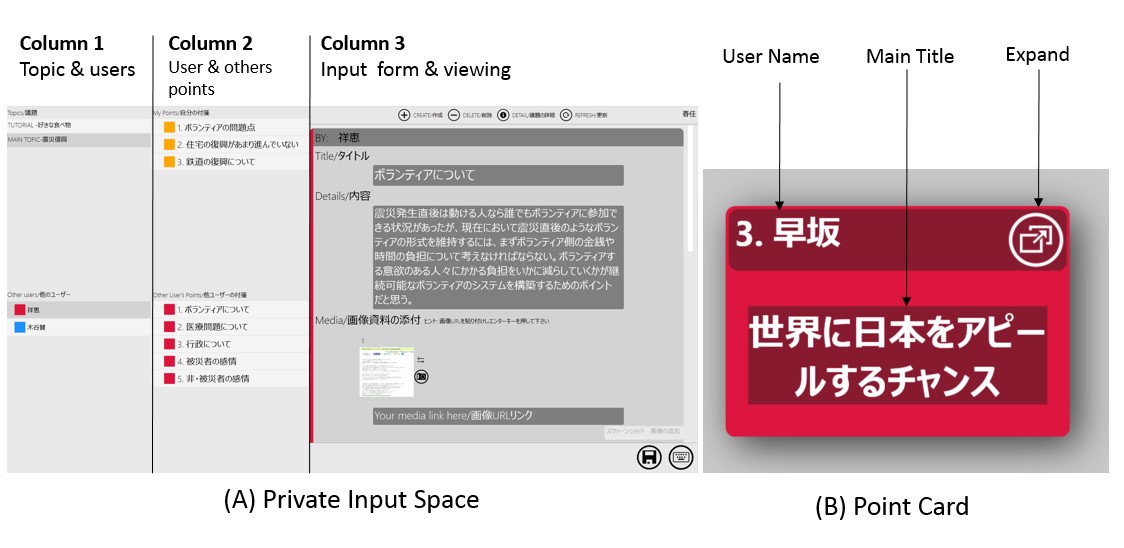


Figure 2. A) Private input UI screen with 3 column design for easy navigation and B) Individual Color-coded point card displayed in distributed UI screen

# Attaching Internet Sources

We also found no current research that focuses on the effectiveness of external source use during computer-mediated discussions. We hypothesize that sources would help users to better explain their points to others, and also make it easier for them to understand others’ points. We think that technology can make attaching information sources to points easier and more efficient. Since the fastest and most familiar source for support documents and media among our population is the internet, we needed to provide an intuitive link between the point creation area and sources of online information. The attachment system that we have designed for DADS allows users to support their points using multimedia sources. One of the simplest functions is the cut-and-paste mechanism (see Fig 3. ) that can allow users to directly quote a website source. Users can also select images to include in the support for their point. We also allow users to capture a screen shot of any website if they wish to quickly include text and images, or the image is not in a standard format. URLs and images are quickly processed into a thumbnail format that can be expanded by the user or viewer. In analog affinity diagram creation, people would be able to easily print materials and paste them on the board next to their point; however, a system that integrates source attachment functions will more efficiently promote the use and understanding of sources that support users’ points.

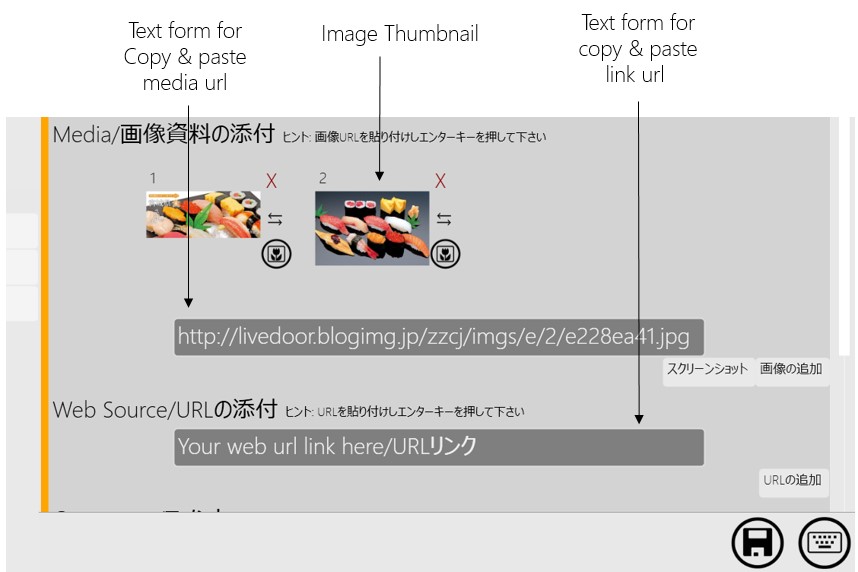


Figure 3. Screenshot to attach source

# (3)Presenting

Technology has often been used to enhance and facilitate the presentation of information to groups. Part of our integrated system is a presentation tool that can be used to make points easier to explain and easier to understand. Our proposed solution capitalizes on users’ separate screens: since each user has her own screen, presentation can take place on a distributed synchronized screen in which discussion points are in card format, with only their title displayed. Details for each point can be expanded on all users’ screens synchronously (see Fig 4. ), under the control of one user at a time, so that each user can present their own ideas. When the system is in presentation mode, card and image expansions are synchronized across all users’ common interactive boards and the central common view screen. While in this mode the opening of images, URLs, and PDFs, zoom-in and zoom-out functions, and scrolling navigation are synchronized across all users so that the documents supporting presentation are part of a seamless shared experience. In addition, a special digital pointer tool lets users highlight areas of the screen or indicate a certain passage of text.

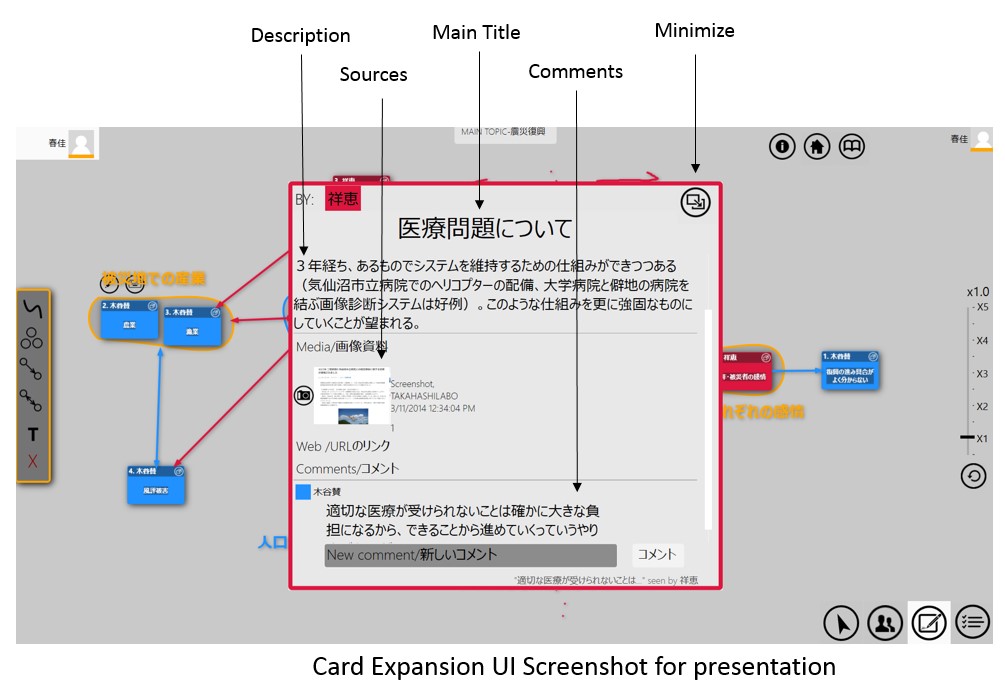


Figure 4. Client, Database, and Server Levels of System Infrastructure: Hardware and Software Setup

**(4)Organizing Comment Exchanges** While we acknowledge that speaking would be an easier and faster way to exchange comments, one drawback of speech is that it is hard to maintain a record of what was said. Since discussions involve multiple points, collaborative discussions can benefit from a method for coordinating which point should be discussed next. text-based comments allow users to simultaneously add their opinions to the discussion without interrupting anyone else, reducing the impact of production blocking and social coordination. However, hand-written comments naturally involve more movement and more resources than electronic comments, making a digital comment system a more efficient option. In our system, users can expand any card and write comments on another person’s points during the comment phase. In addition, users are notified of new comments via the system, which allows users to have real-time exchanges with each other. This is represented as a small dot that shows up on the point creator’s screen when someone leaves a comment. Users can cancel this notification by scrolling to the bottom of the list of comments briefly, at which point the system logs the comment as ”read” and cancels the notification dot until a new comment arrives. While using sticky notes to record comments on other sticky notes is intuitive, an integrated digital system will allow users to more easily organize and document exchanges of comments.

# (5)Flexibility in Creating the Affinity Diagram

Based on our observations that using sticky notes to create a collaborative affinity diagram lacks flexibility and efficiency, we hypothesize that a distributed screen solution can increase the flexibility and efficiency of organizing points using this technique. However, one way that analog affinity diagrams are often more flexible than digital collaboration is that users can arrange sticky notes more naturally, using familiar movements. We have observed that touch interfaces such as smart phones and tablet computers have become more popular because interaction with this interface style is more natural than using peripheral devices. The multitouch interface in DADS gives users intuitive contact with arranging and grouping points and drawing illustrations that help explain their points. Synchronization across distributed screens will also give users a stronger sense of control over card manipulation, observable as more card movements and actions, and lead to higher satisfaction with the final result. These capacities will encourage users to participate more effectively in creating points and discussing them with the group, move and re-arrange points more often, and ultimately lead to greater satisfaction with overall results. We have made grouping, moving, and linking points easier in DADS with three features: vector drawings, smart groups, and arrow linking.(See Fig 5. )

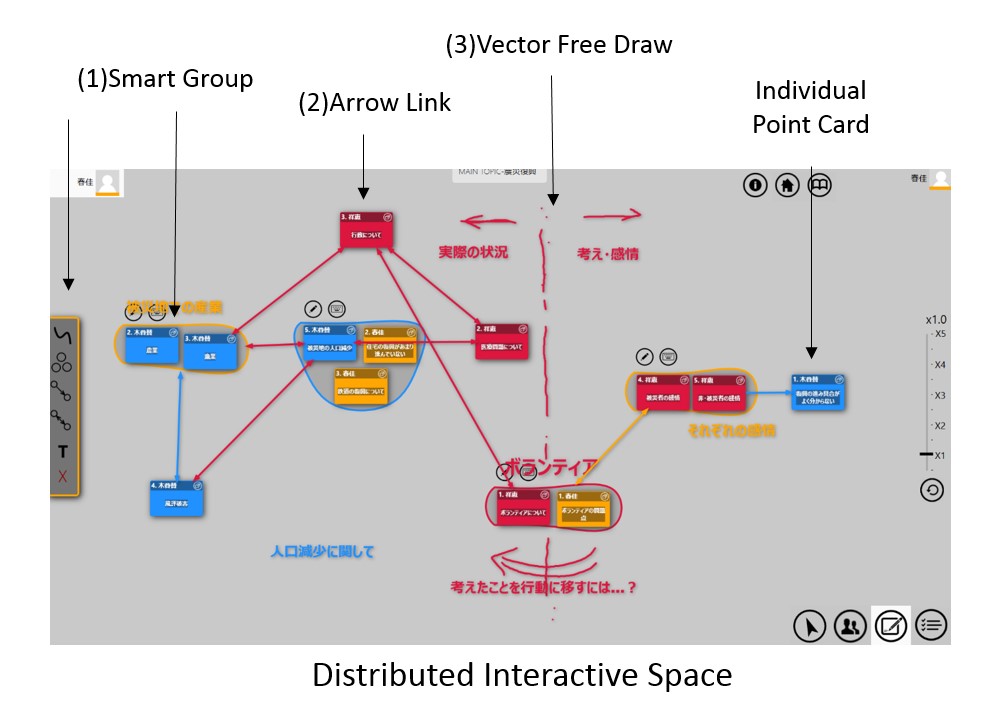


Figure 5. Actual screenshot of affinity diagram conducted in the experiment

*(1)Smart Groups*

Analog affinity diagrams require users to make groups manually that cannot be moved as a unit. However, grouping in computer-based collaboration can involve complex un-grouping and regrouping steps before the new group is formed. DADS uses a group and regroup function that allows users to create and reorganize groups easily. Users can reorganize groups by moving individual cards into or out of the group. The group flexibly re-forms according to each user’s idea of what points should be in each group.

*(2)Arrow Linking*

In analog discussions, users can draw linking arrows between two points or groups to point out causation or relationship. However, this link needs to be redrawn if the group is changed. DADS’ linking system is a simple two-click process that allows teams to have flexibility in constructing and changing the categories in which they group points during discussion. Linking arrows move with the group when users move a group of points, and can be deleted and re-linked as needed during discussion.

*(3)Vector Free Drawing*

Previous digital collaboration systems have allowed users to draw in collaboration, but these drawings quickly become confused and overlapping. Unlike traditional drawing functions in collaborative creation boards, DADS proposes a drawing system that creates vectored image components from each free-form drawing on the common board. Because users’ drawings are converted to vector images that can be moved and resized, users have more flexibility to create and re-shape their own objects on the board and helps users organize their drawing assets.

The hypotheses that we have discussed in detail above have clear implications for the design of a collaborative affinity diagram system. The functions that are required by the system should be reflected by the form of the system’s hardware and software. Below, we explain the system infrastructure that supports the features that we hypothesize will improve user experience in creating a collaborative affinity diagram.

# System Infrastructure

To solve some of these challenges through digital collaboration, we have based the current system on our previously developed system, Discusys[Widjaja et al., 2013], that has already addressed several of these challenges: specifically, separating private idea generation from common collaborative space and using a synchronous touch-screen interface. Our goal in creating this system was to build an integrated system from the ground up that makes collaborative affinity diagrams easier.

The physical setup of the system is shown in Figure 1. Users sit in a U-shaped configuration that we selected based on Kendon’s work on group dynamics [Kendon, 2010] to closely mimic the structure of group discussion. Each user has two screens that separate the private point generation interface (laptop) and common display and presentation space (multitouch 22” Iiyama monitor), and all users have equal visual access to a large version of the common display mounted on the wall. The private input interface is displayed on a laptop to maximize users’ comfort and familiarity. The multitouch monitor is set up at a 30-degree angle from the desk, which serves two purposes: (1) it maximizes joint eye gaze between users during discussion which promotes engagement across the group, and (2) it gives users a more natural angle of access between their hand and the screen for movement gestures and drawing, similar to an artist’s drafting board. The multitouch monitors are fully synchronized across all clients and with the large common display, so that collaborative manipulations can be seen by all users at once.

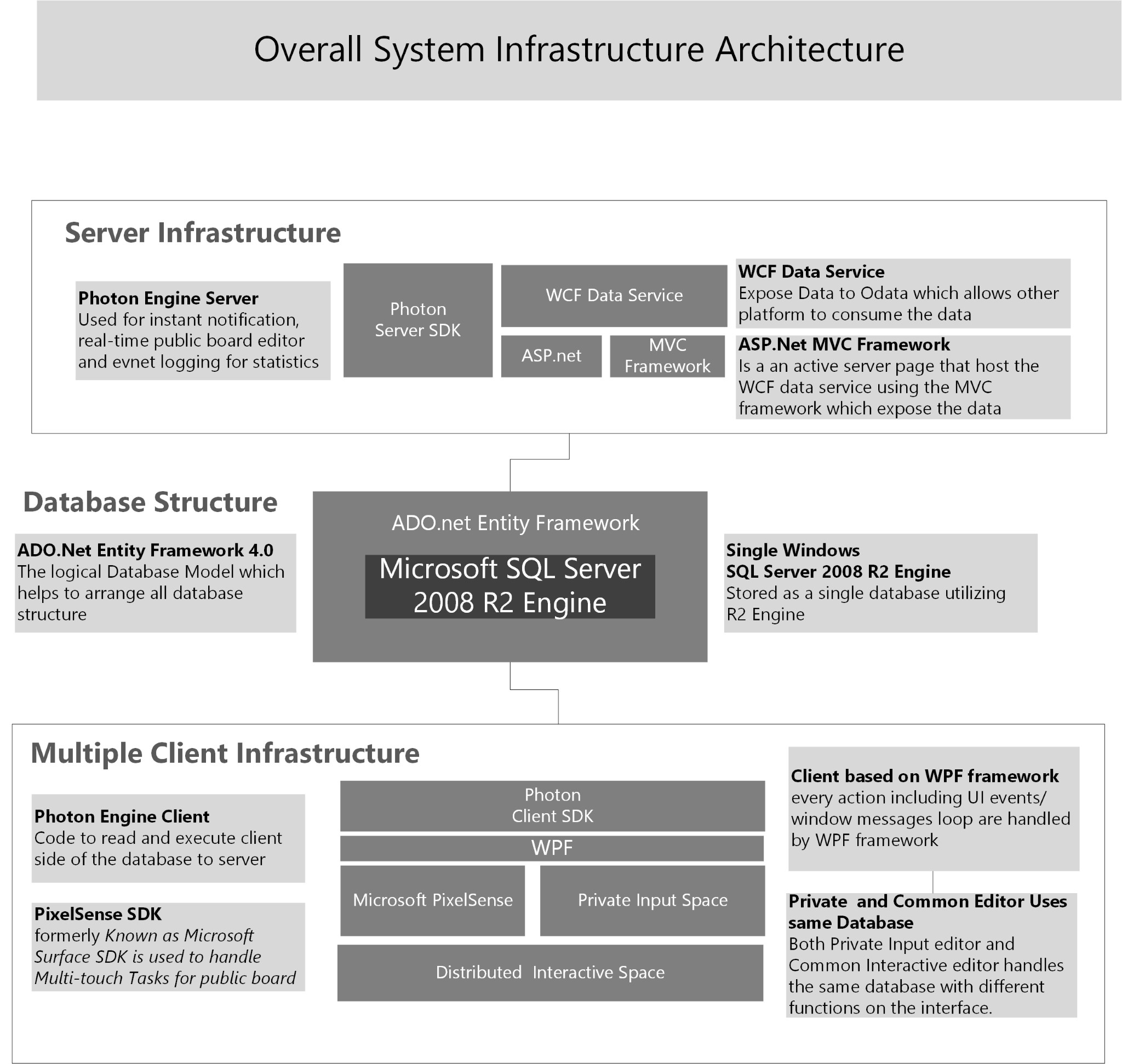


Figure 6. System Infrastrcture Stack between Client and Server

*Multiple Client Infrastructure*

The client interface was designed using Windows Presentation Foundation (WPF), which we selected so that the system could render high quality graphics (vector graphics and typography) and runtime animation, important for the ”drag and drop” movements involved in grouping virtual cards using the multitouch monitor. We developed the touch screen interface using Microsoft PixelSense SDK, for backward compatibility with Microsoft Windows 7. Microsoft PixelSense builds in a basic multi-touch gesture repertoire including pinch to zoom, flick, drag, and double-tap to support user interaction. DADS client interface also has traditional mouse support, used in the private input interface, giving users two options for manipulation. Data from actions in the client interface (e.g. moving a card or creating a new point) is stored in a database and synchronized by the network socket server.

*Database Infrastructure*

All data is housed in a single database utilizing SQL Server 2008 R2 engine. The database map is designed using ADO.net Entity Framework, an object-relation mapping framework. ADO.net Entity Framework is designed for managing information using four basic commands: Create, Read, Update, and Delete (CRUD). SQL database frameworks are quickly developed and easily maintained compared to more powerful tools like MS LINQ. Since this system is a prototype, we propose to use the simpler and faster methods. Using SQL Server also maintains compatibility across all levels of the infrastructure that use Microsoft tools. Data in the database can be easily created, read, updated, and deleted from the client end of the system, and the network socket server can access and synchronize these updates across all users.

*Server Infrastructure*

The server performs two important functions in DADS: storing the database and its structure, and manipulating the database synchronously across all client instances using a network socket service. The database is stored using the XML file format, which is web ready and highly portable between programs. The network socket service used for DADS is the Photon Engine, developed by Exit Games [Exit Games]. We use Photon Server SDK in the server application to automatically update changes between clients and servers, for notification and synchronization between client displays. Rather than a traditional system, we chose Photon Engine for its ability to instantly notify all other parties in a distributed group of clients. Data exposure is managed by ASP.net and Windows Communication Foundation data services (WCF). WCF data services exposes this data through the Open Data (OData) protocol, making it easy for DADS to provide data to other platforms such as Android, iOS, or a web portal. At present the data services integration is peripheral to the performance of the system, but this cross-platform flexibility lays the foundation for future development and extension, e.g. into tablet applications.

# Distributed Common Interactive Space

DADS is constructed to achieve two larger goals: to have higher usability, and to be more efficient than analog methods of collaborative affinity diagram creation. The technical details of the client interface, database, and server system are only useful to the degree that they improve our users’ experience and make it easier to accomplish the tasks involved in creating a collaborative affinity diagram. The following section discusses the method by which we compared our system to the analog version of the same activity, to see which system is more usable and more efficient.

# EXPERIMENT METHODS

While our previous research [Widjaja et al., 2013] focused on validating the Discusys system and its basic usability, the current paper explores usability with a more accurate survey and analysis of performance. This experiment tests the hypotheses we have explained above, concerning content creation and navigation, attaching sources, presentation and explanation, comments, and flexibility. We also explore several correlations that suggest interesting relationships between usability reports and users’ performance with the system. To strengthen our claims, we have increased the number of discussion participants per group from 2 to 3 in order to simulate a more realistic discussion. We have also tried to standardize the data output for more stable results by structuring the discussion using recommended topics and a modified Nominal Group Technique sequence.

Previous researchers have tested digital tools against their analog counterparts and shown that the digital tool has better usability. Researchers have compared analog methods of gallery writing to electronic gallery writing and found that it was a much more efficient method [Ghosh and Aiken, 2013]. In order to test the usability of our system against traditional methods, we set up two different idea exchange environments which we refer to as Analog and Digital methods. The affinity diagram process was originally developed to overcome some of the problems inherent in brainstorming and exchanging points [Kawakita, 1991]. The Nominal Group Technique (NGT) refines the process used in creating affinity diagrams. This technique is particularly helpful for groups that have not worked together before, are imbalanced in terms of extroversion/introversion, when the topic is controversial, and when some quantitative output is desired. Groups using NGT reported that this technique increased feelings of accomplishment and group efficiency compared to freely interacting groups[Gallagher et al., 1993].

During the creation of a collaborative affinity diagram, users go through 5 phases in the NGT:

1. Introduction and explanation
2. Silent generation of points
3. Sharing points by posting them on an open white board
4. Group discussion
5. Voting and ranking

This process was developed for problem-solving situations where groups had to agree on actionable next steps [Delbecq and Van de Ven, 1971]. Our process focused more on generating points and making associations between them rather than arriving at the best solution for a specific problem. Therefore, we eliminated the Voting and Ranking phase and ended our process after the creation of point groups by users.

Discussion topics for this experiment were designed to provoke many points and perhaps even heated discussion that would be moderated both by the NGT structure and by the use of digital tools. Both natural disaster recovery and nuclear energy have become important topics in Japan since the 2011 Tohoku earthquake. Naturally, most people in Japan have an opinion about how these crises have unfolded, so we thought these would be productive topics for discussion. We asked participants to discuss one of two topics in each experiment condition:

Nuclear Energy - Discuss your opinion about the current Nuclear Energy policy in Tohoku

Disaster Recovery - Discuss your opinion about the current Disaster Recovery situation in Tohoku

In the analog method, traditional tools are used including a whiteboard, colored sticky notes, and colored markers. Points and comments are written on numbered sticky notes. Users write their point’s title using a marker, and comments are written using a black pen. Comments are pasted below the relevant point note, creating a physical connection between the points and comments. To make the experiment equally balanced, we allowed users in the Analog condition to search the Internet and print out any sources they wished to use. Magnets were provided to help connect these printed materials to the point notes. In the Digital environment, users interacted with our system to conduct the experimental tasks. The U-shaped seating design in the Digital condition optimizes the dual monitor system and proximity between each participants. In our 3-person experiments, a large monitor acts as a shared common screen visible by all participants. Before the experiment, participants were given instructions and a tutorial on how to use the system. Upon completing the experiment, participants were invited to participate in a survey comparing and evaluating the system’s usability on a number of dimensions.

# EXPERIMENT PROCEDURE

Materials and experiment instructions were in Japanese, the native language of all participants. Because of our sample’s homogeneity, we could rely on participants’ cultural identity to drive some interest in the discussion topics we chose. In this experiment, we compared the Analog and Digital methods of Affinity Diagramming within groups of 3 individuals. Having an odd number of participants per group prevents subgroups from forming, e.g., 2 groups of 2. Users were given a different topic for each affinity diagram; they either discussed the recent earthquake recovery process in Japan, or recent developments in Japan’s nuclear energy policy. Each group created one Affinity Diagram using the analog method and one in the digital method. These topics and the order of conditions was counterbalanced across groups, so that an equal number of teams had the Digital condition + Earthquake topic first, an equal number had the Analog condition + nuclear topic first, and so forth. Each diagram went through 4 phases of creation, listed below.

Participants were welcomed and introduced to the purpose of the study and given a 20-minute overview of the tasks involved. If the group was assigned to the Analog condition first, they were given 30 minutes of instruction on how to create a collaborative affinity diagram using sticky notes, markers, and a white board. If they were assigned to the Digital condition first, they were given a 60-minute tutorial on using the digital system to create, support, present, comment on, navigate between, group, and link ideas in the creation of a collaborative affinity diagram. Participants had ample time during this tutorial to try out the system in the tutorial mock-up scenario, to become familiar with the system, and discover system functionality. The digital condition required a longer tutorial time in order to ensure that users were reasonably comfortable with the system. Users confirmed their familiarity with the system’s functions in a mock experiment at the end of the tutorial. Groups completed all four phases of one condition using one discussion topic before switching to the other condition and using the other discussion topic.

* Task 1**: Point Generation and Attach Source**= Users create new points based on his or her opinions and perceptions of a single topic. During the Point creation phase for both Analog and Digital conditions, users can employ internet searches in two different typical behaviors. One is to start writing immediately, and use internet-based information to verify their points. The other behavior is to search for inspiration on the internet first, and write down points based on information discovered in their search. Both point creation patterns benefit from attaching internet sources (images or URLs) to their points. Each user was given 30 minutes to create 5 points on the assigned topic. They are given access to the Internet to search for support for their points. **Time: 30 Minutes**
* Task 2**: Present points to others =**Even though points are always visible on the common interactive space board, having points explained by their creator is more persuasive than merely reading them. Therefore, our goal was to build a system that helps users explain their points easily. We believe that the availability of sources to support points will increase users’ confidence. A system that allows users to control and broadcast their sources during the explanation phase will increase ease of explanation for point creators, and increase the whole team’s level of understanding. Each user was given time to present and explain their points to the other participants. **Time : Free**
* Task 3: **Exchanging comments=** ach point is unique, and in order to understand them team members must be free to question and critique. Thus, easy navigation to specific points must be a priority in the commenting system. The exchange of points is facilitated by the same turn-taking mechanisms that operate during conversation. Therefore, our system notifies users when a new comment has arrived, and users are able to reply immediately. The group was given 20 minutes to comment on others’ points and reply to any comment posted to them. They were asked to write as many comments as possible across all the points, until the time was up. **Time: 20 Minutes**
* Task 4**: Affinity Diagram creation=** After commenting on each other’s points, participants were asked to collaborate on an affinity diagram using the points created earlier, grouping similar points and linking related points. Groups are are given unlimited time to complete the tasks without any constraints and were instructed to raise all of their hands when everyone is satisfied with the final diagrams. **Time :Free**

After completing the task in one mode (Digital or Analog), groups were given a long break before returning to complete the experiment. All participants completed all 4 tasks in both analog and digital methods, for a total of 8 tasks in each experiment group. The total time needed to complete both conditions of the experiment (with breaks) was 5.5 hours per group, on average. To show the equivalence of the two methods, a snapshot of each experimental condition is shown in Figure 7.

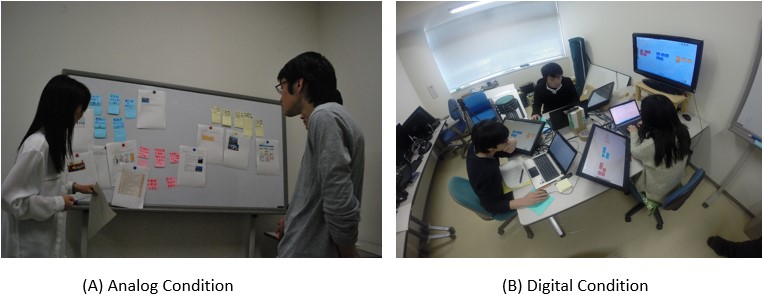


Figure 7. (A) Analog Condition: using sticky note and white board (B) Digital Condition using DADS with auxiliary multi-touch monitors.

# INSTRUMENTS AND METRICS

We operationalize usability in three ways: a usability survey administered to participants, a more general System Usability Score, and several behavioral performance metrics. Our analyses test which of the two methods (Digital or Analog) is more preferred by users, and whether DADS supports users in exchanging and organizing their points. We created a detailed survey about specific features and capabilities of the digital and analog conditions, and we also used the Standard Usability Survey to measure the digital system’s usability and learnability. We collected data on both user performance and users’ self-reported experience, and our analyses highlight the relationships between performance and experience. Correlations between survey items and performance metrics are presented to explore some interesting relationships within the system.

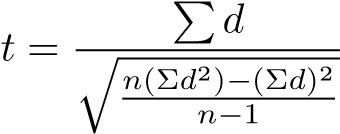
*Survey and Performance Analyses*

We designed a post-experimental questionnaire with 15 1-5 Likert scale questions addressing how easy or hard it was to perform certain tasks using either the Digital or the Analog method. Low scores indicate less reported usability. Questions are listed below with the short item code that we use in graphs and tables throughout this paper. Our specific questions were:

1. How easy was it to create points using this system? (Create)
2. How easy was it to edit points using this system? (Edit)
3. How easy was it to attach sources to your points using this system? (Attach)
4. How easy was it to explain resources and support for your points in this system? (Show)
5. How easy was it to explain your points to others? (Explain)
6. How easy was it to read other users’ points in this system? (Navigate)
7. How well did you understand other people’s points? (Understand)
8. How easy was it to read comments on others’ points in this system? (Read)
9. How easy was it to write comments on others’ points in this system? (Write)
10. How easy was it to reply to comments in this system? (Reply)
11. How flexible was it to create an Affinity Diagram in this system? (Flex)
12. How easy was it for you to cooperate with other group members to achieve the task in this system? (Teamwork)
13. How comfortable did you feel during the experiment? (Comfort)
14. How useful do you think this method is compared to the other method you experienced? (Useful)
15. How satisfied are you with the final Affinity Diagram that you made using this method? (Satisfy)

We measured several aspects of individual users’ behavior in order to get a clearer picture of what users did in each of the conditions: number of attachments made, number of comments made, number of times a card was moved, grouped, or linked, and the amount of time groups took to create the final affinity diagram. We also calculated the number of moves per minute for each user. Groups’ behavior was recorded automatically in the digital environment. In the analog condition we used a GoPro HERO3+ digital video camera mounted on a jib in order to provide visual access to all users’ movements. Video from the analog condition was hand coded by the first author.

We used paired t-tests to analyze the users’ responses and determine whether the two experimental conditions were different from each other. This test is appropriate for our analysis because each participant generates data in both experiments; in essence, we are sampling each participant twice for each measure of the questionnaire and performance metrics. The analysis is based on the differences between the values of each pair - that is, one subtracted from the other. In the formula for a paired t-test below, this difference is notated as d.



We chose p = 0.05 as our threshold for significance for the paired t-tests, reflecting a value of *α* = 0.05. This value is commonly used in experimental psychology and human factors research, since these fields deal with hypotheses that do not require high levels of certainty in order for significant claims to prove true [Bross, 1971].

# OBSERVATION RESULTS

Data collected from 8 groups of 3 for this experiment (N=24; 12 female and 12 male) is presented in the current paper. Participants were university students between the ages of 18 and 24. All subjects used a mouse with their right hand, which is best suited to the system design; one student was left handed but reported being able to use the mouse comfortably with his right hand. Twenty students were in the STEM (Science, Technology , Engineering and Math) disciplines, and 4 students were Social Science majors. Twenty of 24 participants reported that they own a smart phone. Eleven students reported using a dual-screen monitor in the past, and 13 students had not used a dual-screen setup before.

We divide the results into 3 areas: 1. Survey measures of usability and user experience, 2. Performance measures of efficient system actions, 3. Correlations between usability and performance. In the first section, we discuss the results of the survey, focusing on 1) Content Creation and Presentation, 2) Exchanging Comments, 3) Navigating, Understanding and Affinity Diagram, and finally 4) Affective Measures in the results. We then check the System Usability Scale for a different perspective on DADS’ usability, and to show that the system is scientifically usable based on multiple instruments. In the third section we discuss users’ performance based on the performance metrics. Finally, we present a correlation analysis of measures from the survey and user performance to check the strength of our initial hypotheses.

*Content Creation and Presentation*

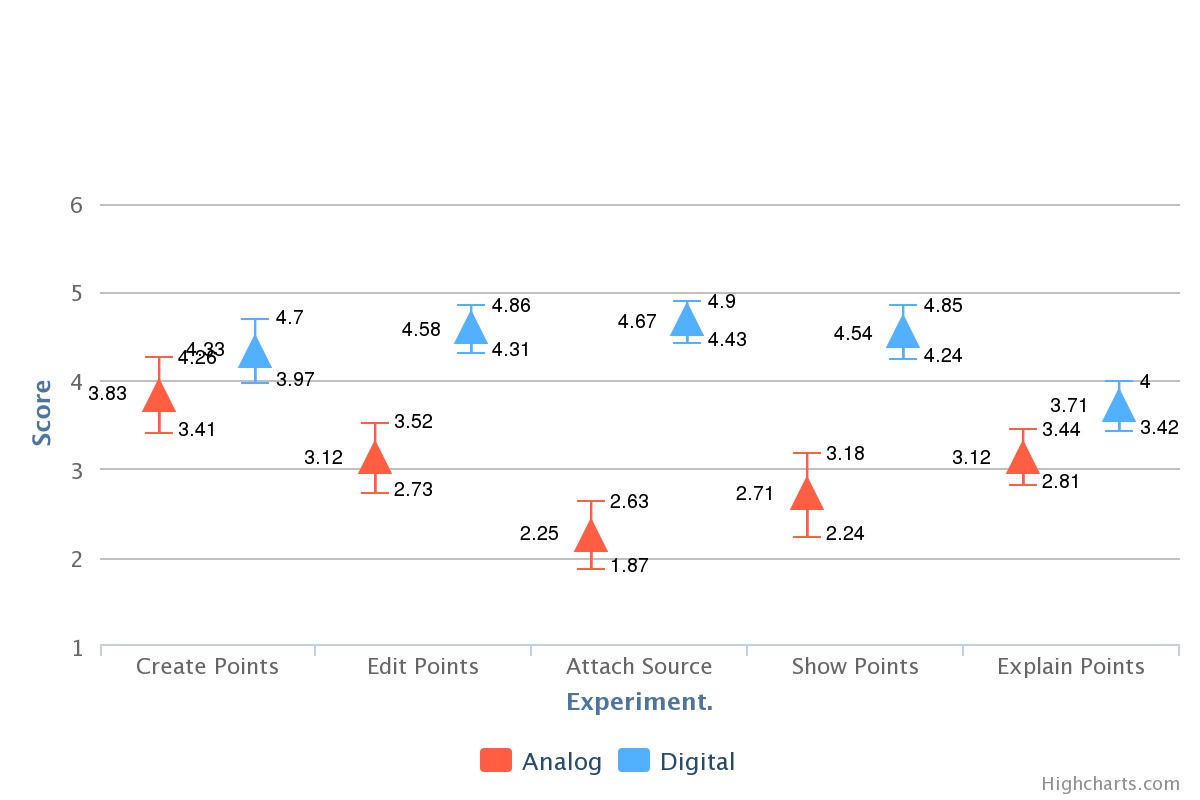


Figure 8. Usability Survey T-test Results

An analysis of participants’ answers to the usability survey questions is shown in the figures below. Paired t-tests were used to determine the relative strengths and weaknesses of the two methods (Analog and Digital), as perceived by participants.

Participants reported that it is easier to conduct the tasks involved in creating an affinity diagram using the Digital method compared to the Analog method. The ease of creating, editing, and attaching sources to points was deliberately engineered into our system, as reflected by responses to the usability survey. Participants found it marginally easier to create Digital points (M= 4.33, SD = 0.87) compared to Analog points (M=3.83, SD = 1.01) p *<*0.09). Participants also found that editing in digital environment is significantly easier (M = 4.58, SD = 0.65) than the same process in the Analog condition (M = 3.12, SD = 0.90 p *<*0.01). Editing points in the digital environment involves familiar text manipulation, while in the analog environment participants must erase and re-write information, which may discourage them from editing. Our system’s Attachment function proved quite usable, with rating for Attaching sources is significantly higher for the Digital format (M=4.67, SD =0.56) compared to Analog ( M=2.25,SD=0.90, p*<*0.01). The presentation and expansion features of our system both contributed to better usability, since users report that showing points is easier in the Digital environment (M=4.54, SD=0.) compared to Analog (M=2.71, SD=1.12 P*<*0.01)). Users also report that it is easier to explain ideas in the Digital environment (M=3.71, SD=0.69) compared to Analog (M=3.12, SD=0.74, P *<*0.05). Participants’ scores show us that the private input interface improves the most difficult tasks in the affinity diagram process: creating, supporting, and finding other users’ points. These relationships are further explored in the correlation analyses.

*Organizing Comments*

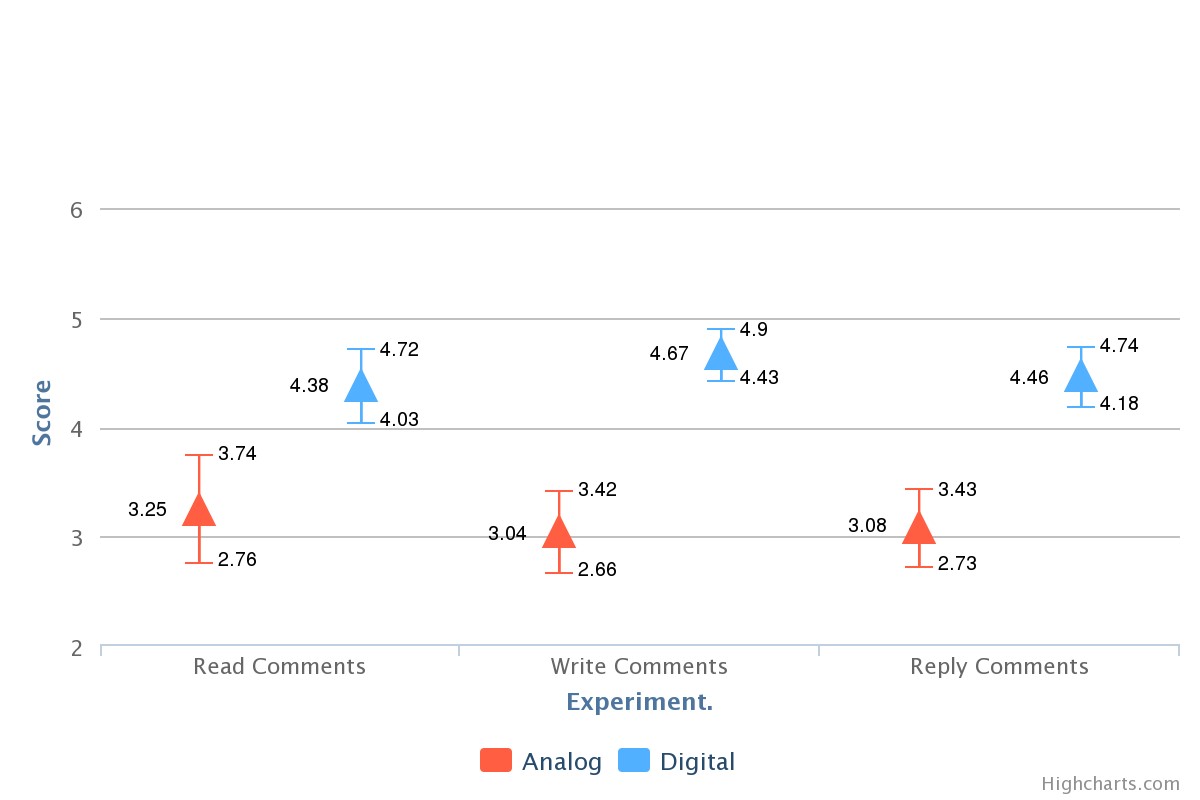


Figure 9. Analog and Digital T.test result for commenting

We evaluate the three tasks related to commenting in both methods: 1) Ease of reading comments 2) Ease of writing comments and 3) Ease of replying to comments. Users report a difference between the ease of reading in analog and digital. While all comment notes are shown in analog, in the digital version the comments are actually hidden until users choose to read them, and our data shows a significant difference between Analog ( M=3.25, SD= 1.15) and Digital (M=4.38, SD= 0.88, p *<*0.01). Subjects also reported that it is easier to write comments in the Digital system (M=4.67, SD=0.56) compared to writing manually (M=3.04, SD=0.91 *<*0.01). Finally, when evaluating the notification system for replying to comments, users rate the Digital format higher (M=4.46, SD= 0.66) than Analog (M=3.08, SD= 0.83) suggesting that it is easier to reply to comments using the digital interface.

*Navigating, Understanding and Flexibility*

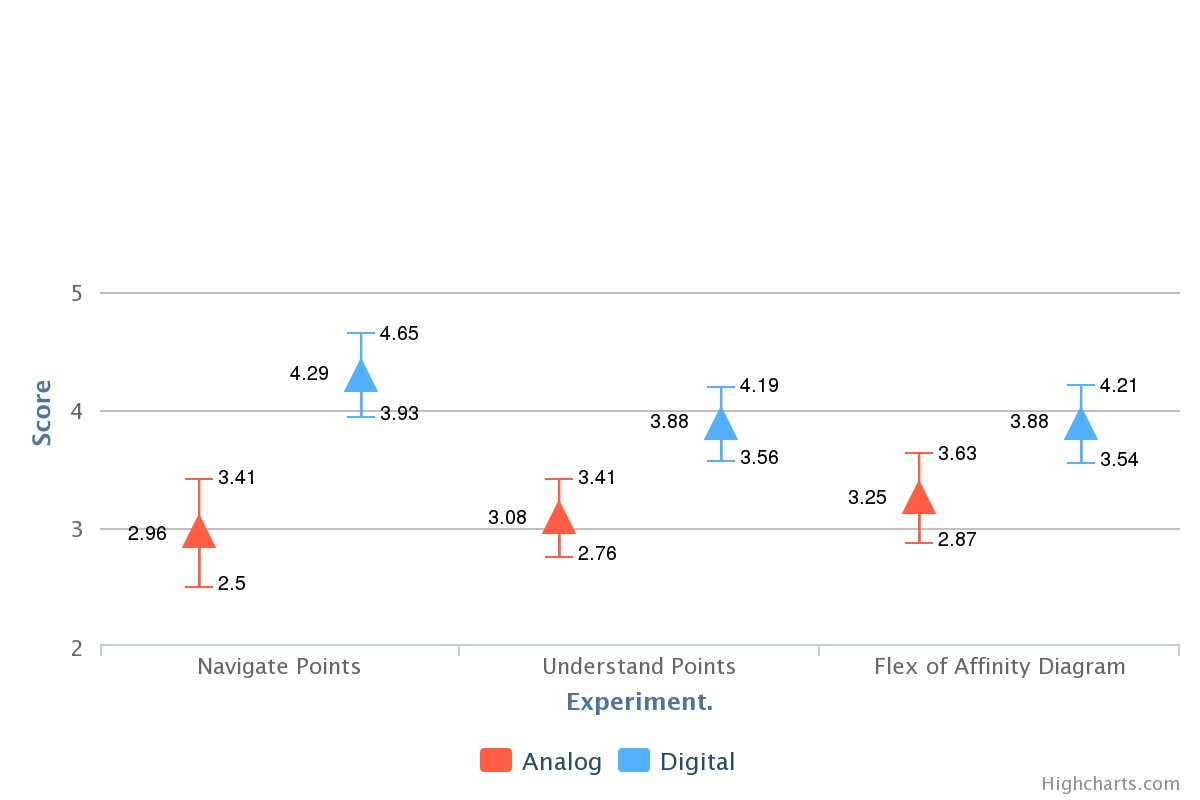


Figure 10. (A) Analog and Digital t-test result for navigation

Navigation (how easy it was to read others’ points) reflects the usability of DADS’ Private Input Space user interface. Digital environment navigation is preferred by users (M=4.29, SD=0.86) compared to Analog (M=2.96, SD=1.08, p *<*0.01) which is one of our primary goals. When we asked about how well users understood others’ ideas, we discovered that it is easier to understand the discussion points in digital format (M=3.88, SD=0.8), compared to Analog format (M=3.08, SD=0.78, p *<*0.01). Collaborating on the affinity diagram was more flexible using the digital system (M = 3.88 (0.8)) than the analog system (M = 3.25(0.9), p *<*0.05), which conclude that according to the experiment, users felt slightly more flexible organizing cards to form affinity diagram in Digital conditions compared to analog conditions.

*Affective Metrics*

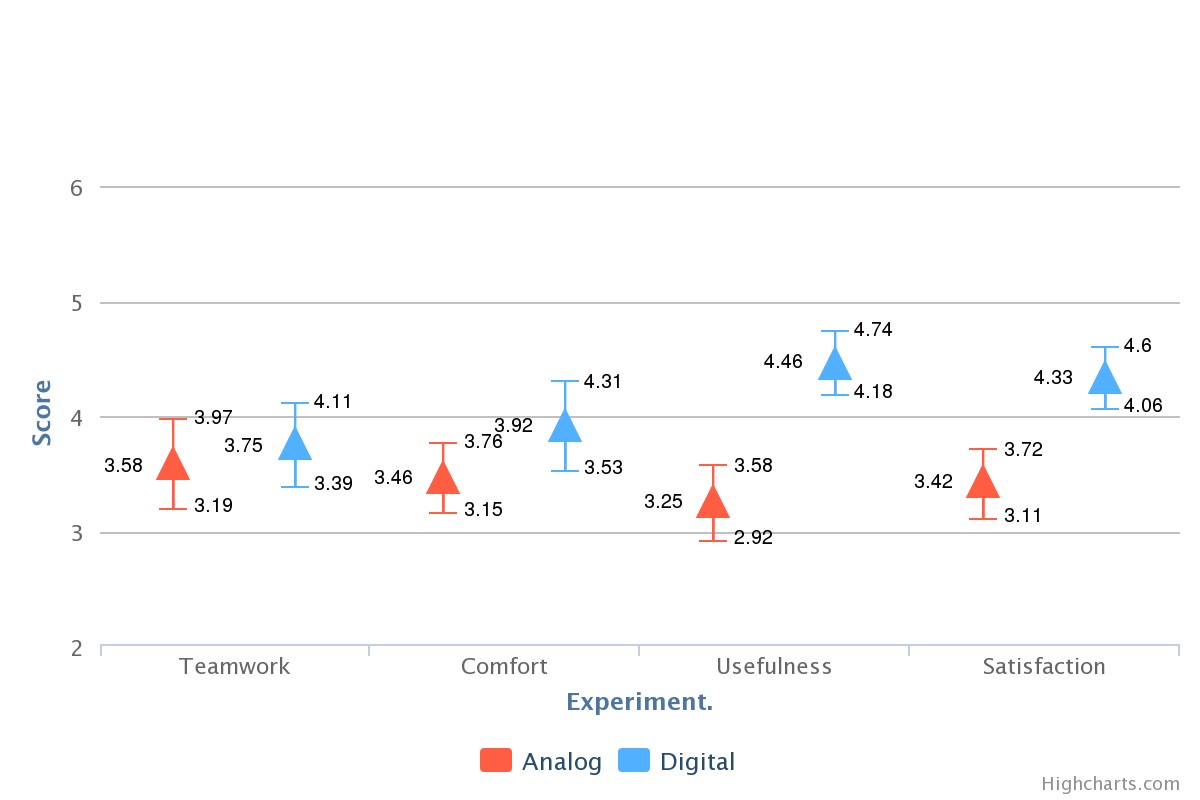


Figure 11. Analog and Digital t-test result for performance measures

We also asked users several questions that reflected their general feelings about the system: how useful it was compared to the other system, how comfortable they were using it, how satisfied they were with the end result, and whether they thought the system improved their group’s teamwork. Users felt that the Digital system was more useful than the Analog system for affinity diagram creation (Digital M = 4.46 SD = 0.64, Analog M = 3.27 SD = 0.96, p *<*0.01). They also report that they feel slightly more comfortable in the Digital system (M = 3.92, SD = 0.93) than the Analog system (M = 3.46, SD = 0.72, p *<*0.05). Users also reported higher satisfaction with the outcome of the affinity diagram from the Digital process (M = 4.20, SD = 0.68) compared with the Analog process (M = 3.42, SD = 0.72, p*<*0.01). Finally, we discovered no statistical difference between the rating of teamwork in Digital (M = 3.87, SD = 0.83) and Analog( M=3.53, SD=1.06, p = 0.31). Some participants’ responses are ambiguous between which of the two conditions promote more cooperation or teamwork. We believe this may be due to users being unable to properly coordinate the creation of the affinity diagram using the distributed screen system. As a result, our data shows that users do not differentiate between the two systems based on their feeling of collaboration.

Table 1. Full Experiment Self reporting Usability Survey result

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | N | Mean | SD | SE | 95%LCI | 95%UCI | T-Value | P-Value |
| Create Points AN | 24 | 3.83 | 1.01 | 0.21 | 3.41 | 4.26 | 1.77 | 0.09 |
| Create Points DI | 24 | 4.33 | 0.87 | 0.18 | 3.97 | 4.70 |  |  |
| Edit Points AN | 24 | 3.12 | 0.95 | 0.19 | 2.73 | 3.52 | 5.88 | 0.00 |
| Edit Points DI | 24 | 4.58 | 0.65 | 0.13 | 4.31 | 4.86 |  |  |
| Attach Source AN | 24 | 2.25 | 0.90 | 0.18 | 1.87 | 2.63 | 10.4 | 0.00 |
| Attach Source DI | 24 | 4.67 | 0.56 | 0.12 | 4.43 | 4.91 |  |  |
| Show Points AN | 24 | 2.71 | 1.12 | 0.23 | 2.23 | 3.18 | 6.26 | 0.00 |
| Show Points DI | 24 | 4.54 | 0.72 | 0.15 | 4.24 | 4.85 |  |  |
| Explain Points AN | 24 | 3.12 | 0.74 | 0.15 | 2.81 | 3.44 | 2.93 | 0.01 |
| Explain Points DI | 24 | 3.71 | 0.69 | 0.14 | 3.42 | 4.00 |  |  |
| Navigate Points AN | 24 | 2.96 | 1.08 | 0.22 | 2.50 | 3.42 | 3.97 | 0.00 |
| Navigate Points DI | 24 | 4.29 | 0.86 | 0.18 | 3.93 | 4.65 |  |  |
| Understand Points AN | 24 | 3.08 | 0.78 | 0.16 | 2.76 | 3.41 | 3.87 | 0.00 |
| Understand Points DI | 24 | 3.88 | 0.74 | 0.15 | 3.56 | 4.19 |  |  |
| Read Comments AN | 24 | 3.25 | 1.15 | 0.24 | 2.76 | 3.74 | 8.28 | 0.00 |
| Read Comments DI | 24 | 4.38 | 0.82 | 0.17 | 4.03 | 4.72 |  |  |
| Write Comments AN | 24 | 3.04 | 0.91 | 0.19 | 2.66 | 3.43 | 6.95 | 0.00 |
| Write Comments DI | 24 | 4.67 | 0.56 | 0.12 | 4.43 | 4.91 |  |  |
| Reply Comments AN | 24 | 3.08 | 0.83 | 0.17 | 2.73 | 3.43 | 2.93 | 0.00 |
| Reply Comments DI | 24 | 4.46 | 0.66 | 0.13 | 4.18 | 4.74 |  |  |
| Flex Affinity Diagram AN | 24 | 3.25 | 0.90 | 0.18 | 2.87 | 3.63 | 2.39 | 0.03 |
| Flex Affinity Diagram DI | 24 | 3.88 | 0.80 | 0.16 | 3.54 | 4.21 |  |  |
| Teamwork AN | 24 | 3.58 | 0.93 | 0.19 | 3.19 | 3.98 | 0.66 | 0.52 |
| Teamwork DI | 24 | 3.75 | 0.85 | 0.17 | 3.39 | 4.11 |  |  |
| Comfort AN | 24 | 3.46 | 0.72 | 0.15 | 3.15 | 3.76 | 2.11 | 0.04 |
| Comfort DI | 24 | 3.92 | 0.93 | 0.19 | 3.52 | 4.31 |  |  |
| Usefulness AN | 24 | 3.25 | 0.79 | 0.16 | 2.91 | 3.59 | 6.06 | 0.00 |
| Usefulness DI | 24 | 4.46 | 0.66 | 0.13 | 4.18 | 4.74 |  |  |
| Satisfaction AN | 24 | 3.42 | 0.72 | 0.15 | 3.11 | 3.72 | 6.26 | 0.00 |
| Satisfaction DI | 24 | 4.33 | 0.64 | 0.13 | 4.06 | 4.60 |  |  |

# Performance Metrics

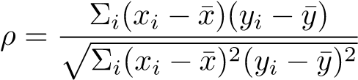
We coded participants’ behaviors in both the Analog and Digital conditions, counting the number of times cards were moved and re-grouped, as well as the number of supporting attachments users made to their points. The calculation on Table 2 is counted as the average of creation individually per discussion. Self-reports from our usability survey are supported by data from users’ behavior. Participants felt that it was easier to attach sources in the digital system (M = 4.53, SD=0.67) than in the analog system (M= 2.13, SD=0.85) p *<*0.01). This is supported by a greater number of attachments associated with each idea in the digital system (M = 5.50, SD=3.49) than in the analog system (M = 1.75, SD=1.19), p *<*0.01). Navigation and attachment required fewer and smaller physical movements by the participant in the digital environment. Since the environment was itself digital, it was less cumbersome to look up new sources on the Internet to support points.

Time spent in discussion also relates to the number of card manipulations and re-grouping done by participants. Card movements per participant are calculated per minute during the fourth phase of the experiment, Affinity Diagram Creation. On average, users carried out 18.63 (SD = 13.63) card movements in the Digital condition compared to 6.79 (SD = 1.98) card movements in the Analog condition. In simple terms, this means that users engage in more activity in Digital compared to Analog. In addition, users spent more time as a group in the Digital condition, at a mean of (M= 15.07 min, SD = 11.10), compared to (M = 7.11 min, SD = 2.11) in the Analog condition. This may be related to more time spent on creating the result and moving cards. The average Digital participant’s satisfaction score was higher (M = 4.33, SD =0.72) than their score in the Analog system (M= 3.42, SD = 0.64). The result do not show a difference in efficiency, since both Analog and Digital conditions average about 1 move per minute. When we probed further about why more time was spent on the Digital affinity diagram, we saw that subjects reported that they felt more flexible with the digital system, and may have been more inclined to extend their discussions and explore more possibilities in the creation of the affinity diagram. Instead of viewing this extended time as a drawback, it may represent participants engaging more with others’ ideas in the Digital system.

Table 2. Performance Metrics Results

|  |  |  |  |
| --- | --- | --- | --- |
|  | AN (SD) | DI(SD) | p-value |
| No of source Attached | 1.75(1.19) | 5.50(3.49) | *<*0.01 |
| No. of comments  Generated | 6.79(1.98) | 13.25(5.52) | *<*0.01 |
| No. of Car movement | 8.04(4.31) | 18.63(14.54) | *<*0.01 |
| Avg Time to complete Affinity Diagram (min) | 7.11(2.11) | 15.07(11.10) | *<*0.01 |

# Correlation Analysis

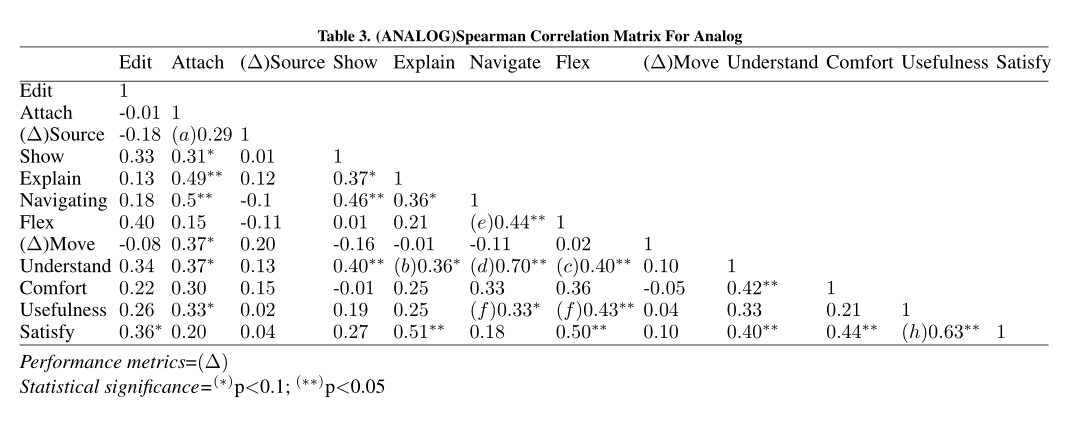
Correlations is a measurement that inform us whether changes in one metrics have any relation to changes in another result. Several of our hypotheses propose that there is a relationships between different factors in our system. In order to test these hypotheses, we computed Spearman’s rank correlation coefficients for select items from the questionnaire and performance metrics in order to explore the relationships between these factors. We chose this analysis method because initial exploration of the data showed some non-linearity. In order to be conservative, we chose a method that does not assume linearity or consistent variance throughout the data, Spearman’s rank correlation (*ρ*)[Myers and Well, 2003]. Spearman’s rank correlation is preferred over Pearson’s r in cases where consistent variance in the data cannot be assumed. Spearman’s *rs* is calculated as follows:

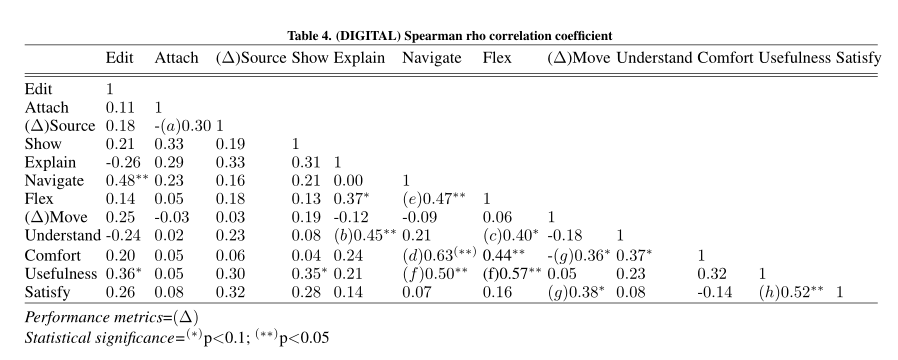
Statistical significance for Spearman’s *rs* , as for all correlation tests, is determined by comparing the current test to a ’null hypothesis’ that there is no relationship between x and y. P-values of 0.05 or lower will indicate that a correlation is statistically significant. Spearman’s *rs* indicates the strength of the relationship between x and y; if this relationship is shown to be significant, then there is less than 5% likelihood that the strength of this relationship is due to chance.

While correlations alone cannot show us a direction of causation, they suggest good possibilities for further research and analysis. We believe that in group discussion, usability of a system will cause more efficient thinking, greater understanding, and greater satisfaction among participants. These factors will be reflected in the productive actions of users. Our analysis focuses on the correlations that are not only the most significant, but also the most relevant to our hypotheses. We present correlations only between items that showed a significant difference between Analog and Digital conditions in the t-test analysis; this excludes Create Points and Teamwork which show not significantly different between the two. Comment data (create, read, and edit) is also excluded because we view the comment system as peripheral to the content creation, presentation, and affinity diagram grouping and movement functions of DADS and will be analyzed in the future research.

Based on our analysis, we discovered that subjects’ performance ratings are more stable in Analog compared to Digital. In all of the correlations we measured, there are 13 relationships that are significantly correlated in the Analog condition compared to only 8 in the Digital condition. We believe this is due to familiarity issues with the system; most participants are familiar with analog methods, and they perform and respond in a more similar manner compared to a newly developed system that relies on each participant’s level of experience with new technology.

The analyses we present focus on providing further support for our hypotheses: a) Ease of attachment helps to increase the number of sources used; b) Number of sources improves explanation and understanding; c) Flexibility also improves understanding; d) Ease of navigation improves understanding; e) Flexibility is related to higher numbers of moves and greater satisfaction with the system. We address each hypothesis in turn below, as well as three additional correlations (f, g, and h) addressing the more subjective questions in our usability questionnaire: usefulness, satisfaction, and comfort.



****

1. **Ease of attachment and number of attachment** -Our correlation analysis does not provide any evidence that the ease with which the affinity diagram was created had any relation to the number of sources attached. Despite the results of the paired t-test analysis in which users reported that the digital environment made attachments easier, some users did not utilize the attachment function during the experiment, although all users tested this function during the training phase. The perceived usability of the attachment system seemed to have no effect on whether or not the attachment function was used. This may be related to the depth or specific content of users’ points, which is outside the scope of this analysis.
2. **Number of sources, explanation and understanding** -The number of sources used is not correlated with ease of explanation or ease of understanding in either condition. However, ease of explanation and understanding are correlated significantly for the Digital condition (*rs* = 0.45, p=*<*0.05). The evidence for this relationship is present, but less clear in Analog (*rs* = 0.36, p=*<*0.1). Rather than the volume of sources, users’ ease of explanation could play a larger role in their ability to understand other parts of the discussion.
3. **Flexibility and understanding-** The level of flexibility in creating the affinity diagram is correlated with levels of understanding in both conditions. For both Analog and Digital users, this coefficient is (*rs* = 0.40, p=*<*0.05). This suggests that, rather than a relationship between understanding and number of sources, the flexibility of creating the affinity diagram is a stronger predictor of understanding during group discussion.
4. **Navigation, understanding and comfort** -Ease of navigation is also strongly correlated with understanding in the Analog condition (*rs* = 0.70, p=*<*0.05). However, this correlation is not strong or significant in the Digital condition. Instead, Digital users’ ease of navigation is more closely related to comfort (*rs* = 0.63, p=*<*0.05). Thus, it appears that ease of navigation in the two conditions may have different impact: directly on understanding for Analog users, and on affective impressions of the system for Digital users.
5. **Navigation and flexibility** -We also confirmed that ease of navigation between discussion points and flexibility are slightly correlated in both Analog ((*rs* = 0.44, p=*<*0.05) and Digital conditions (*rs* = 0.47, p=*<*0.05).
6. Navigation, flexibility and usefulness Ease of navigation is highly correlated with usefulness for Digital users (*rs* = 0.49, p=*<*0.05), while for Analog users the correlation is still positive, but not as strong (*rs* = 0.325, p=*<*0.1). However, both conditions showed a relationship between the flexibility of organizing the affinity diagram and perceived usefulness. For Digital users, flexibility and usefulness correlated highly at (*rs* = 0.57, p=*<*0.05) and Analog at (*rs* = 0.40, p=*<*0.05). This helps us define flexibility as one of the key elements of usefulness, and thus closely related to usability.
7. **Movement, comfort and satisfaction-** We were not able to establish a correlation between flexibility and the number of moves in either condition. However, we found interesting potential evidence in the Digital condition where the number of moves is negatively correlated with the level of comfort (*rs* = -0.36, p=*<*0.1) but positively correlated with the level of satisfaction (*rs* = 0.3.8, p=*<*0.1). This could mean that more movements result in less comfort, but give users more satisfying results. However, Digital user performance varied more than Analog, so these data should still be considered preliminary, prompting further exploration of this new hypothesis.
8. **Usefulness and satisfaction-** Our last correlation shows that Level of Usefulness and Level of Satisfaction are highly correlated for both Analog ((*rs* = 0.63, p=*<*0.05) and Digital users (*rs* = 0.52, p=*<*0.05). This indicates that perceptions of usefulness and satisfaction are closely linked, and that improving the perception usefulness of a system can also improve user satisfaction.

# System Usability Scale

Even though methods for testing the usability of groupware and collaborative discussion systems have emerged in recent years, we chose the industry-standard quantitative SUS over qualitative or behavioral methods, so that DADS can be more easily compared to other systems by researchers across the field of e-collaboration. The System Usability Scale (SUS) is a evaluation methods that provides an efficient way for users to evaluate a new system based on its usability and learnability[Brooke, 1996] from a industrial standardize point of view. It measures the usability, simplicity, integration, learnability and confidence of users in operating any new system. The SUS has 10 questions, making it easy for participants to understand. It is reliable even in small sample sizes, and has been widely validated. The SUS asks participants to score the following 10 items with one of five responses that range from Strongly Agree to strongly disagree:

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

SUS is calculated as follows based on the 10 items listed above (A-J):

***SUS***= X(*a* −1)+(5− *b*)+(*c* −1)+(5− *d*)+(*e* −1)+

(5− *f*)+(*g* −1)+(5− *h*)+(*i* −1)+(5− *j*)∗2*.*5

Positive and negative items are balanced by subtracting one from the score or subtracting the score from five, respectively. This scales all values from 0 to 4, with 4 being the most positive response. All scaled answers are summed and multiplied by 2.5 to convert the range of possible values to 0-100 instead of 0-40. The industry average score for evaluating the usability of a system, based on previous research[Lewis and Sauro, 2009], is between 68 and 78 out of 100.

Table 5. System Usability Scale Score Table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Statistic | N | Mean | St. Dev. | Margin of Error |
| Usability | 24 | 78.0 | 12.49 | 5.54 |
| Learnability | 24 | 58.3 | 16.7 | 12.13 |
| SUS Score | 24 | 74.1 | 10.45 | 5.96 |

We measured the internal consistency of participants’ answers to the 10-question SUS using Cronbach’s Alpha, which is 0.76 for these data. This reflects high internal consistency between all 24 respondents. We also calculated results based on the above formula to get 3 objective scores for the digital system: Usability Score, Learnability Score, and finally the overall System Usability Score (SUS) score[Brooke, 1996]. Based on the results, we report that usability score of the total respondent was 78.0, however the score for learnability is 58.3. We believed that users are still not accustomed to the setup for the first time and report the need for more time to familiarize themselves with the system. Finally, our overall score for the system usability scale (SUS)was 74.1, Comparing the industry standard of 68[Lewis and Sauro, 2009], we can report that the system has a very high usability score and should be considered a strong potential alternative to traditional methods in supporting groups in collaborative affinity diagram activities.

# CONCLUSIONS

Overall, exchanging and organizing ideas using the Distributed Affinity Diagram System (DADS) is generally better received by users in terms of usability, compared to analog methods. We also found that users are more productive based on the number of sources and comments created in the digital system. However, despite this preference among users, our research also shows that users’ behavior in the analog system is less varied, reflecting their familiarity with pen-and-paper collaborative work. Our research has clarified the usability and efficiency of the design of DADS, and our correlation analysis suggests new hypotheses that can inspire further research in the field of e-Collaboration. Below we present what we have learned from this study, and what our work means for future groupware researchers and system designers.

In some ways, DADS did exactly what we expected it to do by making user interactions with the system easy and efficient. Specifically, our system increased the efficiency of content editing and navigation, improved the capacity for supportive information sources, made presentation easier, organized and documented comment exchanges, and increased users’ flexibility in collaborating on an affinity diagram. However, we also found some results that we did not expect:

1. Teamwork is not increased by the digital system. Users did not report different levels of teamwork in the digital platform compared to the analog. We believe that doing work on separate screens might be an issue for some users, but not others. We think this discovery is worth further research.
2. The number of attachments does not correlate with users’ level of understanding. Despite our hypothesis, we discovered that for increasing understanding, the number of attachment available is less important than the ease of presentation. This result leads us to propose a theory that clear presentation, not more information, is the best way to increase users’ understanding.
3. In the digital platform, increased movements during affinity diagram creation is negatively correlated with level of comfort. This means that the more moves a user makes, the more uncomfortable they are. However, the number of movements may also raise a user’s level of satisfaction, which means that even though it is less comfortable to engage in more movement, the users felt more satisfied with the final results.

Understanding the content is of paramount importance in any discussion, so we propose that future research should focus on two things: 1) designing a user interface that is easy to navigate between users’ points, and 2) presentation features that support better explanation and understanding. These two functions are important to consider when building future systems. We also suggest that scholars do further research on the relationship between distributed screens and teamwork. We believe that the digital system’s individual screens might prevent higher levels of perceived teamwork, so future experiments will test different screen configurations to see if the single-screen collaborative brainstorming systems promote more teamwork. Future research can explore which parts of distributed interaction systems affect teamwork most.

Digital collaboration systems have already addressed some of the efficiency and usability problems that analog tools have. DADS is an integrated system for digital collaboration that builds on previous work to improve a common way that teams share ideas, making communication more efficient during affinity diagram creation. Our research into the usability of this system provides insight into the way distributed workspaces can improve collaborative discussion by making it easier to share ideas, easier to present and support these ideas with accessible material, and increase satisfaction with the outcome of the discussion. The insights in our detailed usability study can provide new and fruitful directions for researchers in computer-mediated interaction to explore.

# ACKNOWLEDGMENTS

We would like to extend our gratitude to Tohoku University for providing laboratory facilities, various open-source communities that supported the software development, numerous volunteers that participated in the testing phase by giving constructive feedback to improve system, and the Japanese Ministry of Education, Culture, Sports, Science and Technology for government funding and scholarship.

# REFERENCES

Photon Engine exit games.

[**https://www.exitgames.com/en/Realtime**.](https://www.exitgames.com/en/Realtime) Accessed: 2014-04-12.

Ballendat, T., Marquardt, N., and Greenberg, S. (2010). Proxemic interaction: designing for a proximity and orientation-aware environment. In *ACM International Conference on Interactive Tabletops and Surfaces*, pages 121–130. ACM.

Benford, S., Bowers, J., Fahlen, L. E., Mariani, J., and Rodden, T. (1994). Supporting cooperative work in virtual environments. *The Computer Journal*, 37(8):653–668.

Brooke, J. (1996). Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189:194.

Bross, I. D. (1971). Critical levels, statistical language and scientific inference. *Foundations of Statistical Inference*, pages 500–513.

Clayphan, A., Collins, A., Ackad, C., Kummerfeld, B., and Kay, J. (2011). Firestorm: a brainstorming application for collaborative group work at tabletops. In *Proceedings of the ACM international conference on interactive tabletops and surfaces*, pages 162–171. ACM.

Delbecq, A. L. and Van de Ven, A. H. (1971). A group process model for problem identification and program planning. *The Journal of Applied Behavioral Science*, 7(4):466–492.

Diehl, M. and Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of personality and social psychology*, 53(3):497.

Gallagher, M., Hares, T., Spencer, J., Bradshaw, C., and Webb, I. (1993). The nominal group technique: a research tool for general practice? *Family Practice*, 10(1):76–81.

Gallupe, R. B., Dennis, A. R., Cooper, W. H., Valacich, J. S., Bastianutti, L. M., and Nunamaker, J. F. (1992). Electronic brainstorming and group size. *Academy of Management Journal*, 35(2):350–369.

Gartrell, M., Beach, A., Ramanarayanankrishnaniyer, J., Xing, X., Han, R., Lv, Q., Mishra, S., and Seada, K. (2010). Integrating wikipedia and facebook context into collaborative e-brainstorming. Technical report, Technical Report CU-CS-1073-10, University of Colorado at Boulder.

Geyer, F., Pfeil, U., Hochtl, A., Budzinski, J., and Reiterer,¨ H. (2011). Designing reality-based interfaces for creative group work. In *Proceedings of the 8th ACM Conference on Creativity and Cognition*, CC ’11, pages 165–174, New York, NY, USA. ACM.

Ghosh, K. and Aiken, M. (2013). An empirical investigation of two group idea generation techniques: Manual versus electronic gallery writing. *International Journal of e-Collaboration (IJeC)*, 9(2):61–77.

Grudin, J. (1994). Cscw: History and focus. *COMPUTER-LOS ALAMITOS-*, 27:19–19.

Gutwin, C. and Greenberg, S. (2000). The mechanics of collaboration: Developing low cost usability evaluation methods for shared workspaces. In *Enabling Technologies: Infrastructure for Collaborative Enterprises, 2000.(WET ICE 2000). Proeedings. IEEE 9th International Workshops on*, pages 98–103. IEEE.

Holtzblatt, K., Wendell, J. B., and Wood, S. (2005). *Rapid contextual design: a how-to guide to key techniques for user-centered design*. Elsevier.

Jeners, N., Prinz, W., and Franken, S. (2013). A meta-model for cooperation systems. In *Collaborative Systems for Reindustrialization*, pages 239–246. Springer.

Judge, T. K., Pyla, P. S., McCrickard, D. S., Harrison, S., and Hartson, H. R. (2008). Studying group decision making in affinity diagramming. Technical report, Citeseer.

Kawakita, J. (1991). The original kj method. *Tokyo: Kawakita Research Institute*.

Keary, A. and Redfern, S. (2012). Future directions of the conferencing and collaboration field. *International Journal of e-Collaboration (IJeC)*, 8(2):47–70.

Kendon, A. (2010). Spacing and orientation in co-present interaction. In *Proceedings of the Second International Conference on Development of Multimodal Interfaces: Active Listening and Synchrony*, COST’09, pages 1–15, Berlin, Heidelberg. Springer-Verlag.

Kock, N. (2005). What is e-collaboration. *International Journal of E-collaboration*, 1(1):1–7.

Lewis, J. R. and Sauro, J. (2009). The factor structure of the system usability scale. In *Human Centered Design*, pages 94–103. Springer.

Miura, M., Sugihara, T., and Kunifuji, S. (2011). Gkj:

Group kj method support system utilizing digital pens. *IEICE transactions on information and systems*, 94(3):456–464.

Mugge, R. and Schoormans, J. P. (2012). Product design and apparent usability. the influence of novelty in product appearance. *Applied ergonomics*, 43(6):1081–1088.

Munemori, J. and Nagasawa, Y. (1996). Gungen: groupware for a new idea generation support system. *Information and Software Technology*, 38(3):213–220.

Myers, J. L. and Well, A. D. (2003). *Research Design and*

*Statistical Analysis*. Lawrence Erlbaum Associates, New Jersey.

Nijstad, B. A., Stroebe, W., and Lodewijkx, H. F. (2003). Production blocking and idea generation: Does blocking interfere with cognitive processes? *Journal of Experimental Social Psychology*, 39(6):531–548.

Paul, S. A. and 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, pages 1771–1780, New York, NY, USA. ACM.

Radle, R., Jetter, H.-C., and Reiterer, H. (2013).¨ Twistersearch: A distributed user interface for collaborative web search. In *Distributed User Interfaces: Usability and Collaboration*, pages 53–67. Springer.

Ramakers, R., Vanacken, D., Luyten, K., Coninx, K., and Schoning, J. (2012). Carpus: a non-intrusive user¨ identification technique for interactive surfaces. In

*Proceedings of the 25th annual ACM symposium on User interface software and technology*, pages 35–44. ACM.

Tse, E., Greenberg, S., Shen, C., Forlines, C., and Kodama, R. (2008). Exploring true multi-user multimodal interaction over a digital table. In *Proceedings of the 7th ACM Conference on Designing Interactive Systems*, DIS ’08, pages 109–118, New York, NY, USA. ACM.

Widjaja, W., Yoshii, K., Haga, K., and Takahashi, M. (2013). Discusys: Multiple user real-time digital sticky-note affinity-diagram brainstorming system. *Procedia Computer Science*, 22:113–122.

Williams, K. D., Nida, S. A., Baca, L. D., and Latane, B.´ (1989). Social loafing and swimming: Effects of identifiability on individual and relay performance of intercollegiate swimmers. *Basic and Applied Social Psychology*, 10(1):73–81.