Intent-Gesture Relationships for Collaborative Information Visualization

by

Ravina Samaroo

A thesis submitted to the Faculty of Graduate and
Postdoctoral Affairs in partial fulfillment of the requirements for
the degree of

Master of Arts

in

Human-Computer Interaction

Carleton University
Ottawa, Ontario

© 2015, Ravina Samaroo

Abstract

In this study we look at the relationship between gestures and intents when pairs of participants are collaborating around a large display with a graph. We aimed to find out what gestures paired with which intents, which gestures participants would find suitable for various intents, and how our findings could influence designing interactions with graphs being used for collaborative analysis work. We studied 8 pairs of participants and found 10 frequent gestures and 11 frequent intents. An exploration of the relationship between these gestures and intents found 15 frequent co-occurrences. We analyzed these findings and then proceeded to make design suggestions for enabling co-located collaboration interaction using large multi-touch displays. Throughout, we used a theory of technical intersubjectivity to guide our research. In particular, this helped us to position large multi-touch displays as enablers of intersubjective interactions, which facilitated our design process.

Acknowledgements

I would like to thank my supervisors Robert Biddle and Judith Brown for making the completion of this thesis possible. Your time and detail are very much appreciated. Thank you for all of your hard work and for the extensive amount of knowledge that you both have bestowed upon me.

I would also like to also thank the members of my committee: Sonia Chiasson, Kasia Muldner, and the Chair, Lois Frankel for your time and feedback.

Lastly, thank you to my parents for their support as I would not be in this position without them.

Table of Contents

Abstract.....	ii
Acknowledgements	iii
Table of Contents	iv
List of Tables	vii
1 Chapter: Introduction	13
1.1 Motivation	13
1.2 Research Agenda.....	17
1.3 Objective	18
1.4 Contributions	19
1.5 Outline:.....	20
1.6 Publications:	21
2 Chapter: Background.....	22
2.1 Collaboration	22
2.1.1 Technology-enhanced collaboration	24
2.2 Human gestures	27
2.2.1 Deixis	38
2.3 Natural user interfaces.....	41
2.4 Touch interaction on a large surface.....	47
2.4.1 Beyond touch interaction on the screen	50
2.5 Background summary and connections to our study	52
3 Chapter: Study Design	57
3.1 Developing tasks for our study.....	57
3.2 Participants	62

3.3	Apparatus.....	62
3.4	Materials	62
3.4.1	Parallel coordinates graphs	63
3.4.2	The parallel coordinates graph in our study	64
3.5	Procedure.....	66
3.6	Analysis Plan.....	68
4	Chapter: Analysis.....	70
4.1	Fine grained analysis	70
4.2	Open coding process.....	70
4.3	Axial coding process	75
4.4	Gestures	77
4.5	Intents	83
4.5.1	Emergent versus prompted intents	84
4.5.2	Intent coding granularity	84
4.6	The relationships between gestures and intents.....	86
4.7	Important take-away	90
5	Chapter: Design Approach	92
5.1	Design patterns	93
5.2	Outline of chapter.....	94
5.3	The problem with collaborating on an interactive screen.....	95
5.4	Interaction Techniques & Collaborative Design Strategies for the Intents	96
5.5	Intent relationships	110
5.6	Summary	111
6	Chapter: Discussion and Conclusions.....	113
6.1	Summary of our study	113

6.2	Explaining the co-occurrences from our analysis.....	115
6.3	Design Concerns.....	118
6.3.1	Concerns regarding our collaborative design strategies.....	121
6.4	Limitations and future work	123

List of Tables

<i>Table 1.</i> Yi et al.’s (2007) intent categories.....	58
<i>Table 2.</i> Frequency of use of Yi et al.’s intent categories in our list of study tasks	59
<i>Table 3.</i> Length in time of each video analyzed in this study.....	70
<i>Table 4.</i> Gesture Categories (including number of occurrence in brackets) with their respective gestures.	77
<i>Table 5.</i> Intent categories and their respective intents	83
<i>Table 6.</i> Comparison of Yi et al.’s (2007) intents to ours. Multiple of our intents fit into one of theirs.....	85
<i>Table 7.</i> Co-occurrence table depicting the frequency with which each gesture is performed to achieve each intent	86
<i>Table 8.</i> List of gestures with corresponding intents for those intents that were prompted from the hypothetical tasks	89
<i>Table 9.</i> Categorization of our gestures with Lao et al.’s (2009) three basic hand movement types. Hand motions and Pointing are not included (from our gesture set) as they are not gestures performed on the screen. 2-person gestures fall into both “pressing” and “dragging” because this category usually involved a combination of placing fingers on the screen and tracing lines.	117

Table of Figures

<i>Figure 1.</i> Large interactive display used to analyze the digital X-ray of a small gas turbine. Source: Wypych et al., 2011.....	14
<i>Figure 2.</i> ACH-Walkthrough running on a large multi-touch screen. Source: Wilson et al., 2014.....	16
<i>Figure 3.</i> Parallel Coordinates visualization for the ACH analysis. Source: Wilson et al., 2014.....	17
<i>Figure 4.</i> Okada's hierarchical collaboration model. Source: Grudin & Poltrock, 2012..	22
<i>Figure 5.</i> Johansen's Four-Square model of collaboration types based on different and same times and places. Source: Johansen, 1991.	23
<i>Figure 6.</i> Johansen's Four-Square model with examples provided for each square. Source: Grudin & Poltrock, 2012.....	24
<i>Figure 7.</i> Jonathan Grudin and Steven Poltrock's Model of Technology Development and Use. Source: Grudin & Poltrock, 2012.	24
<i>Figure 8.</i> Three basic hand movements when gesturing on a screen: pressing, tapping, and dragging. Source: Lao et al., 2009.	30
<i>Figure 9.</i> Five types (or elements) of dragging gestures on a screen. From left to right: Moving fingers close to each other, moving fingers apart from each other, moving fingers in the same directions, using one finger to press while the other drag, and lastly, dragging multiple fingers bi-directionally. Source: Lao et al., 2009.....	31

<p><i>Figure 10.</i> General model for touch interactions, showing what types of surface gestures are associated with various user motivations for a single-user photo manager application. Source: Lao et al., 2009.....</p>	32
<p><i>Figure 11.</i> Microsoft Surface table top display and a user interacting with it using her hand. Source: Wobbrock et al., 2009.....</p>	34
<p><i>Figure 12:</i> The user-defined gesture set. For example, the first picture depicts that a user would typically drag something off-screen to delete it. Source: Wobbrock et al., 2009.....</p>	35
<p><i>Figure 13:</i> Four group members perform the “enlarge” gesture by simultaneously touching the corners of a photo. Source: Morris et al., 2006.....</p>	36
<p><i>Figure 14.</i> (a) The “throw-and-receive” gesture is one technique for passing photos long distances. (b) The</p>	37
<p><i>Figure 15.</i> Various methods of "directing-to"; for example, by using one's head to nod at something. Source: Clark, 2003.....</p>	40
<p><i>Figure 16.</i> Blended Interaction Framework. Blended interaction helps to explain whether users perceive a UI as natural or not. It includes a combination of a user’s known concepts of reality, and the computational power of the system. There are four domains of design of blended interaction: individual interaction, social interaction, workflow, and physical environment, as seen at the bottom of the diagram. Source: Jetter et al., 2014.</p>	45
<p>Figure 17. A parallel coordinates plot that measures several aspects of U.S. counties such as Home Value. Here, Alameda County in the state of California has been highlighted in green. Source: (Few, 2006).....</p>	64
<p>Figure 18. Our parallel coordinates graph depicting arrest rates by age and offense categories in the U.S. Offense categories are displayed on the left axis, arrest rates are displayed</p>	

on the middle axis, and age is displayed on the right axis. The color scheme is as follows:
aggravated assault in blue, burglary in red, driving under the influence in yellow, drunkenness in
green, larceny (theft) in purple, vandalism in orange, and weapons in cyan..... 65

Figure 19. Prior to beginning the study, participants were presented with this slide that
explained how a parallel coordinates graph worked. They were provided with additional verbal
explanations by the researcher and were allowed to ask questions for clarification. 67

Figure 20. The first task/question asked of participants in our study, displayed from a
Power Point presentation. 68

Figure 21. Creating a code in Atlas. On the left is the video, in the middle is the dialogue
box for creating a code, and on the right is the timeline where codes have already been associated
with segments of the video. In this section of the video, we can see approximately 70 codes. ... 72

Figure 22. The final codes on a section of one video. This is similar to *Figure 21* except
the codes have been dragged over the video section of the page, allowing one to see a greater
number of codes that were off to the right hand side. In this image, we can see about 120 codes,
or about double the amount that can be seen in *Figure 21*. 73

Figure 23. List of codes. This is referred to as the “Code Manager” in Atlas, and displays
an alphabetical list of every code used within all of the coded videos. This is a snapshot of about
32 of the 234 codes. 74

Figure 24. Creating and sorting the Code groups (what Atlas refers to as Family codes).
On the top is the list of group codes. As seen in this figure, selecting the code group “Hand
Motions” shows a list of codes that are in the group (on the bottom right), and the rest of codes
that are not in the group (on the bottoms left). 76

<i>Figure 25.</i> Code Groups in the “Code Manager”. On the left side is a list of all code groups (families) and on the right are all of the codes within the selected group, “Collaborative Gestures” (later renamed to 2-person gestures).	77
<i>Figure 26:</i> Example of the 2-person gesture "Both placing fingers on screen"	78
<i>Figure 27:</i> Example of the gesture "Dragging an item". The participant touches the bottom offense (Weapons) and drags it up the axis to the top.	79
<i>Figure 28:</i> Example of the hand motion "Moving hand up and down the graph". Here, the participant moves his entire hand up and down the middle axis.	80
<i>Figure 29:</i> Example of the “Pointing” gesture.	80
<i>Figure 30:</i> Example of the gesture “Spreading index fingers apart from each other”.	81
<i>Figure 31:</i> Example of the gesture “Swiping motion with fingers” (from left to right on the screen).	81
<i>Figure 32:</i> Example of the gesture “tracing line with index finger”. The participant on the left hand side starts tracing a line from the left of the screen, following it through to the right of the screen.	82
<i>Figure 33.</i> Example of the pattern, "Breadcrumbs", including a description of what the pattern is and when to use it. Source: Tidwell (2011).	94
<i>Figure 34.</i> Example of using arrows to show the direction the graph is being panned. The further away the arrow is from the fingers, the more ghostly it becomes until it disappears.	101
<i>Figure 35.</i> Example of a ghostly presence being left behind of the deleted item “Burglary” at the top of the graph.....	102
<i>Figure 36.</i> Example of the blinking circle when selecting a new point to placehold.	108

Figure 37. Example of encircling an item when it is selected to gain more information
about it. 109

1 Chapter: Introduction

1.1 Motivation

Collaboration in the work environment involves multiple people working together towards a common goal (Buttler, Janeiro, Lukosch, & Briggs, 2011). With the introduction of large, multi-touch screens, we can expect collaboration to evolve around these technological advances.

One popular field of research deals with *collaborative visualizations*, which “enhance the traditional visualization by bringing together many experts so that each [expert] can contribute toward the common goal of understanding the [joint] object, phenomenon, or data under investigation” (Raje, Boyles, & Fang, 1998, p. 1079). Isenberg, Elmqvist, Scholtz, Cernea, Ma, and Hagen (2011) define collaborative visualization as “the shared use of computer-supported, (interactive) visual representations of data by more than one person with the common goal of contributing to joint information processing activities” (p. 312).

One example of collaborative visualization has been demonstrated in an application that involves the fields of structural, mechanical and aerospace engineering, where Wypych, Yamaoka, Ponto, and Kuester (2011) looked at the inspection of high-resolution image data within a custom-built X-ray radiography system. X-ray images are analyzed by these professionals to classify the health or structure of a particular specimen during quality or state-of-health assessments. However, the improving resolution of images from these devices, which produces significantly more data from these machines, has not resulted in equivalent high-resolution methods for cataloging or visualizing the data. Data analysts are in need of a way to effectively keep up with the rapidly expanding rate of data production in order to inspect, manage, and analyze the results in an efficient manner. Wypych et al. (2011) have developed a

system in light of this concern called HIPerSpace, a 286 mega-pixel resolution display made up of 70 networked 30-inch LCD display panels, spanning 9.9 meters in width and 2.6 meters in height (see *Figure 1*). It provides analysts with a large surface to view X-ray images, and incorporates a variety of visualization features to support the analysis process, such as being able to manipulate, annotate, and markup X-ray images, while being able to view very fine details that accurately reflect the higher resolution of the images.

Figure 1. Large interactive display used to analyze the digital X-ray of a small gas turbine. Source: Wypych et al., 2011.

Isenberg, Fisher, Paul, Morris, Inkpen, and Czerwinski (2012) discuss the paradox that even though many visual analytics tasks are best solved in groups working face-to-face (where they can, for example, challenge each other immediately with facts), there are not many technologies to support this type of co-located visual analytics. They studied pairs working together on a tabletop system to solve the VAST 2006 Challenge, an intelligence analysis task exploring a large collection of text documents. Features of the system allowed participants to share and organize information easily, and the use of this display allowed for easy face-to-face interactions. The ability to share and to have face-to-face communication enhanced

collaboration. Participants could easily point to and manipulate documents that their partners could also see, as they did not need to turn away from the table in order to speak face-to face. One participant mentioned that the ability to share the display diminished any fighting over who gets to use the computer at any time, therefore leading to a collective feeling of participation and group awareness of the progress of the analysis by all members. In their study of this technology, Isenberg et al. found that participants who collaborated more performed better on the task, as determined by their overall score. The authors suggest that when designing collaborative systems, there is a high value on close collaboration, and therefore these systems should include awareness features that bring teams closer together by encouraging conversation about the content of the analysis. These features “could make obvious visual connections between common information that team members are reading, that are placed in close proximity, or that have been stored in the workspace” (p.186).

More recently, Wilson, Brown, and Biddle (2014) examined how large interactive surfaces can support collaborative intelligence analysis. They drew from an analysis technique called ACH, or *Analysis of Competing Hypotheses*, where analysts consider and disprove multiple hypotheses using evidence and assigning each hypothesis a score based on this evidence. The process aims to reduce cognitive biases. Their previous research suggested that collaborative support would be beneficial during an ACH session, as the process would be assisted by team discussions, and would help to further reduce cognitive biases. Making the process of ACH collaborative allows for a broad spectrum of perspectives to be incorporated, enhancing the identification of critical issues. They named their system ACH-Walkthrough (see *Figure 2*), as it is based on a collaborative practice called cognitive walkthroughs developed for interface designers (Wharton, Bradford, Jeffries, & Franzke, 1992), where members of a team

walk through (discuss and execute) the various analysis steps together, with each member contributing from their own perspective.

Figure 2. ACH-Walkthrough running on a large multi-touch screen. Source: Wilson et al., 2014.

There is a significant visualization component to ACH-Walkthrough, as Brown, Wilson, and Biddle's (2014) research prior to this had shown that there was a need amongst the analysts to consider the overall patterns in rating evidence, and in scoring of hypotheses within the analyst community. They supported this need by designing a visualization of the ACH data using a visualization technique for displaying data called *parallel coordinates* (Inselberg, 1997), as this is an established technique to support the exploration of multi-dimensional data. An example of the parallel coordinate's visualization can be seen in **Figure 3**. Each line represents one piece of evidence and the location where the line crosses a line represents a value on one of the four following axes: credibility, relevance, hypothesis one, and hypothesis two. Following a single line allows the user to see ratings for that piece of evidence. The overall pattern of lines provides a general sense of the strength of the analysis as a whole.

Figure 3. Parallel Coordinates visualization for the ACH analysis. Source: Wilson et al., 2014.

1.2 Research Agenda

One can see that research on collaborative visualizations is in its infancy, but that many data analysis tasks are enhanced by collaboration. We can further see that many researchers are exploring the use of large screens as a means for displaying collaborative visualizations, as they provide the physical space that is necessary for collaboration around visualizations to occur.

While we can see that there is considerable research on collaborative visualizations, there are no studies that we know of that examine the implications that a large vertical screen displaying an interactive graph has on collaboration (i.e., working together towards common goals). There is a lack of research on how people collaboratively interact with graphical information in order to achieve the various lower-level intentions that they create in the process of achieving their joint goal. We are therefore interested in examining the field of collaborative visualizations, specifically looking at the types of gestures that people use to complete various intentions while collaborating on a large display with a graphical visualization. Intentions can be seen as the smaller steps towards achieving a common goal or towards accomplishing a shared

task (for example, rearranging variables on a graph to achieve an ordering that is more suited to the larger goal at hand).

We are further interested in exploring the possible design implications stemming from the information provided from the gesture-and-intent relationships we may observe. Tang (1991) stresses the importance of re-assessing the design assumptions that are made when designing tools for individual use, when designing tools for collaborative use, as assumptions in the former situation need to be challenged in the latter situation. This is necessary in order to assess the different requirements of a group using a tool collaboratively, versus the requirements of an individual user using a tool on his/her own. When designing collaborative technology, designs should stem from an understanding of how collaborative work is accomplished by users. By understanding what resources the collaborators use and what obstacles and limitations they encounter as they work, tools can be designed to enhance the available resources and decrease any complications within the collaborative activity. Tang proposes that research done to observe and understand collaboration when engaged in an activity is an important step towards guiding future designs to support collaborative activity. Drawing from this, we precede our design suggestions with an observation phase. In this phase we observe pairs of users collaboratively using a graphical visualization in order to gain insights into what gestures they perform when working together on a large screen. This will provide us with valuable evidence-based data for gesture design suggestions in the context of collaborative visualization analysis.

1.3 Objective

We aim to assess how pairs of participants collaborate together on a large vertical touch screen while analyzing a graphical visualization. The type of graphical visualization that we will use for our study is a parallel coordinates' graph. Parallel coordinates “transform multivariate

relations into 2D patterns” (Inselberg, 1997, p. 100). These types of graphs make comparison and pattern analysis visually appealing and efficient when dealing with moderately-sized datasets.

Although they are not ubiquitous, they are interesting, can present a lot of data, and can be used in multiple domains. They will be explained in more detail in the materials section of chapter 3.

In this study, we are exploring how gestures relate to user’s intentions while analyzing a parallel coordinates plot in a collaborative setting. While there has been an increase in the number of studies looking at collaboration with visualizations, there have not been any particular studies that we know of that look specifically at how gestures relate to intents when collaborating on large displays, and how knowledge of these gesture-intent relationships can affect the design of collaborative systems. Our research questions are as follows:

- 1) Which gestures are related to which intents?
- 2) Which gestures do people believe to be suitable for different intents?
- 3) How can knowledge of these gesture-intent relationships influence design ideas for interactive collaboration with a parallel coordinates graph on a large touch screen?

The goal is to determine what gesture-intent pairs will emerge when people analyze the parallel coordinates data together; and furthermore, to determine designs that support the gesture-intent pairs that are natural to collaborators.

1.4 Contributions

The contributions made by this study are as follows: We created a study design that provides evidence-based results for designing gestures to support collaborative analysis work. We will show that our study found a clear link between various gestures and intents. Some of these relationships between gestures and intents will add to the literature as they have not been identified previously. Due to our focus on a graphical visualization, we identified gestures that

are more specific to graphical data analysis, which has not been studied in detail previously. We will demonstrate how this knowledge of gesture and intent pairings for graphical visualizations can be leveraged to design gestures that are fitting to a collaborative task analysis on large vertical touch displays.

1.5 Outline:

- Chapter 2: Background Information. This chapter details related work relevant to our study in 4 different areas (collaboration, human gestures, natural user interfaces, and touch interaction on large screens).
- Chapter 3: Study design. This chapter details how the study was designed and implemented, along with general information about the participants of the study.
- Chapter 4: Results and Analysis. This chapter describes the analysis process, and presents the findings of our study.
- Chapter 5: Design. This chapter suggests design recommendations and ideas for the design of gestures for a collaborative analysis activity around a shared parallel coordinates' visualization, and is based on our results.
- Chapter 6: Discussion/Conclusion. This chapter sums of the major findings and relates them to previous work by other researchers, looks at the higher level generalizability of our results, and provides limitations of our study and possible directions for future work.

This paper is written in first person, where “we” refers to Ravina Samaroo, under the advice of Robert Biddle and Judith Brown.

1.6 Publications:

Publications from our work to date relating to this thesis include the paper *The day-in-the-life scenario: A technique for capturing user experience in complex work environments* (Samaroo, Brown, Biddle, & Greenspan, 2013).

For this CEWIT paper, we developed a visualization for people studying behavior in operation centers. This type of visualization was designed to be displayed on a large interactive screen in order for software developers to view and analyze the attention levels of the operators that they are designing products for. It is a type of visualization which software developers and designers might collaborate around in order to expose and work through their own design issues.

2 Chapter: Background

This chapter looks at background work related to various topics that this thesis covers. We review the topics of collaboration, human gestures (with a special focus on deixis), natural user interfaces, and touch interaction on a large surface.

2.1 Collaboration

Buttler, Janeiro, Lukosch, and Briggs (2011) define collaboration as multiple people working together towards a common goal. The main benefit they highlight is that one person may not have enough resources, skills, knowledge, or time to reach the goal on their own.

Figure 4. Okada's hierarchical collaboration model. Source: Grudin & Poltrock, 2012.

Okada (2007 in Japanese; as cited in Grudin & Poltrock, 2012) proposed a more detailed, hierarchical framework of collaboration, as seen in *Figure 4*. Okada's hierarchical collaboration modelIn his framework, collaboration has two dimensions: assertion and cooperation. Collaboration is strongly influenced by the degrees of assertion and cooperation exhibited by participants. In his view, low levels of both assertion and cooperation result in compromise; high levels of both result in coordination; more assertion than cooperation results in collision; and

more cooperation than assertion results in concession. This 'layer' of collaboration is supported by a sharing level: collaborators share views and opinions, knowledge and information, and work and operations. The sharing layer is supported by an awareness level. Awareness includes awareness of other collaborators, the environment where the work occurs, and the objects and tools that are used. Awareness in turn is affected by the coexistence layer, which depends on temporal and spatial factors.

Johansen (1991) outlined 4 different types of collaboration options where interaction can occur synchronously or asynchronously, and may be collocated or distributed (see *Figure 5* and *Figure 6*):

- 1) Same time, same place (e.g. collaborating in a meeting room)
- 2) Same time, different place (e.g. collaborating remotely from different countries)
- 3) Different time, different place (e.g. collaborating through email)
- 4) Different time, same place (e.g. collaborating through shift work)

Figure 5. Johansen's Four-Square model of collaboration types based on different and same times and places.

Source: Johansen, 1991.

		Place	
		Same	Different
Time	Same	Electronic meeting room	Application sharing
	Different	Shift work, team rooms	Email, newsgroups

Figure 6. Johansen's Four-Square model with examples provided for each square. Source: Grudin & Poltrock, 2012.

Johansen's model inspired researchers to assess the temporal and spatial dimensions that influence collaboration and the different ways group activities can be supported in these four separate elements.

2.1.1 Technology-enhanced collaboration

Grudin and Poltrock (2012) describe a model of technology development and use within the field of computer supported cooperative work—CSCW (A core premise in CSCW is that technology can increase the effectiveness and the experience of collaboration). Their model (See *Figure 7*) describes the basic process of how technology becomes adopted in a collaborative setting. At the top of the figure, the process begins with people collaborating, whether using technology or not. Moving along clockwise on the diagram, the next step is to determine the user needs for the technology that people will use to collaborate. Next, a technology investigation must be conducted to find the needed technology (if it already exists), or to design a new piece of technology if it doesn't already exist. Once this phase is completed, the technology can then be

Figure 7. Jonathan Grudin and Steven Poltrock's Model of Technology Development and Use. Source: Grudin & Poltrock, 2012.

created. It is then deployed to the employees, or marketed to the desired population. In the final stage, people can adopt the technology and choose to use it. This brings us back around to the top of the diagram. When collaborating, people may or may not use the technology as intended, and over time their use evolves as the technology is understood better or is used alongside other new technologies or processes. Then it may be time to start the process all over again, by examining the new needs of the people collaborating.

With collaboration there are challenges for group members to create a common understanding, and more generally, it is hard for designers to devise methods of interaction that allow for effective collaboration (Buttler, Janeiro, Lukosch, & Briggs, 2011). These challenges of collaboration are further enhanced and complicated by the use of technology, as the technology is another aspect of the process of collaboration that needs to be carefully assessed. Vatrapu and Suthers (2009) examine the interconnections and dynamics between persons and the technology they are using while collaborating, and they discuss designing for human-to-computer interaction, as well as designing for human-to-human interaction. They use the term *technological intersubjectivity* (TI).

Intersubjectivity is communication and participation by two or more people — it is the fabric of social interactions. This term was defined similarly by Suthers although he called it intersubjective learning (2006). According to Suthers, *intersubjective learning* is the sharing of interpretations and information between collaborators, where the interpretations can be *jointly* created through interaction. Suthers argued that learning and the process of meaning-making are accomplished through the social interactions of the participants, and not simply through information sharing. However, Suthers' definition provides no role for technology in this social process.

Vatrapu and Suthers (2009) address this issue. They define *technological intersubjectivity*, as:

A technology-supported interactional social relationship between two or more people. It forms from a dynamic interplay between the technological relationship of participants with artifacts and their social relationship with others. (p. 156)

Technological intersubjectivity views technology as a medium through which our interactions with other participants are informed – i.e., the technology potentially *enhances* our communication. When technology is used to support collaboration, the interface becomes another entity in the social interaction equation.

Just as there are challenges in non-technologically supported collaboration in creating common understandings between all members, and allowing for effective interaction methods, these challenges are also applicable to the interface in the social equation. The interface must therefore be designed to facilitate person-to-person interactions in a way that enhances collaboration, while also displaying shared information effectively, and enabling the achievement of a common understanding by all participants.

In a more detailed analysis of the topic, Suthers (2006) proposes that within computer supported collaborative learning (CSCL), there should be a focus on the design and study of social interactions that are informed by the affordances and limitations of technology. Intersubjective learning requires CSCL systems to be social in nature: the technology should mediate and encourage social interactions that establish learning (therefore intersubjective learning has technological requirements similar to those of collaborative analysis). This should be done through designs that draw on the unique capabilities of the technology, rather than trying to fit the technology into a specific mold for which it is not well suited. One way technology can

be used to implement support for collaborative learning is through the scripting of interactions (detailing the smaller components observed in an interaction); for example, by observing people as they go through the natural phases of collaboration, and describing their interactions at a detailed level, in order to create guidelines for making and evaluating technological applications that support social interactions. This detailing work may be done through videos and transcripts of participants to discover the methods they use to accomplish learning/meaning-making. This is a data-driven, often micro-analytic approach that involves examining short episodes of videos in fine detail, with the objective of identifying significant interactions and determining theoretical groupings in this data. Overall, the author concludes that the study of technological affordances (in collaborative settings) should be constantly done with reference to the activity that it is trying to support—i.e., intersubjective meaning making or learning.

2.2 Human gestures

When humans interact with each other, we are constantly using visible bodily movements to inform others of our feelings, intentions, focus of attention, and ideas; such as pointing to indicate to others to shift their focus of attention to what is being pointed at (Kendon, 2004). These gestures may be used simultaneously with speech, or they may complement, or substitute speech; such as a head nod to indicate agreement without saying “yes.” Kendon defines *gesture* as:

a label for actions that have features of manifest deliberate expressiveness. They are those actions that, or those aspects of another’s actions that, having these features, tend to be directly perceived as being under the guidance of the observed person’s voluntary control and being done for the purposes of expression rather than in the service of some practical aim. (p. 15)

In this definition of gesture, the intentionality of a gesture is completely dependent on the situation and the people involved, as they are the ones who interpret the flow of actions (i.e., what is seen as a gesture versus an incidental movement).

Kendon further discusses gestures in terms of their components, or what he refers to as gesture units. A gesture unit consists of a combination of excursions, or movements of the body (i.e., articulator) to complete the gesture. For instance, when waving a hand, the excursions are as follows: first, the hand moves from a rest position towards a region of space, and then it is moved back to a position of rest. Broken down even further, the movement of the hand away from the rest position is labelled as *preparation*, while the movement nearest the apex of the gesture is called the *stroke*, and signifies the expression of the gesture being accomplished. In some cases, the articulator (for example, the hand) is maintained in the position from the end of the stroke, which is labeled as the *post-stroke hold*. Finally, the end phase where the rest state is returned to is called *recovery*. In the hand waving example, preparation would be lifting the hand up, the stroke would be waving the hand, a post-stroke hold would be leaving the hand in the wave position shortly after the wave ends, and the recovery phase would be placing the hand back down.

In a more technology-specific analysis of gestures, Wu, Shen, Ryall, Forlines, and Balakrishnan (2006) propose the concepts of *gesture registration*, *relaxation*, and *reuse* as principles for gesture design on direct-touch surfaces. *Gesture registration* is the first phase of a gesture operation, and it sets the context for succeeding interactions. This is akin to Kendon's preparation phase of a gesture unit. *Gesture registration* is important when multiple interaction styles and tools are being used, such as a surface where both cursor-based and free hand-based gestures can be used. The *gesture registration* phase can be used in this situation to delineate the

transition of a tool from one interaction style to another, or from performing one intent to another. For example, a stylus can be changed from being a pointer for selecting an object, to being a writing tool, by changing the gesture registration phase. The principle of *gesture relaxation* aims to create more physically ergonomical gesture techniques. Most freehand gestures require the hand posture to remain the same throughout the gesture being performed, which imposes a burden on the user, as precise hand postures and muscular tension must be maintained. *Gesture relaxation*, on the other hand, allows gestures to be performed with minimal constraints after it is registered in the gesture registration phase. This places all of the physical tension in performing a gesture in the gesture registration phase, allowing the user to relax the shape and dynamics of the gesture afterwards. Finally, *gesture and tool reuse* refers to using the same gesture (including hand postures, finger touches, or stylus touches) to complete different intentions. Having a large set of *gesture primitives*, “the basic components that define a gesture, such as hand postures or gesture dynamics, given that a gesture can consist of continuous motions and be compounded with more than one gesture primitive” requires users to memorize all of these gestures, and also requires the system to recognize these gestures (p. 187). Reusing gesture primitives allows more gestures to be created without requiring more primitive gestures to be defined.

Lao, Heng, Zhang, Ling, and Wang (2009) proposed a comprehensive understanding of the possible types of touch used on various touch-sensitive devices by constructing a design model for single user touch interaction. In describing simple gestures, one of their major classifications is the type of movement(s) the hand makes while interacting with the screen. They describe three basic hand movements: pressing, tapping, and dragging, as seen in *Figure 8*. Pressing involves touching the surface and leaving the finger there, as seen in the right side of

Figure 8. Tapping involves touching the surface and lifting the finger right away, as demonstrated in the middle of *Figure 8*. Lastly, dragging involves touching the surface and moving the finger along the surface, as seen in the left side of *Figure 8*. Dragging can be further broken down into the 5 elements seen in *Figure 9*. In the first element (A), two fingers are dragged towards or closer to each other. In element B, two fingers are dragged away or apart from each other. In element C, fingers are dragged in the same direction (e.g., to the left). In element D, one finger may press, while fingers on the other hand may drag. Lastly, in element E, multiple fingers from both hands may be dragged bi-directionally.

Figure 8. Three basic hand movements when gesturing on a screen: pressing, tapping, and dragging. Source: Lao et al., 2009.

Figure 9. Five types (or elements) of dragging gestures on a screen. From left to right: Moving fingers close to each other, moving fingers apart from each other, moving fingers in the same directions, using one finger to press while the other drag, and lastly, dragging multiple fingers bi-directionally. Source: Lao et al., 2009.

Drawing from an extensive set of observations of touch applications on the DiamondTouch, iPod Touch, and other touch devices, Lao et al. (2009) provided a general model for touch interactions, showing what types of surface gestures are associated with various user motivations (their interest was in a single-user photo manager application). These associations are displayed in *Figure 10*. As an example, looking at the first gesture in *Figure 10* of dragging fingers apart from each other, we can see that it is associated with zooming out as well as with creating a classification. The authors aimed to allow re-use of gestures and promote consistency in user interaction across applications and devices with this model of mapped actions (gestures) and motivations (intentions).

Figure 10. General model for touch interactions, showing what types of surface gestures are associated with various user motivations for a single-user photo manager application. Source: Lao et al., 2009.

Hinrich and Carpendale looked at how interpersonal interaction is mediated by the system and examined the types of gestures explored by people using a large touch screen in a public setting (using an interactive tabletop exhibit from the Vancouver Aquarium). They conducted an ethnographic study, observing participants naturally interacting with the screen as they toured the aquarium (using video cameras set up to record their gestures). To analyze the data,

Each gesture was coded according to intended action (determined by context, previous and subsequent gestures, and visitor comments), number of hands used, hand posture (including the fingers involved and number of touch points), hand and finger movement, and the targeted

interface element (media item, button, or tabletop surface). [They] occasionally transcribed activities and verbal comments co-occurring with gestures...

At the end of their analysis, they were left with 926 gestures. They determined that intents were made up of lower level actions, of which they identified seven categories: drag/move (the most commonly performed action), enlarge/ shrink, rotate, tap, sweep, flick, and hold. To perform an action, users applied a large variety of gestures; for example, dragging/moving action was completed using single-handed, bimanual, single-finger and/or multi-touch gestures. While they found many general trends in the choice of gestures for these low level actions, they also found that the choice of gesture is influenced by social factors such as number of visitors at the aquarium as well as the social relationships between the visitors. Another interesting finding was that 67% of all bimanual gestures were mostly used symmetrically: with both hands holding the same posture, with the same fingers from each hand touching the surface, and with both hands engaging in the same action on the same object. Other times, people would use each hand to engage in different low-level actions at the same time; for example, holding a digital item in place using the left hand while also flicking through other digital items using the right hand. A different way people used their hands asymmetrically was to achieve different low-level actions on the same object. For example, digital items would sometimes slip away due to interactions by other visitors, so to prevent this while also targeting a button that provides information, people would use both hands in a combined hold and tap action: one hand holding the digital item in place while the other hand tapped the information button. This type of combined action was also commonly performed collaboratively to manipulate a single media item. They found that for these collaborative gestures, they were performed particularly to serve the group: for example, when holding a digital item, people

would use their fingertips rather than their entire hand, and also would hold the tip of the item, so as to not block other people's view of it. They suggest that when designing for multi-touch systems, social encounters should be considered, as they might have an important impact on other situations that include large direct-touch displays, such as work environments (where collaboration is common).

Wobbrock, Morris, and Wilson (2009) developed a user-defined gesture set by having 20 non-technical participants perform gestures on a Microsoft Surface table-top prototype while working at the table on their own (see *Figure 11*).

Figure 11. Microsoft Surface table top display and a user interacting with it using her hand. Source: Wobbrock et al., 2009.

Participants were shown different gesture effects or interaction techniques, such as an object moving across the table, and were asked to perform the gesture they believed would create that effect (such as tapping and sliding the object, for example). They labelled the effect of a gesture as the *referent* to which the gestural sign refers. Their gesture set was based on the participants' gesture choices: the most frequently used and agreed upon gesture would be assigned to that gesture referent. *Figure 12* depicts some of the user defined gestures from the

study. For example, the gestural sign for the referent to delete something was to drag it off of the screen.

Figure 12: The user-defined gesture set. For example, the first picture depicts that a user would typically drag something off-screen to delete it. Source: Wobbrock et al., 2009.

Morris, Huang, Paepcke, and Winograd (2006) introduced the term *cooperative gesturing* as “a multi-user interaction technique for co-located single display groupware systems. Cooperative gestures are interactions where the system interprets the gestures of more than one user as contributing to a single, combined command” (p. 1201). An example of cooperative gesturing would be if all collaborators used a swiping motion on their shared screen to delete the entire content displayed, rather than a scenario where one user swiped on a certain area to delete only that area of content. This type of gesture may be used to increase participation/collaboration

among members, and the specific coordination that is needed between the collaborators can increase group cohesion and teamwork. An example from their study is shown in *Figure 13*: Four group members perform the “enlarge” gesture by simultaneously touching the corners of a photo. where four participants touch the corners of a picture in a collaborative artistic program they called CollabDraw, to enlarge the picture.

Figure 13: Four group members perform the “enlarge” gesture by simultaneously touching the corners of a photo. Source: Morris et al., 2006.

Cooperative gestures can also help prevent accidental interactions, such as quitting an application, since all members need to perform the gesture for the intent to be completed.

Another major benefit to this type of gesturing occurs when collaborating on large surfaces. Items on the screen may be out of reach to some participants, and therefore, interactions with these items can be accomplished using cooperative gestures, where one collaborator specifies the target item, and another collaborator specifies the action/gesture to achieve the desired intent with the target item. This method also avoids the awkward scenario of reaching across other collaborators (into their personal territory) to reach the desired item. *Figure 14* depicts two methods implemented by the researchers in their study that enable users to locate and give pictures to other members across the table. In *Figure 14 a)* user 1 touches the photo with 3

fingers and makes a throwing gesture, while user 2 taps an empty location on the table with 3 fingers to signal where the photo should land. In *Figure 14 b*) user 1 places the edge of her hand on the table, aimed toward the desired image, and the desired image then begins to blink. User 2 then taps on the blinking picture to select it, and user 1 can then move her hand toward herself, to drag the picture with her.

Figure 14. (a) The “throw-and-receive” gesture is one technique for passing photos long distances. (b) The “pull” action (where a partner selects the desired target for the puller) is another option for moving photos long distances. Source: Morris et al., 2006.

Another benefit to collaborative gestures relates to privacy, as there may be personal items or documents that a collaborator might not want disturbed. With cooperative gesturing, sensitive actions such as editing a document can require a cooperative gesture from both the document’s owner and anyone desiring to modify the document. Users then have complete

control over their personal documents, even when they are being shared on a shared display groupware system, allowing for ease of mind and security for the document owner.

While the incorporation of multiple natural gestures into user interface design has led to an array of possibilities over a wide range of devices, Norman (2010) argues that gestures lack the critical clues that are necessary for successful human-computer interaction. Since gestures are transient, they do not leave behind a record of their path. Therefore if a gesture is made and gets no response or the wrong response, there is little information available to help understand why this occurred. This is an important issue when thinking about designing interfaces, especially those that will be used in a collaborative setting, and will be discussed further in the analysis and design sections of this paper.

2.2.1 Deixis

In fields such as linguistics and pragmatics, deixis has been defined as the part of grammar that uses deictic terms based on the notion that spoken or written language has a sender of a message and a receiver of that message (Waszink, 1988). We are focused specifically on gestural deixis where an object/person is referred to physically, such as by holding one's finger up while saying, "this finger" (Levinson, 2008). We are therefore less concerned with the grammatical uses of deixis (in text) and this section will review only gestural deixis, and most specifically, pointing.

Pointing is a common gesture people use in their everyday interactions with others. Pointing is performed to indicate an object, location, or direction that may be visible at the moment, or may exist in the real world but cannot be seen at the moment, such as something behind a building (Kendon, 2004). It usually accompanies speech to decrease spatial ambiguity, as it helps listeners understand spatial information more clearly (Alibali, 2005). Pointing is

usually performed with the hands, however, it may also be done using the elbows, using eye or head movements, protruding the lips, or with the foot. Pointing is a movement toward some specific target, and this type of gesture involves the body part that is pointing being moved in a well-defined path (Kendon, 2004). Many times, there is a post-stroke hold during a pointing gesture, where the body part that is pointing is briefly held in the final position before returning it to the rest phase.

Clark (2003) categorizes pointing as a technique for indicating, or what he refers to as *directing-to*. During an interaction between two or more people, *directing-to* occurs when the speaker uses a signal to direct the others' attentions towards an object. This is done to establish a connection between the index and the object, and is usually a spatial connection, as the direction of the index directs one's attention to the desired object of interest. This technique relies on the observer recognizing the intrinsic relationship between the index and the object; or the referent and the site. In *Figure 15* Clark displays various methods of *directing-to*, which are not limited to pointing in the traditional sense. For example, one can use one's head to nod at something in order to direct attention to that object, or one can use one's finger to tap on an object that one is referring to.

TABLE 10.5
Methods of Directing-To

<i>Instrument</i>	<i>Index i</i>	<i>Example</i>
Finger	Pointing at <i>o</i>	" That is the book I want."
Arm	Sweeping <i>o</i>	" All this is yours."
Head	Nodding at <i>o</i>	"She was standing there ."
Finger	Tapping on <i>o</i>	" This is the book I want."
Foot	Tapping on <i>o</i>	[of carpet samples] "I like this best."
Torso	Turning to <i>o</i>	"Let us talk."
Face	Directing at <i>o</i>	[looking up from papers on desk] "Can I help you ?"
Eyes	Gazing at <i>o</i>	"I want you [person A] and you [person B] to come with me."

Figure 15. Various methods of "directing-to"; for example, by using one's head to nod at something. Source: Clark, 2003.

Kranstedt, Lücking, Pfeiffer, Rieser, and Wachsmuth (2006) define the combination of a pointing gesture and a referring verbal expression as a *complex demonstration*, where pointing adds content to a definitive description. This is done by specifying an object independently of the definite description (*object-pointing* to draw someone's attention to an object), or by narrowing down the description's restrictor or spatiotemporal region (*region-pointing* where the gesture refers to a region).

With regard to deixis in collaborative settings with visualizations, Hill and Hollan (1991) point out that "the social use of graphics relies heavily on pointing behaviors, those behaviors that direct visual attention by whatever means available toward salient aspects of graphics, often in coordination with speech and text" (p. 316). They claim that deixis is fundamental to visualizations, especially within social situations, such as when a presenter points to various presentation graphics using his hand, a pointer, a laser, a mouse and cursor, through narrative reference, or by some embedded visual technique. In order for a visualization to be effective, it must therefore include efficient methods of deixis. Since visualizations allow for graphic transformation, there are a diverse number of techniques that can be used to direct attention towards some aspect of the visualization. These deixis-enhancing visualizations would have visually appealing, yet subtle forms of deixis cues, which could be achieved by using, for example, lighting, textures, highlighting, point of view, size, transparency, cutouts, arrows, blurs, or eye-catching movements. Since visualizations allow for graphic transformation, diverse techniques can be used to direct attention towards some aspect of the visualization. This will be discussed further in the design chapter, as it is essential to collaborative visualization designs.

2.3 Natural user interfaces

With the introduction of multi-touch technologies, and other modern interactive techniques, natural user interfaces have become a prominent interface style. The *natural* part of a NUI refers to mimicking the real world (Wigdor & Wixon, 2011). Various examples of NUIs include software products designed for input modalities such as voice, gesture, eye gaze, and body-based interaction (Vetere et al., 2014). Wigdor and Wixon (2011) explain that the term ‘natural’ does not refer to the interface at all, but instead to the way users interact with the product and how they feel about the product. In other words, it refers to what the users do and how they feel while they are using the interface. A natural user interface should mirror users’ capabilities, while at the same time, taking full advantage of their capacities to meet their needs. With NUIs, users must feel the intuitiveness of the product right away, and not after multiple uses and a big learning curve. There are many parallel ideas and theories to NUIs, most of which stem from an era known as post-WIMP. We begin this section with the history of this era and then move on to reviewing theories and frameworks leading from this era with regards to NUIs.

In his 1997 paper, Andries van Dam discusses the four stages of user interfaces throughout history. The first period, in the 1950s and 1960s, consisted of computers used in batch mode with punched-card input and line-printer output, and did not have a user interface. The second period, in early 1960s through early 1980s, was the era of timesharing on mainframes and minicomputers, where users could interact with the computer by typing in commands with parameters. In the third stage, Xerox PARC started a new wave of user interfaces—WIMP GUIs (graphical user interfaces based on windows, icons, menus, and a pointing device, typically a mouse), which is still highly popular today. Van Dam coins the fourth stage of user interfaces as post-WIMP, where menus, forms, or toolbars aren’t used anymore, but instead, aspects, such as

gesture and speech recognition, are used. In this post-WIMP phase, the goal is to decrease the cognitive distance between what the user intends to do, and how the user carries out that intention focusing on the task at hand, without thinking about the technology. The ideal interface would be to have no interface at all—where the computer senses what the user wants to do and acts accordingly, but as of the time of his publication (1997), van Dam describes a post-WIMP interface as one that contains at least one interaction technique not characterized by traditional WIMP techniques such as menus or icons.

Jacob, Girouard, Hirshfield, Horn, Shaer, Solovey, and Zigelbaum (2008) propose that these emerging interaction styles in the post-WIMP stage can be understood as a new generation of HCI through the concept of Reality-Based Interaction (RBI). These new interaction styles utilize interface objects that are more realistic and seen in our daily lives, and allow users to interact more directly with them in ways similar to real world interaction. They list four real world themes of RBI:

- 1) Naïve Physics (NP): common sense knowledge about the physical world and the way things operate, such as gravity. User interfaces are increasingly simulating or directly using these characteristics of the physical world. For example, mobile phones use inertial sensing to orient content: when the user turns the phone from landscape to portrait view, the content displayed on the screen is turned as well to remain right side up.
- 2) Body Awareness and Skills (BAS): the perception and understanding of one's own body and how it moves; for example, coordinating one's legs to take steps to move forward. Interfaces now support a variety of input techniques based on these skills; for example, rolling a bowling ball using one's own arm while playing bowling on the Nintendo Wii.

- 3) Environment Awareness and Skills (EAS): understanding the clues embedded in the environment that facilitate our sense of orientation and spatial understanding, such as shadows providing depth cues, and also the skills used to manipulate objects, such as picking up and positioning a camera on a table. Interaction styles often draw upon these skills; for example, tapping with a finger or pen on a touch screen to select an item.
- 4) Social Awareness and Social Skills (SAS): awareness and communication towards others in the environment, such as when collaborating on a task together. Many interaction styles encourage social awareness and remote or co-located collaboration; for example, virtual conference rooms that allow meetings to occur among people in different places.

Jacob et al. (2008) believe that viewing interaction through the lens of RBI might provide insights for design and opportunities for future research. There are design trade-offs that need to be made in some cases however; when mimicking reality alone is not enough, and other characteristics are more important. These characteristics include:

- 1) Expressive Power: the functionality of a system.
- 2) Efficiency: how fast users can perform a task.
- 3) Versatility: if users can perform many tasks from different application domains, (however, a system like a tangible user interface that is more realistic may not afford as much versatility as a computer that is less realistic).
- 4) Ergonomics: if users can perform a task without physical injury or fatigue.
- 5) Accessibility: if users with a variety of abilities can perform a task; such as disabled individuals.
- 6) Practicality: if the system is practical to develop and produce.

Jetter, Reiterer, & Geyer (2014) take the concepts of reality-based interaction and natural user interfaces further by introducing the concept of Blended Interaction. This conceptual framework helps to explain whether users perceive user interfaces (UIs) as natural or not. They believed this framework was necessary since HCI still lacks an understanding of why some interfaces are successfully viewed as invisible or natural, while others fail. They claim that this uncertainty stems from a lack of any theory or model about the cognitive processes that occur when we perceive a UI as natural or not. Blended Interaction includes a blend (see *Figure 16*, center) between *concepts from the users' familiar reality*, including already well-established digital concepts such as two-finger zooming (see *Figure 16*, left side); and the *expressive power of digital computation* (see *Figure 16*, right side). These *concepts of reality* include basic experiences and building blocks of human thought (e.g., affordances, and image schemas such as object permanence), and the *power of digital computation* refers to the 'magic' that post-WIMP interfaces offer (e.g., high-resolution displays, wireless networks, and smart objects).

Figure 16. Blended Interaction Framework. Blended interaction helps to explain whether users perceive a UI as natural or not. It includes a combination of a user's known concepts of reality, and the computational power of the system. There are four domains of design of blended interaction: individual interaction, social interaction, workflow, and physical environment, as seen at the bottom of the diagram. Source: Jetter et al., 2014.

They list four domains of design for Blended Interaction (see bottom of *Figure 16*):

- 1) Individual interaction: this domain is at the individual level and refers to the individual interacting with an artefact to manipulate it, which demands a need for an adequate amount and variety of interaction techniques.
- 2) Social interaction and communication: this domain considers the social norms, protocols, and practices that we have developed when socializing and collaborating. When

designing collaborative systems, these practices need to be considered in order to support, and not interfere with, natural social interaction, territoriality, and communication. An example is providing a visual-tangible representation of the task at hand that increases group awareness and affords interaction by all collaborators.

- 3) Workflow: this domain considers the idea that “the structure and sequence of individual and social interactions are often guided by a higher-level workflow that defines the typical phases of collaboration and their intended outcomes” (p. 1150). Designs must consider supporting the users’ actual tasks and work processes and must not be concerned solely with individual and social interactions. This includes determining the best tradeoffs “between enforcing clearly defined workflows by providing functionality in sequences, templates, or page flows and, on the other hand, creating designs that are open for unanticipated use and encourage more inventive and creative results” (p. 1150).
- 4) Physical environment: this domain considers the physical arrangement and architecture of the environment, including the room and the objects inside, as well as the form factors of digital devices such as display sizes. In some cases the physical environment already exists and only minimal alterations can be made such as changing the lighting or sound level.

Widgor and Wixon (2011) focused more on the meaning of the term *natural* as well as how to design for NUIs. They state that achieving a natural user interface is not best accomplished by mimicking other experiences, or by relying on familiar metaphors. Instead, they suggest that NUIs must be designed, researched, and engineered. This is a process that does not involve using previous designs by others, but instead, designing a completely new experience based on detailed research. For example, direct multi-touch devices allow more degrees of freedom and

expressiveness than with a mouse alone, and the goal of creating a NUI for these systems is to appropriately leverage this potential. Instead of copying WIMP interfaces that are well-known to most users, a completely new interface must be designed with new input actions and affordances that are authentic to the technology itself. To achieve this NUI, the simpler interactions (those with few possibilities and a clear and familiar model of interaction) should first be designed and tested, and once perfected, then extrapolated on to create more complex interactions. This is what they mean when they suggest that NUIs need to be carefully designed, researched, and engineered. More of Wigdor and Wixon's ideas will be used in the design chapter later on.

In summary, NUIs are a thing of the post-WIMP era which involves a change in user interface styles. There are many factors to consider when designing NUIs, such as the trade-offs that need to be made and most importantly, how to create an interface that is easy to learn and feels natural to its users. This is best achieved through user research and careful design.

2.4 Touch interaction on a large surface

Large displays allow many people to see detail simultaneously and allow for pointing and discussion, even if there is no direct interaction with any software (Brown, Wilson, Gossage, Hack, & Biddle, 2013). Displaying data visualizations on a large *interactive* screen provides users with the ability to move objects around like they would naturally do in the real world. They can do this via gestures such as pushing and dragging to achieve intents such as rearranging, and categorizing objects. The view of these digital objects can also be manipulated in additional ways, with possible intents being to zoom, pan, highlight, filter, or select, for example.

Multi-touch sensing enables a user to interact with a system with more than one finger at a time, and can also accommodate multiple users simultaneously, which is useful for larger, shared-displays (Han, 2005). Large displays usually use some type of computer vision to

determine touch coordinates. Multi-touch technologies first appeared in 1982 when the University of Toronto's Input Research Group created the first human-input multi-touch system, using a frosted glass panel with a camera behind it (Buxton, 2007). Touch input was detected by the camera as a black spot placed against an otherwise white background, and was pressure-sensitive, as the size of the dot depended on the pressure of the finger placed on the screen.

More recently, and after much research and many working models had been developed, Han (2005) popularized the technique of detecting multi touch gestures using Frustrated Total Internal Reflection (FTIR). He describes the technique as follows:

When light encounters an interface to a medium with a lower index of refraction (e.g., glass to air), the light becomes refracted to an extent which depends on its angle of incidence, and beyond a certain critical angle, it undergoes total internal reflection (TIR). Fiber optics, light pipes, and other optical waveguides rely on this phenomenon to transport light efficiently with very little loss. However, another material at the interface can frustrate this total internal reflection, causing light to escape the waveguide there instead (p. 116).

In a touch display, beams of infrared light are shone into the edge of a clear acrylic display, and when a finger is placed on the screen, it deforms one side of the display. The beams of light are then projected down into a video camera that only detects infrared light. While Han's technique provided a great potential for collaboration on touch displays using fingers or objects made for screens, FTIR displays are not ecologically friendly, as they use cameras and projectors, therefore requiring a large installation behind or below the display (Brown et al., 2013). This is because the camera and projector have to reach the entire screen, so they must be positioned quite far behind or below the screen.

A newer, less bulky method is Microsoft's PixelSense technology (Microsoft, 2012) which allows for multi touch recognition of fingers, hands, and objects, without the use of cameras. It is a liquid crystal display with a large grid of sensors attached to it. The individual sensors/pixels in the display detect, process, and interpret touch information from the screen. These sensors can be placed in a layer behind the screen, therefore eliminating the bulk that accompanies FITR methods. The origin of this technology is ThinSight, developed at Microsoft Research in Cambridge, which enables multi-touch sensing using novel hardware embedded behind an LCD (Hodges, Izadi, Butler, Rrustemi, & Buxton, 2007).

PixelSense works as follows: infra-red back light provides light that hits the point of contact on the display (i.e., a finger). This light is then reflected back from the contact and seen by the sensors, which then converts the light signal into an electrical value. Values reported from all sensors are used to create a picture of what is on the screen which is analyzed using image processing techniques. This final output (the sensor image and contact types—for example, fingers, blobs, or hands) is then sent to the PC.

Unfortunately, the PixelSense sensors, being located very close to the screen surface, pick up all light coming from external sources as well as ambient light. It then becomes easily flooded and blinded by any external IR light (MultiTouch Ltd, 2013).

Microsoft recently released a new product, Microsoft Surface Hub, for team collaboration that can recognize up to 100 touch points with precision, allowing multiple people to interact with the screen simultaneously (Microsoft, 2015). This system uses a Projective Capacitance optically bonded sensor, which dramatically improves display system readability when ambient light levels are high (Review Display Systems Ltd, 2012).

Another method that reduces the bulk of FITR systems is to embed infra-red lights and detectors around the edges and the top of the display within a frame (Brown et al., 2013). For example, PQ Labs (PQ Labs, 2015) provides a 4K multi-touch display series, or a variety of multi-touch wall displays. Frame sizes are scalable, and can range from desktop screens to wall-sized displays, but the downside is a limitation on the number of independent touch events that can be interpreted (Brown et al., 2013).

2.4.1 Beyond touch interaction on the screen

While the focus has been placed mostly on creating software that is not bulky and that can determine multiple points and types of contacts on a screen, research is expanding in scope. Lozano, Jindrich, and Kahol (2011) focused on the ergonomics of multitouch, and the effects on the user's musculoskeletal system. They stated that poor design strategies can cause biomechanical and physical stress to the musculoskeletal system, which can cause musculoskeletal disorders. Their goal was to provide designers with objective data on the impact of multi touch interactions on the musculoskeletal system, to make multi touch systems not only easy to use, but also easy on the body in the longer-term.

Murugappan, Vinayak, Elmqvist, and Ramani (2012) were interested in going beyond simply determining touch points, to determining the identity, posture, and handedness of the user as well, which they called *extended multitouch interaction*. The aim was to increase the capabilities of multi-touch interaction and collaboration.

Jota, Nacenta, Jorge, Carpendale, and Greenberg (2010) were interested in going beyond the surface of the screen to determine effective ways of detecting pointing gestures towards the screen. Pointing is a common mid-air gesture used to bring other collaborators' attentions towards a point of interest, and is different from other gestures because it is not

performed directly on the screen. In their study, they compared 4 types of *ray pointing techniques*, which are techniques using ray casting (the intersection of a ray with a surface) to determine the point on a surface that is being pointing at from a distance. They considered ray pointing as natural interaction techniques because they do not require a physical surface to operate on (such as a mouse). Their technique draws on the user's concepts of reality and their body awareness skills, and is easily understood since it draws upon everyday pointing gestures. Ray pointing also has the additional advantage of allowing multiple users to interact in the same space without their bodies physically getting in the way, as might occur when interacting directly on the screen with touch gestures.

Aslan, Uhl, Meschtscherjakov, and Tscheligi (2014) focused more specifically on mid-air gestures, stating that since we use our hands and fingers to manipulate real-world objects, mid-air gestures can improve interaction with ubiquitous systems because they allow users to interact with the digital world as they would naturally interact in the real world. They studied mid-air gestures for authentication in clean rooms that have constraints on interaction modalities and require clean and ideally touchless ways to interact with workstations. They designed two authentication gestures, which were perceived as easy to perform and suitable for all relevant situations, and also provided hand biometry information for touchless authentication.

Reetz & Gutwin (2014) also focused on mid-air gestures— in particular they wondered what factors of mid-air gestures affect group awareness in co-located environments. They cited Don Norman (1993), who discussed the usefulness of big controls and actions for shared work, stating that when engaging in shared tasks, the critical component is to keep everyone informed about the complete state of progress. This is easily done using large gestures because they enable collaborators to maintain awareness of their colleagues' actions through

observation. Norman used observational studies to identify gesture size as the primary factor of gesture visibility. Reetz & Gutwin (2014) sought to identify other factors that also influenced gesture visibility, which is important for creating team awareness. Their research showed that gesture size was an important factor in creating awareness (bigger gestures were more easily seen), but that other factors were also important, such as the direction of a gesture—gestures that led away from the body seemed to be more observable, and the ability of users to easily discriminate between purposefully executed gestures and unintentional gestures/movements.

While there is a lot of research on midair gestures (especially since the technology to support this kind of interaction has become much more accessible with devices like the Microsoft Kinect) the research focus of our study is on touch gestures only.

2.5 Background summary and connections to our study

This chapter reviewed research in five different fields: collaboration, human gestures and

deixis, natural user interfaces, and touch interaction on large surfaces. These are all fields of study relevant to this thesis paper that looks at the relationship between physical gestures executed by collaborators on a large touch screen and their intents. The goal is to find natural gesture-intent pairings that occur in collaborative settings and to recommend design suggestions based on these findings that can lead to a natural user interface for collaborators.

Collaboration is generally defined as multiple people working together towards a common goal. Johansen (1991) described four types of collaborative situations involving time and place. In our study, we will focus on the type of collaboration “same time, same place”, as our participants will be collaborating in pairs in the same room (co-located collaboration), and

therefore, our analysis and design suggestions will take into account only collaboration situations where group members are collaborating in the same space, at the same time.

Grudin and Poltrock (2012) showed a model of technology development and use within the field of computer supported cooperative work, detailing how technology becomes adopted in a collaborative setting. The important aspect of this model is that a technology investigation must be conducted to find the needed technology that will assist with collaboration (if it already exists), or to design a new piece of technology if it doesn't already exist. Butler et al., (2011) discuss the idea of the challenges for group members in a collaboration in creating a common understanding, and more generally, for designers to devise methods of interaction that allow for effective collaboration. Furthermore, when using technology to collaborate, the technology may enhance, but also complicate the collaboration. Vatrappu and Suthers (2009) use the term technological intersubjectivity to denote the idea that when technology is used to support collaboration, the interface becomes another entity in the social interaction equation. Suthers (2006) discusses how the study of technological affordances (in collaborative settings) should be constantly done with reference to the activity that it is trying to support — for example, a collaborative goal. There should also be a special focus on the design and study of the social interactions that are informed by the affordances and limitations of the technology. The technological intersubjectivity framework was used in our study to assess the various categories of natural gestures and intents that people use in a collaborative task. It was also used in the design phase where we understand the large multi-touch display as an active entity in the social interaction equation and design techniques that can enhance communication.

We moved on to discuss human gestures and deixis, where an object/person is referred to physically, most often done by pointing. Kendon (2004) defines gestures as

purposeful, deliberate actions that have expressive power. He further breaks gestures down into gesture units: *preparation*, which begins the gesture; *stroke*, which is closest to the apex of the gesture, and *recovery*, which is returning back to the initial state of rest.

Wu et al., (2006) coined three terms related to designing gestures for touch surfaces: *gesture registration*, *relaxation*, and *reuse*. The main idea here is that users should be able to perform gestures with minimal constraints after the gesture is registered in the gesture registration phase (the beginning phase of a gesture), and be able to reuse the basic components of a gesture to complete different intentions. In this way, the user has to remember less gestures.

Lao et al. (2009) constructed a design model for single user touch interaction based on three simple hand movements: pressing, tapping, and dragging. Similarly to our study, they looked at which gestures corresponded with which intents, but with a focus on a single-user photo manager application. They found some interesting gestures, such as a fist tap to refresh the page. Wobbrock et al. (2011) also determined various gesture-intent pairings, but for a table top system, and drew sketches to depict all of the gestures they saw. This was a study we drew on for creating our own study.

Morris, et al. (2006) introduced the term *cooperative gesturing* where people collaborating gesture in sync to complete various intents. This type of specific coordination that is needed between the collaborators can increase group cohesion and teamwork when working together.

We discussed Norman's (2010) argument that gestures are transient and therefore if a gesture is made and gets no response or the wrong response, there is little information available to help understand why this occurred. This is an important issue that we looked at in our design chapter when thinking about designing interfaces for collaboration.

The third area of research we discussed was natural user interfaces (NUIs), which Wigdor and Wilson (2011) describe as being interfaces that mirror users' capabilities, while at the same time, taking full advantage of their capacities to meet their needs. They focus on the idea that *natural* concerns how the user feels when interacting with the interface, rather than referring to the interface itself. This was important for our study, as we wanted to look at the natural types of gestures people would use on the screen, in order to create informed designs that create a natural interface.

Jacob et al. (2008) described a framework called reality based interaction where interface objects that are more realistic and seen in our daily lives should be used, while allowing users to interact more directly with them in ways similar to real world interaction. Their framework takes into account naïve physics, body awareness and skills, environment awareness and skills, and social awareness and skills, and they believe that viewing interaction through this lens might provide insights for designing natural user interfaces. Similarly, Jetter et al. (2014) discuss their framework called Blended Interaction, where an interface is perceived as natural based on two factors: including concepts from the users' familiar reality, and the expressive power of the technology being used.

We ended this section coming back to Wigdor and Wixon (2011) who suggest that NUIs must be designed, researched, and engineered in process that does not involve using previous designs by others, but instead, designing a completely new experience based on detailed research. This is the approach we wanted to take in our study, as we looked at first researching the gestures that people find natural, and then coming up with design suggestions based on these findings.

Lastly, we discussed touch interaction on large surfaces with a focus on multi-touch, as our study looks at multiple people interacting with a screen simultaneously. We discussed how multi-touch technologies first appeared in 1982 when the University of Toronto's Input Research Group created the first human-input multi-touch system. Then later in 2005, Han popularized the FTIR method with bulky cameras behind the screen to detect touch points. Microsoft's PixelSense technology (Microsoft, 2012) later allowed for multi touch recognition using a liquid crystal display with a large grid of sensors attached to it.

We also discussed interaction beyond the screen, or mid-air gestures, with a focus on pointing techniques, as pointing is a common aspect when interacting with others around a visualization. While there are many techniques for detecting interaction beyond the screen, such as ray pointing, designing for this type of interaction was beyond the scope of our research.

3 Chapter: Study Design

In this study, we are exploring how gestures relate to users' intentions while analyzing a parallel coordinates plot in pairs. We chose to conduct a qualitative study, as we wanted to explore, understand, and analyze unstructured data to provide us with descriptive results of collaborative gestures based on observations and feedback from participants. This study was approved by the Carleton University Research Ethics Board: Project #102137.

3.1 Developing tasks for our study

To determine what types of tasks participants would be asked to perform, we chose to use a general model of common interactions around information visualizations by Yi, Kang, Stasko, and Jacko (2007) as a guide for task development. They define interaction techniques as “features that provide users with the ability to directly or indirectly manipulate and interpret representations” within a visualization (p. 1225). Their model draws from their belief that existing interaction taxonomies either “focus strongly on interaction techniques and are relatively system-centric, [or...] focus on user goals without a main focus on interaction” (p. 1225). They believed it would be beneficial to bridge these two efforts, and so their model is a taxonomy of intent-based interactions for information visualization systems that connects user objectives (intentions) with information visualization interaction techniques. By surveying 59 papers and 51 systems, they collected 311 individual interaction techniques that are commonly used in information visualization systems, and grouped similar interaction techniques together, iteratively refining the groups as necessary. The major category used to group the techniques was the user's intent in performing the technique — what the user wants to achieve. At the end of the grouping process, they were left with seven categories of interaction in information visualization, based on user intent. We will refer to them simply as intent categories. Each intent category

includes an identifying name and a description that captures the user’s intent in performing the interaction. Yi et al.’s intent categories are found in **Table 1. Yi et al.’s (2007) intent categories.**

Table 1. Yi et al.’s (2007) intent categories.

Intent name	Intent description	Example technique to achieve the intent
Select	Mark something as interesting	Highlighting
Explore	Show me something else	Panning
Reconfigure	Show me a different arrangement	Rearranging the order
Encode	Show me a different representation.	Changing a pie chart to a histogram
Abstract/Elaborate	Show me more or less detail	Zooming in
Filter	Show me more or less detail	Range sliders to select a range of data to be displayed
Connect	Show me related items	Highlight relationships between data items

While this is not a collaborative model, the purpose of our study is to explore how these typical intents for visualizations would be executed in a collaborative setting. We believe that this taxonomy will help us to create realistic tasks, with a good coverage of a range of tasks relevant to information visualization. Questions/tasks for our study incorporate at least one of Yi’s intent categories each. **Table 2** depicts how many times each of Yi’s intents were used within our set of tasks.

Table 2. Frequency of use of Yi et al.'s intent categories in our list of study tasks

Intent Category	# of times used in tasks
Select	8
Explore	1
Reconfigure	1
Encode	0
Abstract/Elaborate	2
Filter	1
Connect	6

“Select” and “Connect” were used the most for designing tasks because these are highly common intents when analyzing graphical data; however a task is not always limited to these 2 intents, as participants may also perform other intents while completing these tasks, such as exploring. We found it difficult to generate analysis tasks that would account for all seven of Yi’s intents, since we were using a static diagram and not an interactive diagram. We decided on a static image rather than an interactive image because we wanted to see the natural gestures participants would perform, without the complications introduced by trying to figure out how to use an interface. All of our initial tasks were realistic and analysis based, involving participants working together to find answers to specific questions about the data. To incorporate more of Yi’s intents into the study, we created a second set of tasks which hypothetically asked participants how they would achieve a specific intent, such as reconfiguring the graph. In order to maintain the same rhythm as the analysis task questions, we phrased these latter questions as tasks as well, instead of directly asking what gestures the participants would use to accomplish the desired intent. For example, instead of asking “How would you pan on this visualization?”

we asked instead, “Imagine this was only a part of the graph being seen. Explore the visualization to see other parts of the graph”. We wanted to elicit responses that were as natural as possible, by allowing participants to pretend to interact with the visualization as if it were interactive, rather than simply verbally explaining the type of gesture they would want to perform to complete an intent. This aligns with our goal of determining the natural kinds of gestures that participants would find suitable for various intents on a parallel coordinates graph, as stated in our research question.

We labelled the first set of analysis-specific tasks as *realistic tasks*, and the second set of tasks as *hypothetical tasks*, since they relied on participants supposing the graph was interactive. By having these two types of questions in our study (realistic and hypothetical tasks), we were able to incorporate a variety (6) of Yi et al.’s intents (encode was not used in our study as it was a more complex intent that involved changing the way the data was presented, and we felt that this was not an important intent to focus on in our study). Having both types of tasks would allow us to not only observe which gestures were paired with which intents during task completion, but also to determine which gestures participants would feel natural using for various other intents that would not have come up during the realistic task analysis phase of the study (such as zooming in on the display, for example). This would allow for greater feedback for design purposes later on in the study. Our study tasks are as follows with the corresponding intent(s) shown in parentheses:

Realistic tasks:

- 1) How many arrests were made for larceny among 21 year olds? **(Select)**
- 2) Compare the arrest rates for weapons between 15 year olds and 23 year olds. What trend do you see with age? **(Select) (Connect)**

- 3) Compare the arrest rates for 17 year olds and 21 year olds on both driving under the influence and drunkenness. What trends do you see here? **(Select) (Connect)**
- 4) Between burglary and vandalism, where do we see the biggest decrease in arrest rates with increasing age? **(Select) (Connect)**
- 5) What 3 crimes account for the most arrests among the younger population? **(Select) (Connect)**
- 6) What age/age group or age range would you target for a burglary intervention program that you are developing? **(Select) (Connect)**
- 7) Kale is 16 years old. According to this graph, what crime do you think he would be most likely to commit? **(Select) (Connect)**

Hypothetical tasks:

- 8) Highlight the offense that each of you is most interested in examining. **(Select)**
- 9) Give me the definition of ‘theft’ used by the creators of this data. **(Abstract/elaborate)**
- 10) Imagine that you could find out more detail about a portion of the visualization by enlarging that section of it. Find out more details about the upper left hand corner of this diagram. **(Abstract/Elaborate—zooming)**
- 11) The graph contains information that is not relevant to you. Delete the two offenses at the very ends of the graph (theft and vandalism). **(Reconfigure or Filter)**
- 12) Categorize the offenses into two groups: one for more serious offenses and one for less serious offenses. Place the more serious offenses at the top of the graph and the less serious offenses at the bottom of the graph. **(Reconfigure)**
- 13) Imagine this was only a part of the graph being seen. Explore the visualization to see other parts of the graph that are off-screen. **(Explore—panning)**

14) How would you change this visualization to be more effective? (**Could be any of the techniques as there are multiple intents here**)

3.2 Participants

Participants were recruited through posters and word of mouth at Carleton University. There were 7 females and 9 males, a total of 8 pairs of participants. Twelve participants were undergraduate students, and 4 were at the graduate level. Thirteen were right handed and 3 were left-handed. All participants had an extensive amount of prior use with touch screens, but none had experience with parallel coordinates plots. Seven of the 8 pairs came together (as friends), while one pair of participants did not know each other prior to the study. The latter situation had no effect on the study; the sole difference between this pair of participants and others was some initial awkwardness and hesitation between these participants.

3.3 Apparatus

In this study, the parallel coordinates plot was displayed on a 46 inch vertical screen. We chose this large screen size to allow for a sufficient viewing size of the graph for collaboration based on our literature review detailing the benefits of collaboration around large surfaces. The screen was not touch enabled or interactive, and therefore gestures on the screen had no effect on the graph. A laptop was connected to the screen, and the study was operated from a PowerPoint presentation on the laptop.

3.4 Materials

The materials used in the study included a short demographics questionnaire, the parallel coordinates graph, and the tasks asked of the participants (refer back to section 3.1). The questionnaire consisted of two short sections: demographics information and prior experience information. Demographics included age, year of study, sex, program of study, and right/left

handedness. The prior experience section asked for a rating of past use of both touch screens and parallel coordinates plots using the following rating scale: no experience, some experience, moderate experience, and a lot of experience. This information was collected to possibly explain any unusual results that were found, but was not used to analyze the data.

3.4.1 Parallel coordinates graphs

The study began due to our interest in data graphs, and in particular, parallel coordinates plots. These plots solve the problem of visually displaying multiple-dimensional numerical data, and can allow meaningful multivariate patterns and comparisons to be seen when used interactively for analysis (Few, 2006). While a parallel coordinates plot may look very similar to a line graph, the two are very different in the meanings displayed: in a line graph, lines encode time-series data, where the slopes of the lines depict changes through time from one value to the next. By distinction, in a parallel coordinates plot, the lines do not indicate change, but they connect a series of variables, which can be measured on different dimensions. Examples of variables include Type of product or Country. The graph displays these variables on different vertical X-axes, and they can each have different dimensions (i.e., units), such as dollars, acres, percentages, years, etc. In Figure 17, a parallel coordinates plot displays 7 variables of U.S. counties (the items). From left to right the variables are Home value in dollars, Farm acres, Average income in dollars, Population, +65 population, College graduation percentage, and Life expectancy. In this graph, the scales for these variables (i.e., the dimensions) have been normalized as percentages. Highlighting one line in the graph in Figure 17 shows the data for Alameda County (an item) on all of the 7 variables. While the slopes of the line are not meaningful, one can see the data values for each variable for the item Alameda County, and it is easy to compare Alameda County to other counties on any of these 7 variables. For example, it is

easy to see that Alameda County has a lower percentage of older adults (65+) than most of the other counties.

Figure 17. A parallel coordinates plot that measures several aspects of U.S. counties such as Home Value.

Here, Alameda County in the state of California has been highlighted in green. Source: (Few, 2006).

3.4.2 The parallel coordinates graph in our study

The first step in the design of our study was to create a parallel coordinates graph that was not too cluttered for this study. We found an interesting dataset online of arrest rates by age in America on a number of different offenses. This data was publically available online from the U.S. Federal Bureau of Investigation and was retrieved from their website (Uniform Crime Reports, 2011). We decided to keep the number of offenses and age categories relatively small (4-6) to reduce clutter on our graph, as we wanted to display a static image. Offenses were selected based on the ranges of their arrest rates: we did not want the range of the graph to be too large, so offenses that had arrest rates between 0 and 60,000 were selected. We chose to focus on younger ages (15, 17, 19, 21, and 23) because participants might find this information more interesting. Once all of the data were chosen, a parallel coordinates plot was developed by another member of our lab, Jeff Wilson, using Python. Together we chose and altered the color scheme and curves of the data points to design a graph that was as easy to comprehend as possible. We chose a different color to represent each offense category: red, orange, cyan,

yellow, green, blue, and purple. We made the colors as saturated as possible, as Ware (2013) suggests that using more saturated colors when color coding small areas such as thin lines, makes the colors easier to distinguish. Our graph can be seen in Figure 18. Our parallel coordinates graph depicting arrest rates by age and offense categories in the U.S. Offense categories are displayed on the left axis, arrest rates are displayed on the middle axis, and age is displayed on the right axis. The color scheme is as follows: aggravated assault in blue, burglary in red, driving under the influence in yellow, drunkenness in green, larceny (theft) in purple, vandalism in orange, and weapons in cyan.

Figure 18. Our parallel coordinates graph depicting arrest rates by age and offense categories in the U.S. Offense categories are displayed on the left axis, arrest rates are displayed on the middle axis, and age is displayed on the right axis. The color scheme is as follows: aggravated assault in blue, burglary in red, driving under the influence in yellow, drunkenness in green, larceny (theft) in purple, vandalism in orange, and weapons in cyan.

3.5 Procedure

The study was conducted in the Multimodal lab at Carleton University in the Human Computer Interaction building at Carleton University. Participants first read and signed a consent form (detailing the procedure, requirements, and amount of compensation) and then filled out the demographics questionnaire. Participants were then taken to the large screen and told how to operate the PowerPoint slides using the connected laptop. They were then instructed to move through the introduction slides and ask questions for clarification. These slides included a brief overview of the study and an explanation of how to read a parallel coordinates plot with an example provided. The parallel coordinates explanation slide is shown in *Figure 19*. Prior to beginning the study, participants were presented with this slide that explained how a parallel coordinates graph worked.

Figure 19. Prior to beginning the study, participants were presented with this slide that explained how a parallel coordinates graph worked. They were provided with additional verbal explanations by the researcher and were allowed to ask questions for clarification.

Further explanations were provided by the researcher when requested. When both participants were comfortable they moved on to the first task of the study, as the researcher stood behind them with the video recorder. The participants worked together to answer the 14 tasks required of them, which involved talking to each other, and interacting together on the screen. When necessary, participants communicated concerns or questions to the researcher, as there was always an open line of communication and discussion between all parties. For the hypothetical task questions, the researcher held more open discussions with participants to gain a better understanding about how and why they would perform certain gestures and how they would desire the system to respond. For example, a few participants misunderstood what the following task was asking of them: Highlight the offense that each of you is most interested in examining. They thought that this was another analysis task and would try to find out what the arrest rate was for an offense of interest. In this case, the researcher would explain that it was asking them what they would do to highlight a line for example, and how they would want the system to respond. After the study, participants were thanked for their participation and were compensated \$10 in cash each. *Figure 20* shows the first task presented in the Power Point presentation, where the question/task is placed above the graph. Each slide shows the exact same graph, just paired with a new question/task.

Figure 20. The first task/question asked of participants in our study, displayed from a Power Point presentation.

3.6 Analysis Plan

We used Grounded Theory as the basis of our analysis, which is a form of inductive qualitative research with various coding procedures, such as open coding, and axial coding (Lewins & Silver, 2007). Open coding is very fine-tuned coding of small segments of data, generating a large number of codes, and opening up the data in all the possible ways it can be interpreted. Axial coding involves looking through all of the codes created from the open coding process and reconsidering them in terms of their similarities and differences to each other. For example, similar codes can be grouped together or merged into higher level categories that are less specific. This process also involves exploring the relationships between all of the fragmented codes developed in the open coding process. In our study, we moved from the very fine-detailed single codes created in the open coding session, to a more abstract coding created in the axial

coding session, by collecting similar codes together into groups which had more abstract names.

This will be elaborated on further in the analysis chapter.

4 Chapter: Analysis

4.1 Fine grained analysis

Each video recording was analyzed using Atlas.ti (“Atlas.ti Qualitative Data Analysis,” n.d.) by coding the various gestures and intents made by all participants (codes were created per instance, rather than on a per person basis). This process involved watching the videos and using Atlas.ti’s start and stop buttons to mark each segment of the video that included a gesture (which occurred constantly within each video). The length of each video can be found in *Table 3*. Each coded clip ranged anywhere from 2-8 seconds and therefore there were many codes within each video.

Table 3. Length in time of each video analyzed in this study.

Video #	Length of video (in minutes)
1	44
2	19
3	30
4	20
5	19
6	18
7	18
8	30

4.2 Open coding process

When a segment of a video was selected, two types of codes were given to it: a description of the physical mechanics of the gesture, and the intent or reason for performing the gesture. Some segments included more than one intent paired with the gesture, mostly when it was a collaborative 2-person gesture, where both participants were performing gestures on the screen simultaneously. We did not use Yi et al.’s intentions for coding purposes because theirs is a very general model of intents that would not have benefited us in our analysis. The intents that

we observed were more specific, fine-grained intents, and therefore we used very specific coding names and categories (for both intents and gestures). Fine-graining the intents was also necessary for our later purpose of using the data to create design patterns useful for designers developing collaborative applications with graphs on large touch screens. For example, Yi et al.'s intent "Select" is quite broad and may incorporate a variety of sub-intents that the user has, such as selecting something to gain more information about it, or selecting something to highlight it. If Yi et al.'s criteria were used in this example, from a design perspective, "to select" would not be very meaningful, as there would be many different gestures used to complete the various different types of intents within the general intent of selecting something. From a design perspective, intents need to be more specific in order to find the relevant gestures that would accompany them. Therefore we discovered our own intents through the process of grounded theory that were closer to the actual elements of our parallel coordinates graph, and that paired well with the gestures performed on these elements (such as "Tracing a line"—a gesture on a line element, "To find a certain point on the line"—the corresponding intent).

The open coding process was very detailed and time consuming (generating over 300 types of codes (both gestures and intents), most of which were used multiple times. *Figure 21* shows how codes are created in Atlas, and in *Figure 22*, a screen capture of a section of one video is shown, with the hundreds of codes displayed (not all can be seen for this section as some are to the left of the screen, as well as behind other codes due to limited display space on the screen).

Figure 22. The final codes on a section of one video. This is similar to *Figure 21* except the codes have been dragged over the video section of the page, allowing one to see a greater number of codes that were off to the right hand side. In this image, we can see about 120 codes, or about double the amount that can be seen in *Figure 21*.

An example of a gesture we saw multiple times was “Placing an index finger on a point of interest on the graph”, and we observed that the intent was “To placeholder” this spot in order to return to it later for comparison purposes. Other examples of gestures with their observed intents included “Hover pointing” To show the area one is referring to, “Double tapping” To select an item on the screen, “Moving index fingers towards each other” To zoom out, and “Flicking an item with index finger” To delete it. There were multiple gestures that were performed by both participants simultaneously which involved a combination of placing various fingers on the graph and tracing lines on the graph with the observed intents being “To placeholder” and “To find certain points on the lines.” In these simultaneous gestures, participants were coded as “P1” (participant 1) and “P2,” (participant 2); for example, “P1 places thumb on screen, P2 traces with index finger.”

After all 8 videos were coded, we went through a process of cleaning up the codes, which brought the total number of codes down to 234 (again, most of which were used multiple times). Codes were distinguished as gestures (the physical movements) versus intents through labelling: gesture code labels began with a “G”, while intent code labels began with an “I.” Further cleaning involved renaming codes to be shorter, and changing the wording of codes to ensure that similar codes sounded alike and would be easy to group together later (as Atlas.ti lists codes alphabetically in the “Code Manager”). *Figure 23* shows a screen capture of some of the alphabetized codes displayed in the code manager in Atlas.ti.

Figure 23. List of codes. This is referred to as the “Code Manager” in Atlas, and displays an alphabetical list of every code used within all of the coded videos. This is a snapshot of about 32 of the 234 codes.

To ensure inter-coder reliability, after the codes were created by the initial researcher in the open coding process, they were shown and discussed with the researcher’s two supervisors. Together, we went through the majority of the codes, to check that we were all in agreement on what the described intents and gestures were. As the gestures were obviously seen in the videos, and the intents were easy to infer from the videos, we all agreed on the codes that were created. There was one code that was up for discussion however: *placeholder*. The issue that we found was that *placeholder* could be argued for being both an intent and as a gesture. It can be seen as a gesture in the sense that the intent is to compare points of interests that are *being Placeholder*, and it is also an intent because it is the reason for placing a finger on a point of interest. After

discussions, we decided that Placeholder is in fact an intent, as it is the reason for the gesture (what the person intends to do), while the gesture component would be placing a finger on the screen (the physical mechanics of the gesture). However, we noted that there are many different gestures that could be done to placeholder, such as tapping on the screen, or holding a finger on the spot, and therefore, we decided *To placeholder* was an intent.

4.3 Axial coding process

The next step involved grouping codes together, as there were over 200 codes. This is a common step in qualitative analysis, and is a major part of bottom up analysis, which is an analysis beginning at a micro level and moving towards a more abstract level. We wanted to look at the gesture-intent relationships, but initially, we realized that using each single code resulted in too much data that was not very meaningful. Since the codes were so specific, we decided to group them together based on similarity of gesture or intent. This process involved using axial coding from Grounded Theory, as mentioned in the previous chapter. We moved from the very fine-detailed single codes created in the open coding session, to more abstract codes created in the axial coding session, by including similar codes together into groups which had more abstract names. For example, all gestures involving flicking an item on the graph, whether with one finger or two, were grouped together. In this example, the finer detail coding included specific gestures for which finger was used or how many fingers were used, while the grouped coding bunched these gestures together, as they were all essentially the same action, with slight personal variations. An example of intents that were grouped together included all intents “To show something”, such as showing a point of interest, or showing the correct spot on the graph. In the finer detail coding for showing something, the thing that was being shown was specified, while

in the more abstract coding, all intents of showing something were grouped together, regardless of what was being shown.

These categories were made in Atlas.ti by creating group codes (what Atlas.ti calls family codes), as seen in *Figure 24*. *Figure 25* shows all of the code groups (families) on the left, with the selected group's gestures being shown on the right (for the 2-person gesture group).

Figure 24. Creating and sorting the Code groups (what Atlas refers to as Family codes). On the top is the list of group codes. As seen in this figure, selecting the code group “Hand Motions” shows a list of codes that are in the group (on the bottom right), and the rest of codes that are not in the group (on the bottoms left).

Name	Gro...	De...	Author	Created	Modified	Families
G- Each placing index finger on graph	8	0	Super	2014-12-...	2015-03-...	G- Collaborative gestures, Gesture
G- P1 & P2 place index fingers on screen	2	0	Super	2015-01-...	2015-03-...	G- Collaborative gestures, Gesture
G- P1 & P2 places thumb, P2 traces w index	1	0	Super	2015-02-...	2015-04-...	G- Collaborative gestures, Gesture
G- P1 hover points; P2 hover traces a line	3	0	Super	2015-01-...	2015-03-...	G- Collaborative gestures, Gesture
G- P1 places both indexes; P2 places thumb	3	0	Super	2014-12-...	2015-04-...	G- Collaborative gestures, Gesture
G- P1 places both indexes; P2 traces w index	2	0	Super	2015-02-...	2015-04-...	G- Collaborative gestures, Gesture
G- P1 places index&middle; P2 places index & traces w othr index	2	0	Super	2015-01-...	2015-04-...	G- Collaborative gestures, Gesture
G- P1 places index&middle; P2 traces w index	1	0	Super	2015-01-...	2015-03-...	G- Collaborative gestures, Gesture
G- P1 places index, P2 traces w index	7	0	Super	2015-02-...	2015-04-...	G- Collaborative gestures, Gesture
G- P1 places index; p2 places index&middle	1	0	Super	2015-02-...	2015-04-...	G- Collaborative gestures, Gesture
G- P1 places middle; p2 traces w index	2	0	Super	2015-01-...	2015-03-...	G- Collaborative gestures, Gesture
G- P1 places thumb; P2 traces with index	3	0	Super	2015-01-...	2015-03-...	G- Collaborative gestures, Gesture
G- P1 takes P2 position so P2 can compare another line	2	0	Super	2014-12-...	2015-03-...	G- Collaborative gestures, Gesture
G- P1& P2 place indexes; p1&p2 trace w indexes	3	0	Super	2015-01-...	2015-04-...	G- Collaborative gestures, Gesture
G- P1&P2 place indexes;p2 traces with othr index	5	0	Super	2015-01-...	2015-04-...	G- Collaborative gestures, Gesture

Figure 25. Code Groups in the “Code Manager”. On the left side is a list of all code groups (families) and on the right are all of the codes within the selected group, “Collaborative Gestures” (later renamed to 2-person gestures).

Many codes were left out of the final analysis because they were not common (i.e., had 2 or less occurrences, like “backing hand away from the screen”), or were specific to an individual’s personal movements, such as a “parabola motion with the hand”. Other codes were eliminated because they were gestures performed away from the screen or gestures that naturally accompanied speech and weren’t relevant to this study; an example being cracking one’s knuckles, or speaking with one’s palms facing upward. At the end of the grouping process, there were 10 gesture categories and 11 intent categories.

4.4 Gestures

The 10 gesture categories we created from our coding process are shown in *Table 4*:

Table 4. Gesture Categories (including number of occurrence in brackets) with their respective gestures.

Gesture Categories with counts	Specific gestures within the category
<u>2-person gestures on the screen</u> (56) -involved a combination of both participants placing their	Both tracing lines Each placing index finger on graph P1 and P2 place indexes; p2 traces with index P1 hover points; P2 traces a line P1 places both indexes; p2 places thumb

<p>fingers on the screen and/or tracing lines simultaneously (see <i>Figure 26</i>)</p>	<p>P1 places both indexes; p2 traces with index P1 places index & middle; p2 places index & traces with other index P1 places index & middle; p2 traces with index P1 places index; p2 traces with index P1 places index; p2 places index & middle P1 places middle; P2 traces with index P1 places thumb; P2 traces with index P1 takes P2's position with index finger P1 & P2 place indexes; P1 & P2 trace with indexes P1 & P2 place indexes; P2 traces with other index</p>
<div data-bbox="203 764 1427 1068" data-label="Image"> </div> <p>Figure 26: Example of the 2-person gesture "Both placing fingers on screen"</p>	
<p><u>Circling</u> (10) -moving finger/hand in circular motion</p>	<p>Circling a point of interest Circling and tapping a point of interest Hover circling with hand Hover circling with index finger</p>
<p><u>Dragging items</u> (15) -involved pressing on an item (i.e. an offense category on the graph) and dragging it up or down with one's finger (see <i>Figure 27</i>)</p>	<p>Dragging items with index finger Pressing and dragging an item</p>

Figure 27: Example of the gesture "Dragging an item". The participant touches the bottom offense (Weapons) and drags it up the axis to the top.

Hand motions (86)
-involved a variety of movements including moving hands vertically or horizontally shaking hands, or wiggling fingers (see *Figure 28*)

Downward motion w whole hand
Flicking hand upwards
Moving both hands around
Moving finger down the screen
Moving finger up the screen
Moving fists up and down
Moving hand horizontally
Moving hand up and down
Moving hand up the screen
Moving hands towards each other
Moving index fingers towards each other
Shaking hand
Shaking index finger
Shaking thumb and index fingers
Switching between fingers placed on the screen
Up & down motion with both hands
Using all fingers to hover
Waving hand horizontally across screen
Wiggles finger back and forth on screen
Wiggles finger up and down

Figure 28: Example of the hand motion "Moving hand up and down the graph". Here, the participant moves his entire hand up and down the middle axis.

<p><u>Pointing</u> (74) - involved a variety of pointing methods, whether pointing once, or continuously, or pointing with various different fingers and hand positions (e.g. palm up) (see <i>Figure 29</i>)</p>	<p>Hover pointing Hover pointing continuously with hand Hover pointing continuously with index finger Pointing across screen to partner's side Pointing downwards Pointing hand toward screen Pointing with index Pointing with index and middle Pointing with thumb Pointing with thumb and pinky Pointing with index and pinky Pointing with palm up Pointing with 3 fingers Pointing with back of index finger Pointing with index and middle fingers</p>
---	--

Figure 29: Example of the "Pointing" gesture.

<p><u>Spreading fingers apart</u> (13)</p> <p>- involved moving 2 or more fingers from one or both hands away from each other (see <i>Figure 30</i>)</p>	<p>Spreading thumbs and indexes away from each other</p> <p>Spreading all fingers apart (in one hand)</p> <p>Spreading index fingers away from each other</p> <p>Spreading index and thumb away from each other</p>
--	---

Figure 30: Example of the gesture “Spreading index fingers apart from each other”.

<p><u>Swiping or flicking motions</u> (29)</p> <p>- involved using one or more fingers to swipe the screen (see <i>Figure 31</i>)</p>	<p>Double tap & flick an item</p> <p>Flicking items with index and middle fingers</p> <p>Flicking items with index finger</p> <p>pressing and flicking an item</p> <p>Swipe down on a point of interest</p> <p>Swipe horizontally with one finger on a point of interest</p> <p>Using multiple fingers to swipe</p>
---	---

Figure 31: Example of the gesture “Swiping motion with fingers” (from left to right on the screen).

<p><u>Touching a point of interest</u> (38)</p>	<p>Double tapping</p> <p>Pressing on point of interest</p> <p>Rubbing on point of interest</p>
---	--

<p>- involved tapping, rubbing, or pressing on an area of interest</p>	<p>Tapping on point of interest</p>
<p><u>Tracing lines</u> (67) - was done using a variety of fingers either on or away from the screen to follow lines (see <i>Figure 32</i>)</p>	<p>Hover tracing line(s) with index Hover tracing line(s) with index and middle fingers Hover tracing line(s) with pinky Hover tracing line(s) with whole hand Tracing and tapping a line Tracing line(s) with index finger Tracing line(s) with middle finger Tracing line(s) with both index fingers Tracing line(s) with thumb and index fingers Tracing line(s) with thumb and middle fingers</p>
<div data-bbox="203 867 1255 1255"> </div> <p>Figure 32: Example of the gesture “tracing line with index finger”. The participant on the left hand side starts tracing a line from the left of the screen, following it through to the right of the screen.</p>	
<p><u>Placing finger(s) on the screen</u> (65) - involved one or more fingers touching the screen for longer than a tap</p>	<p>Placing index finger on screen Placing index on screen & wiggling it Placing both index fingers on screen Placing index on screen, traces w other index Placing index and middle fingers on screen Placing middle on screen Placing middle on screen; tracing with index Placing pinky on screen Placing thumb and index finger from same hand on screen</p>

	Placing thumb and index fingers from different hands on screen Placing thumb and index fingers on screen while tracing with the other hand Placing thumb and middle finger from same hand on screen Placing thumb on screen Placing thumb and pinky on screen Placing thumb on screen, while using index finger on other hand to trace
--	---

4.5 Intents

The 11 intent categories created from our coding process are shown in *Table 5*.

Table 5. Intent categories and their respective intents

Intent Category	Prompted/Emergent Intent	Intents within the category
To placeholder	Emergent	To placeholder To placeholder 2 points close together To placeholder 2 points far from each other To placeholder 2 points while tracing lines from those points Multiple placeholdering
To show	Emergent	To show a range of lines To show areas that are not relevant To show partner the correct spot To show partner which lines he's referring to To show the area he's referring to To show the right point of interest

To highlight it	Prompted	N/A
To compare them	Emergent	N/A
To delete it	Prompted	N/A
To find a point on the line	Emergent	N/A
To gain more information about it	Prompted	N/A
To pan	Prompted	N/A
To rearrange items	Prompted	N/A
To zoom in	Prompted	N/A
To examine	Emergent	N/A

4.5.1 Emergent versus prompted intents

Out of the 11 final intents, 6 of them were derived from the latter, hypothetical questions in the study. With these questions, we were looking directly at what gesture(s) participants would want to use to achieve a specific intent. To reiterate, these questions were asked in order to gain a wider variety of intents that are common to visualizations that would not have been seen in our study naturally, as we used a visualization that was not touch enabled or interactive. Analysis for these questions was more straight forward, as we knew what the intent was before the gesture was performed, whereas for the task-based reality questions, intents were not directly known, but were assessed and stipulated among the researchers during the analysis process. Intents stemming from the hypothetical tasks can be seen as “prompted intents” as we prompted them from the participants, versus the “emergent intents” that emerged from participants analyzing the data to complete the realistic analysis questions. These prompted intents were: To highlight it, To delete it, To gain more information about it, To pan, To rearrange items, and To zoom in; and the emergent intents were: To placeholder, To show, To compare them, To find a point on the line, and To examine.

4.5.2 Intent coding granularity

After developing all of the codes, we did a final comparison of our intent categories with Yi et al.'s (2007) intents that were used to generate the questions in our study. We determined that we were justified in deciding to create more fine-grained intents than theirs. We ended up with a total of 11 intent categories, while Yi et al. had 7 intents. Due to the more abstract nature of Yi et al.'s intents, many of our intents would fit into more than one of their categories. *Table 6* shows a side by side comparison of Yi et al.'s intents and our intents. One of their intents may span a variety of our intents; this is most notable for the intent "Select", which can be synonymous with "To placeholder", "To show", "To highlight it", and multiple other intents. Examining these two sets of intents showed that their intents are more interaction-oriented, whereas ours are more goal-oriented. Their intents aren't that helpful for our purposes because we found the intents that will bond to the gestures are much more specific, and also, we were not focused on interaction techniques, but on gestures. Their list of intents connects users' intentions with the interaction techniques performed to achieve them. They define interaction techniques in information visualization as "the features that provide users with the ability to directly or indirectly manipulate and interpret representations" (p. 1225). This is an important point, as our study aimed to define gestures according to our users' specific goals. Our intents were strictly defined as what a user wants to accomplish, without incorporating any interaction techniques. The separation of gesture and intent was the focus of our study, and this is the main reason why using their intents in the analysis process would not have given us the relationship between gestures and intents that we were looking for (although it was useful in the study design phase of generating questions).

***Table 6.* Comparison of Yi et al.'s (2007) intents to ours. Multiple of our intents fit into one of theirs.**

Yi's intent categories	Our intent categories that align with Yi et al.'s
Select	To placeholder To show To highlight it To delete it To find a point on the line To gain more information about it To rearrange items To zoom in To zoom out To examine
Explore	To pan
Encode	(Not addressed in our study)
Abstract/Elaborate	To zoom in
Filter	To delete it
Connect	To highlight it To find a point on the line To compare (lines)

4.6 The relationships between gestures and intents

We created a co-occurrence table on Atlas.ti that shows the frequency (co-occurrence) with which each gesture was paired with each intent (*Table 7*). We highlighted co-occurrences that were 5 and above in red, to showcase the common occurrences. We used the cut-off of 5 because it seemed like a natural breakpoint where the data changed. In this way we could identify frequent gesture-intent pairings.

Table 7. Co-occurrence table depicting the frequency with which each gesture is performed to achieve each intent

	G- Circling something	G- 2-Person gestures	G- Dragging item	G- Hand motions	G- Pointing	G- Spreading fingers	G- Swiping/flicking	G- Touching a POI	G- Tracing line(s)	G- Placing fingers on screen	TOTALS
I- To placeholder	0	13	0	0	1	0	0	0	1	53	68
I- To show	1	1	0	1	13	0	0	2	0	0	18
I- To "highlight" it	2	0	0	0	0	0	2	11	3	1	19
I- To compare them	0	26	0	0	2	0	0	0	14	3	45
I- To delete it	0	0	1	0	0	0	9	4	0	0	14
I- To find a certain point on the line(s)	1	15	0	0	2	0	0	0	34	4	56
I- To gain more info about it	0	0	0	0	0	0	1	14	0	0	15
I- To pan	0	0	0	0	0	0	16	0	0	0	16
I- To rearrange them (items)	0	0	13	1	0	0	1	0	0	0	15
I- To zoom in	0	0	1	0	0	12	0	1	0	0	14
I- To examine	0	1	0	5	12	0	0	0	1	1	20

We found 15 major co-occurrences as seen in *Table 7*:

- 1) 2 person gestures to placeholder
- 2) 2 person gestures to find a certain point on the line
- 3) 2 person gestures to compare lines
- 4) Dragging items to rearrange them
- 5) Hand motions to examine things
- 6) Pointing to show something
- 7) Pointing to examine things
- 8) Spreading fingers to zoom in
- 9) Swiping/flicking to delete items
- 10) Swiping/flicking to pan
- 11) Touching a point of interest to highlight it

- 12) Touching a point of interest to gain more information about it
- 13) Tracing lines to find a certain point on the line
- 14) Tracing lines to compare them
- 15) Placing fingers on the screen to placeholder

Grouped by gestures, we explain the co-occurrences in more detail:

- 1) 2 person gestures involved a combination of both participants placing their fingers on the screen and/or tracing lines simultaneously. As seen in the table, these actions were most commonly performed with the intent To placeholder, To compare lines, and To find a point on the line. This was all part of the collaborative process of working together to analyze the data and find answers to the questions that were being asked. Due to the heavy traffic areas on the graph (with many lines being very close in proximity), it was helpful to have a partner in order to keep track of places that needed to be compared, while searching for other points of interests at the same time.
- 2) Dragging items involved pressing on an item (i.e., an offense category on the graph) and dragging it up or down with one's finger. Dragging was almost always done To rearrange the ordering of items: e.g., to move an offense category up or down on the axis.
- 3) Hand motions involved a variety of movements including moving hands vertically or horizontally shaking hands, or wiggling fingers, and these motions were usually done To examine things. For instance, as one participant was looking down the screen to find the right age category, he moved his hand in a downward motion accompanying his gaze.
- 4) Pointing involved a variety of pointing methods, whether pointing once, or continuously, or pointing with various different fingers and hand positions (e.g., palm up) and was done mostly To show something, and To examine things.

- 5) Spreading fingers involved moving 2 or more fingers from one or both hands away from each other, and was done To zoom in.
- 6) Swiping or flicking involved using one or more fingers to swipe the screen and was done To delete items off of the screen (such as offense categories) and also To pan in order to see more of the graph that is not visible on the screen at the moment.
- 7) Touching a point of interest involved tapping, rubbing, or pressing on an area of interest and was done either To highlight it, or To gain more information about it.
- 8) Tracing lines was done using a variety of fingers either on or away from the screen and was mostly used To find a point on the line, and also To compare (lines).
- 9) Placing fingers on the screen involved one or more fingers touching the screen for longer than a tap and was done almost always to placeholder.

It should be noted that of these 15 major co-occurrences seen in the co-occurrence table, 6 of them emanated from the hypothetical set of questions in the study where we determined how participants would go about getting a certain intent accomplished (what we referred to earlier as “prompted intents”). These are displayed in *Table 8*.

Table 8. List of gestures with corresponding intents for those intents that were prompted from the hypothetical tasks

Gesture	Intent
Dragging items	To rearrange them
Spreading fingers apart	To zoom in
Swiping/flicking	To delete To pan
Touching a point of interest	To highlight it To gain more information about it

4.7 Important take-away

The important knowledge to take from this chapter is that our study shows that there are certain gestures that are used in a collaborative context to accomplish certain intentions, and these occurrences are pretty consistent (there is not a lot of variability as seen in the high numbers in the chart). Consistency is further enhanced by the finding that there were many instances where there was a 1:1 ratio mapping of gesture to intent. For example, the following gestures were done to accomplish one specific intent each: dragging items (to rearrange), spreading fingers (to zoom), hand motions (to examine), and placing a finger on the screen (to placeholder).

The gestures *Pointing* and *Tracing lines* were done to accomplish two separate intents each, but these intents were quite similar. For *Pointing*, these intents were *To show something*, and *To examine*. When we point at something, it is generally an area where our attention is focused, and allows the other collaborators to see where our attention is focused at that moment. Therefore, we can think of pointing as being performed to either draw attention to a particular area, or as accompanying our own visual or attentional processes in the moment.

Tracing lines was done *To find a certain point on the line*, and *To compare (lines)*. These are similar intents, as the goal is to determine what information the line is presenting, and how this compares to the information presented by other lines. It is quite common in graphical data analysis to compare data points, and this first involves finding the data points one wishes to compare. We can even view this process as beginning with locating the first data point (corresponding to the intent of *finding a point on the line*), and ending with locating other data points to compare them (corresponding to the intent of *comparing lines*). This end goal of comparing data points may also be performed in one step, by tracing multiple lines

simultaneously to obtain and compare their data, as we saw many participants doing (which is where the mapping of the gesture tracing lines with the intent being to compare them was seen). Even when comparing lines, the ultimate goal is to find out what the data point is (or what the numerical information is — e.g., the arrest rate or age in this study), and therefore, the intents of *finding a point on the line* and *comparing lines* are interrelated. We explore this relationship further in the design chapter: when a participant traces one or more lines, regardless of whether the intent is to find out information for one line, or multiple lines (in order to compare them), the participant's intention would be more readily realized if the system displayed potentially relevant numerical data for them.

5 Chapter: Design Approach

A lot of what was observed during this study involved participants being able to see what the other was doing while they were working to complete their given tasks. This is done during collaboration so that people have an idea of what the other people are doing. For example, when performing gesture-intent pairing of placing fingers on the screen to placeholder, they would attempt to do this in a manner in which the other participant could see what was being done (i.e., by not blocking the view). Sometimes when multiple spots on the graph were being placeholder it was hard to keep track of them all and to see all of the points for comparison purposes, as participants' arms/hands occasionally got in the way. It was therefore an important part of the collaboration to see what the other person was doing. We can think of this as awareness of what the other collaborators are doing. In a collaborative setting, gestures might need to be more noticeable because they are more likely than not being performed with or for an "audience" so to speak. Furthermore, but not seen in this study, since our graph was static and not interactive, we can assume that it is important for collaborators to not only maintain awareness of what the other is doing, but also to maintain awareness of the effects that any actions have on the graph. Therefore, from an interaction perspective, it might be more effective to make the gestures more visible on the screen, rather than needing the collaborators to make their own gestures more elaborate for the rest of the team.

In this chapter, we will discuss this issue in more detail with reference to the gesture-intent pairings we found in the previous chapter, and outline some design ideas that could be helpful in these collaborative situations. This relates back to our final research question of *how knowledge of gesture-intent relationships can influence design ideas for interactive collaboration with a parallel coordinates graph on a large touch screen*. We begin with the idea

of design patterns which we will draw on indirectly as we discuss our design ideas. We will then present our design ideas for the intents we found in the previous chapter, with a special emphasis on collaborative design. We will conclude with a summary of the designs.

5.1 Design patterns

In their book on pattern languages, Alexander, Ishikawa, and Silverstein define patterns within the field of designing for towns, buildings, and construction:

Each pattern describes a problem that occurs over and over again in our environment and then describes the core of the solution to that problem in such a way that you can use this solution a million times over without ever doing it the same way twice (1977, p. x).

More recently, Tidwell (2011) describes patterns specifically within the field of designing interfaces:

Patterns can be a description of best practices within a given design domain. They capture common solutions to design tensions (usually called “forces” in pattern literature) and thus, by definition, are not novel. They aren’t off-the-shelf components; each implementation of a pattern differs a little from every other. They aren’t simple rules or heuristics either. And they won’t walk you through an entire set of design decisions... (p. xviii).

According to Tidwell, patterns allow for the flexibility of being creative because they capture common structures of design (e.g., a list layout) without being too specific about the details. In *Figure 33*, Tidwell describes the pattern called “Breadcrumbs.” In the image we can see a picture of the pattern, as well as a description of what the pattern entails, and when to use it. Furthermore, Tidwell also discusses why and how you would use the pattern, provides

additional pictorial examples, and discusses patterns that may be related or linked to this pattern (not shown in *Figure 33*).

Figure 33. Example of the pattern, "Breadcrumbs", including a description of what the pattern is and when to use it. Source: Tidwell (2011).

5.2 Outline of chapter

This is Tidwell's style of explaining pattern languages, however, for this chapter, we will take a different approach. We will first describe what we identify as a major problem when collaborating around a screen, and list 5 possible design strategies (similar to the concept of patterns, as they aim to resolve a problem) that may resolve this problem. We will then begin the following section with a recap of the intent-gesture pairings found in the analysis chapter, as these will be the basis of our design ideas and patterns. We will then describe each intent-gesture pairing by intent, describing what the user's intent was and how it was accomplished in the majority of cases (i.e., what gestures were performed to achieve the intent), as well as how the system providing these types of interactions could support these gestures, and finally, the design strategies that may be used for the intent.

For each intent, we will describe the following: (in section 5.4):

- Description of the intent
- Gesture Description: Description of the gesture(s) used to perform this intent
- System Functionality: How the system should support these gestures
- Collaborative Design Strategies Used: How the system could enhance the gestures to make them more visible to all collaborators

5.3 The problem with collaborating on an interactive screen

Drawing on the information provided from our background chapter, we identified that one of the major problems in a collaborative setting is for group members to create a common understanding, and for designers to devise methods of interaction that allow for effective collaboration (Buttler et al., 2011). Additionally, when collaborating around technology, the technology can be seen as a medium through which our interactions with other participants are informed – i.e., the technology potentially enhances our communication (Vatrapu & Suthers, 2009). When technology is used to support collaboration, the interface becomes another entity in the social interaction equation that must be taken into account. The interface must therefore be designed to facilitate person-to-person interactions in a way that enhances collaboration, while also displaying shared information effectively, and enabling the achievement of a common understanding by all participants.

We sum up these challenges as a main problem of *allowing all collaborators to understand what has happened with the interface any time someone interacts with it*. Wigdor and Wixon (2011) also identify this as a problem, where the system should support consequential communication by making changes clear to all collaborators. For example, when a person uses two hands to zoom in on a map, everyone should be able to easily see how and why the zoom

level changed. In order to maintain awareness of the analysis process while collaborating around a visualization, it is important that all collaborators are aware of each change that is made to the visualization, whether it is zoomed, or whether various data points are deleted, for example. We can assume that everyone's attention will not always be focused on the visualization every time a gesture is performed (for example, they might look away to take notes or become distracted by something else), and therefore, to display any changes effectively, these changes must be more visible than in a single-user setting. This is a challenge specifically within the field of the collaborative use of technology, and our design suggestions will be the types of patterns that could address this problem.

We determined 5 different strategies to address the main problem of collaboration awareness. These strategies are:

- 1) Leaving a trace of the gesture on the screen
- 2) Showing the motion of the gesture
- 3) Glowing/blinking animations
- 4) Having a small overview box presented on the screen
- 5) Using breadcrumbs to show what has been done on the screen

They will each be explained in the following section with regards to each intent that would benefit from its use. They will be discussed further in the discussion chapter regarding possible design issues.

5.4 Interaction Techniques & Collaborative Design Strategies for the Intents

Referring back to our co-occurrence table in *Table 7*. Co-occurrence table depicting the frequency with which each gesture is performed to achieve each intent there were 15 gesture-

intent pairings, with a total of 11 different intents. These intents are as follows (underlined), with their corresponding gestures:

- 1) Zooming in: Spreading fingers apart
- 2) Panning: Swiping/flicking
- 3) Deleting items: Swiping/flicking
- 4) Rearranging items: Dragging the items
- 5) Highlighting: Touching a point of interest
- 6) Finding a certain point on a line: Tracing lines; 2 person gestures
- 7) Comparing lines: Tracing lines; 2 person gestures
- 8) Placemolding: Placing fingers on the screen; 2 person gestures
- 9) To gain more information about something: Touching a point of interest
- 10) Showing/drawing attention to: Pointing
- 11) To examine: Pointing; Hand motions

1) Zooming

Zooming is an intent that increases or decreases the scope of what is being seen. In parallel coordinates plots, there are often “high traffic areas” that are very concentrated with data lines, making it difficult to see through the clutter. Distinguishing between lines and various points in these areas proved to be a difficult task for participants, causing them to strain their eyes to differentiate and compare separate lines. The clutter of these areas can clearly lead to decreased efficiency of task completion as well as a possible decrease in the accuracy of analyses.

Gesture Description: Participants were asked to accomplish a zooming intent on the screen. They preferred to zoom in by spreading their fingers from opposite hands away from each other due to the big size of the screen. Most participants used both of their index fingers to zoom, while others used both their index and middle fingers of both hands.

System Response: These results suggest that to support zooming on a large touch screen, the system should recognize the gesture of two or more fingers moving apart from each other simultaneously, and adjust the display according to the extent of the movement (the shorter the distance the fingers spread apart, the less zooming that should occur).

Collaborative Design Strategies Used: In a collaborative setting, the zooming interaction may need to remain visible shortly after being performed in order to allow all collaborators to see why the visualization has changed. To reiterate, we must assume in a collaborative setting that not all collaborators will be looking at the screen at all times and therefore an important design consideration in these settings is to have the system make gestures and interactions on the screen more visible. For example, the system could leave a trace of someone's fingers on the screen for a short period, or provide some type of feedback showing others what just happened, so the other collaborators can see why the display is changing in case they were not attending to the other person's fingers.

Another alternative that a couple of participants had suggested involves a less traditional zooming method (and therefore was not a commonly seen occurrence, but we felt it was relevant to mention here in this design section). This method may be used for more specific/precise zooming of a certain area, where users select the exact area they want to enlarge by drawing a box/circle around it and then tapping it in the center. This would cause the area selected to zoom in and cover the whole area of the screen, with a small box on the screen displaying an overview

of the entire graph. This is akin to the focus + context method familiar to visualizations (Baudisch, Good, & Stewart, 2001). This method could potentially be better in a collaborative setting, as it is easier to see what a user did after the interaction was made. The box or circle drawn around the zoomed in area would remain visible in the overview section of the screen, and therefore all collaborators can easily see what section of the visualization was zoomed. This could also be used in combination with the previously mentioned method of zooming by spreading the fingers apart and leaving a trace of the gesture, while also having an overview box presented on the screen.

A possible method that could have great benefit is to use the “Breadcrumbs” strategy, where a bar at the top of the graph could display all the changes that have been made to the graph in the order they occurred (refer back to *Figure 33*). For example, when zoomed, the breadcrumbs bar would say “zoomed”, and in addition, perhaps clicking on this breadcrumb might re-enact the zoom for those that did not see it occurring in the first place. Combining this technique with the box overview method described previously could also enhance understanding of the change to the graph, as having multiple ways to understand the change might be more beneficial than only one method.

2) Panning

Many times a visualization may be larger than the space provided on the screen it is being displayed on, and users may need to pan the screen in order to see parts of the display that are “off screen”.

Gesture Description: When accomplishing a panning intent, participants used a swiping gesture with one to five fingers on one hand in order to “sweep” the display to the side, or up/down.

System Response: The system should therefore be able to support this interaction by recognizing the swiping gesture and moving the visualization accordingly. Because the swiping gesture is also used to delete items off of the graph (refer to the co-occurrence table in *Table 7*), the system should be able to distinguish between these two intents based on where on the graph the user is swiping. When the gesture is performed on a larger area (rather than on a specific item) the visualization should be panned.

Collaborative Design Strategies Used: In order to allow all collaborators to see that a panning interaction has been performed on the graph, the system could leave a trace of the fingers on the screen to depict the panning gesture. Alternatively, arrows may be used to show the direction that the graph was panned (as demonstrated in *Figure 34*), or arrows may be shown pointing towards the direction that the graph was initially panned from. This could also provide an easier way of remembering how to get back to the original view of the graph for all users. As mentioned for zooming, it could be helpful to have an overview section in the corner, showing the entire graph and the section that is currently being displayed on the screen to maintain group awareness of changes.

Figure 34. Example of using arrows to show the direction the graph is being panned. The further away the arrow is from the fingers, the more ghostly it becomes until it disappears.

The Breadcrumbs strategy can also be implemented here, with the breadcrumbs bar saying “panned.” Once again, clicking on this breadcrumb could re-enact the pan, showing the initial position of the graph and its movement to the final panned position. This might become confusing if multiple panning interactions are performed simultaneously, but perhaps the system could list only the most recent pan, or perhaps list all of the panning interactions when clicked on (in a drop down menu for instance).

3) Deleting

During the process of analysis, users may find that there is no longer a need for certain data points on the graph anymore and may want to delete them.

Gesture Description: When deleting an item (an entire offence and all of its data points) from the graph, the majority of participants used a swiping or flicking gesture with their index finger, or with both the index and middle fingers, from right to left, over the name of the offence. In other words, they were “flicking or swiping the item off the screen” (as the screen edge was to the left of the words).

System Response: The system should support the swiping/flicking gesture and recognize at least one finger pattern in a straight line motion from either left to right, or right to left, as long as it is on a meaningful part of the display (such as on an offense category in our visualization). The system would then delete all information connected with the swiped item. Another alternative suggested by a few participants was to select the item to be deleted, and then be presented with a pop up menu, providing the option of deleting the item.

Collaborative Design Strategies Used: Once again, in a collaborative setting we should assume that not all collaborators will see what gesture and interaction was just performed, and might be confused about the change to the visualization after. The design of the delete interaction technique could potentially be enhanced by including a trace of the gesture, or perhaps a ghostly presence of the deleted item(s) for a while after being deleted, for example, as seen in *Figure 35*. This ghostly presence of the item on the screen might explain its absence to the other collaborators.

Figure 35. Example of a ghostly presence being left behind of the deleted item “Burglary” at the top of the graph.

The use of the “Breadcrumbs” strategy could potentially be very useful for this intent, as in the breadcrumbs bar there could be specification of what was deleted. This could be displayed within the bar itself, or when the user clicks on the breadcrumb “deleted”.

4) Rearranging

While the information on the graph is provided in a certain arrangement, users may want to rearrange the items (e.g., the offences) on the graph to suit their needs; for example, to

categorize offences into categories by placement on the axis (more serious offences at the top and less serious offences at the bottom).

Gesture Description: When asked to perform this task, participants used a dragging gesture to place an offense where they wanted it to move. The gesture began with pressing on the name of the offence and then dragging it up or down the axis to the desired position. Some participants preferred to drag the item around the other items (instead of making a direct straight line movement up or down) so as to not cross through other items. (While this was not coded as a separate gesture due to the high rate of similarity, we bring this point up as an example of how gestures may vary slightly, and will be discussed further in the discussion chapter. Participants expected that the graph would automatically readjust itself and the other offences would move up or down to make room for the offence being re-positioned.

System Response: The design of a rearranging interaction technique could work as follows: when the user presses on an item (i.e., *not* a short tap) the item can “bounce up” demonstrating that the user has the affordance to move the item to a position of his or her liking. Wigdor and Wilson (2011) suggest that in designing NUIs, movable elements should use visual feedback (depth) to acknowledge objects or controls that users successfully touch by moving the item toward the user along the z-axis (in our example, this would be akin to the item bouncing towards the user). The item can then be moved with the user’s finger, as long as the gesture is an up or down motion on the same axis. Since some participants preferred to drag the item around the other items, the design should accommodate for such variability of movements as well: as long as the user’s finger ends up on the same axis, the item should be moved accordingly, even if the gesture path includes making a semi-circle rather than a straight line to get to the desired end position.

Collaborative Design Strategies Used: To assist the other collaborators in understanding what gesture was performed, the visualization can leave a ghostly trail behind the item as it moves from its original position to its new position. Perhaps an arrow would be suitable to depict the movement and the direction of the movement more effectively.

“Breadcrumbs” can also be used for this intent, where clicking on the breadcrumb “rearranged” would re-animate the rearrangement interaction, showing the items in their original positions moving to their new positions.

5) Highlighting/Finding a point on the line/Comparing lines

While making comparisons and analyses, users may want to highlight important data points, or points that they are referring to. In this section, finding a point on the line, and comparing lines will be sub-intents of highlighting because, from a design perspective, we can support slightly different intents with the same design. Highlighting lines allows for users to easily find certain points of interest on the lines, as well as making comparison visually more effective. Therefore, this section will discuss highlighting as a joint intent including these 3 intents.

Gesture Description: When asked about highlighting data points, most participants pressed on a point of interest (whether an entire offence category or a specific arrest rate) with their index finger. Therefore, when designing a highlighting function, the system should recognize tapping a line or item (for example, an offence or age category in our graph) as a point of interest, and highlight it. If an offence or age category is highlighted, all data points for that offence will be highlighted, whereas if one line is selected, only that line will be highlighted. When finding a point on a line, or when comparing lines, participants traced the desired lines,

and also placeheld certain points on the lines. Therefore, the system should also complete the highlighting intent when any line is traced. The issue of placeholding for comparison purposes will be discussed later.

The most frequently performed gesture in this study was tracing lines to find a certain point on the line (i.e., an arrest rate or age number). Highlighting lines could cut down analysis time and increase its efficiency, as many participants reported difficulty in reading the exact arrest rate, especially in highly cluttered areas. Instead of having to trace a line to find a desired point, users can easily tap on the line of interest at any point along the line (or trace any part of the line) and the entire line would be highlighted (tapping the line again could take away the highlighting effect). Additionally, users commented that analysis would be less straining if the graph could provide the information for them without them having to search for it. Therefore, if a user highlights a line, the arrest rates and the age could appear on the screen above the line. Instead of having to estimate what the arrest rate is, the graph could automatically provide the number on top of the line when highlighted (along with the age). For example, it could appear as: A(15) 5000; to show that for age 15 the arrest rate is 5000 arrests per year. The color of the text could be the same color as the offense category it belongs to. By doing this, the offense category doesn't have to be written out along with the age and arrest rate as users would possibly recognize it by color. This would save room on the graph in order to decrease clutter and redundant information.

Collaborative Design Strategies Used: The ability to highlight various lines on the graph will not only enhance the effectiveness of the analysis, but will also enhance the interaction between the users (i.e., the collaboration). When one or more users highlights an area of interest, the other collaborators or observers will become aware of what is being discussed or analyzed at

that point of time. If one person is doing the majority of the analysis on the screen, the other collaborators might be able to easily follow along with the analysis process, as the user's interactions will highlight the areas of interest. While discussions may generally aid in the ability of collaborators to understand what is being assessed, the process of highlighting could make understanding even easier, due to the added visual component. In a collaborative setting, highlighting may be shown with more emphasis than in a single user setting. For example, the highlighting effect could be a bright glowing line versus a less visually noticeable effect such as the line being surrounded by dashes (as is a common technique used when something is selected in many applications). Additionally, the glowing effect could be prominent for a short time after the line is highlighted to allow all users to see which line was most recently highlighted (assuming that multiple lines had already been highlighted previously).

“Breadcrumbs” could also be useful, as the breadcrumbs bar would display that something was highlighted, and clicking on it would detail what was highlighted, possibly with re-enactments of the highlighting interaction. This could make it easy for users to understand what has been highlighted recently if they were not paying attention when the highlighting interaction first occurred.

6) Placeholder

Placeholdering was one of the most common gestures performed by participants to keep track of the various points of comparison. Placeholdering is a common gesture we use even when interacting on paper. We use one or more fingers, placed on a spot of interest, so that we can keep a marking of that place, while diverting our gaze or attention elsewhere. It is then made easy to return to that place when we are ready because of our physical marker.

Gesture Description: Placeholding was done by placing a finger on each point of comparison until the analysis was complete.

System Response: The design for placeholding should be as natural to this gesture as possible: users can press on a point of interest for longer than one second (to distinguish it from a tap), and a lighted ring could appear around the point on the axis that was pressed. This is distinct from highlighting a line since the specific point pressed on the line would be encircled (and is only possible on either the arrest rates or age axes), while a tap highlights (“lights up”) the entire line. Users should be able to placehold a point on a line even if the line is already highlighted (one action shouldn’t cancel out the previous action), as there may be instances where only specific points on a line are important at the moment, but in the general analysis all of the data represented in the line is relevant to examine.

Collaborative Design Strategies Used: The point of placeholding is to mark places you are examining, so a circle surrounding these specific data points makes them easily accessible to the eyes when users are ready to come back to these points. The circles might stand out from the rest of the graph, allowing all collaborators to see what the important data points are at the moment. In a situation where one collaborator placeholds a new data point and there are already multiple data points placeheld on the screen, the system can provide feedback to the other users by having the circle blink a few times to display that it is the newly placeheld point on the graph (an example is seen in *Figure 36*), or having it glow more, akin to the highlighting technique mentioned previously.

Figure 36. Example of the blinking circle when selecting a new point to placehold.

The “Breadcrumbs” bar would potentially be very useful to depict placeholding, as clicking on this breadcrumb could display a list of all recently placeheld items, and provide the ability to reanimate these interactions when clicked on.

Special case: There were many instances of both participants tracing lines and placeholding simultaneously (2 person gestures) with the intents being to placehold and to find certain points on the line to compare them. The system should therefore support *collaborative* (2-person) highlighting and placeholding as well. This could work as follows: if user 1 presses on an offense on the offense axis to highlight it, then user 2 can press on specific age categories to be placeheld on the age axis. This will result in the graph automatically placeholding the selected data for that offense category, while still highlighting all lines within the selected offence. For example, if the users select ages 15 and 17 for burglary using this method, then all lines for burglary will be highlighted (by user 1’s actions), and the arrest rates for ages 15 and 17 will be circled (placeheld; by user 2’s actions) to be more easily seen. Additionally if user 2 double taps on ages 15 and 17 then these lines will return to normal lighting (i.e., they will no longer be highlighted), with user 1’s previous highlighting action removed. This type of action would be performed to narrow down the lines within burglary that are highlighted.

7) To gain more information about something

While analyzing graphical information, users may want to gain more information about certain aspects of the graph, such as a definition or a description of one of the data variables.

Gesture Description: The majority of participants used a tapping gesture to touch the point of interest (i.e. the item/data point that they wanted to gain more information about).

System Response: The system could support this intent by providing the user with a pop up screen giving more details about an item or data point that is selected. To distinguish it from placeholding (a longer press), the system could recognize a shorter press (more similar to a tapping gesture) as wanting to gain more information about something, while a long press on a point could placeholder that point.

Collaborative Design Strategies Used: To allow all collaborators to understand why information is being presented on the graph, the selected data point could be encircled or highlighted (similar to when a point is placeheld) in addition to the pop up box beside it. This will allow all users to see exactly what they are getting more information about.

Figure 37. Example of encircling an item when it is selected to gain more information about it.

Additionally, in a “Breadcrumbs” bar, it could be shown that an item was selected for more information, by perhaps saying, “More info: Burglary”. Clicking on this breadcrumb could light up the item that was selected again, in case more visual stimuli are needed for collaborator’s awareness.

8) Showing/drawing attention to something and Examining something

Both of these intents involved gestures that were performed away from the screen (pointing and various hand motions), and therefore are not within the scope of this design chapter of our study, as we are focusing on natural gestures in a collaborative setting performed on the screen.

5.5 Intent relationships

In this section we will outline the various intents described as well as how they might relate to each other.

Zooming and Panning: These intents are related in the sense that when performed, the view of the diagram is changed.

Highlighting, Finding a point on the line, Comparing lines, and Placeholding: As discussed previously, these intents are related because they all share a similar higher level goal of finding various data points, and maintaining an easy view of these points, usually for the purpose of comparing them.

Deleting and Rearranging items: These intents are similar, as it changes the ordering or arrangement of items displayed on the graph, whether because some items are removed from the graph, or because they are reordered in position.

Showing/drawing attention to something and Examining something: These intents are similar because they both involve some sort of bodily depiction of where one's attention is currently focused, or alternatively, where one wants the other's attention to be focused.

Gaining more information about something: This intent is not related to the others as it involves a higher level function of getting information that is not visible on the graph unless requested.

5.6 Summary

We discussed the types of interaction techniques that a system could use to allow for certain intents to be completed. We also determined 4 collaborative design strategies that may help with the problem of collaborator's awareness within collaborative settings. These are:

- 1) Leaving a trace of the gesture on the screen: We suggested that perhaps a gesture can linger on the screen for longer than usual in a collaborative setting. This can include a faint or ghostly presence of the path that the fingers followed on the screen. This design strategy could be used for the following intents: *Zooming, Panning, Deleting, And Rearranging*.
- 2) Showing the motion of the gesture: This strategy is similar to the previous strategy, although perhaps more animated. This could include having more visual animations whenever a gesture is performed, such as having any type of swiping motion on the screen creating arrows that follow the finger or some other type of motion icon, such as the comet shaped icon. This design strategy could be used for the following intents: *Zooming, Panning, Deleting, And Rearranging*.
- 3) Glowing/blinking animations: This type of strategy would signal to viewers that something new has been done to the graph by drawing their attention to the area. This

could be used for *Highlighting, Placeholdering, and possibly for Gaining more information about something* as well. When a line or data point is highlighted, the line or data point can blink a few times to show that it is the newly highlighted or placeheld entity. If something is selected and a popup box provides more information about it, the selected item can blink to signal to others that the system is currently providing them with information about this specific item.

4) Providing an overview box on the same screen: This strategy would provide a small box on the screen that would show the entire graph and depict what area of the graph is currently being seen on the screen. This could best be used for *Zooming and Panning*.

5) Using breadcrumbs to show what has been done on the screen: This is a commonly used strategy in many applications and web browsers to show a path of where the user has travelled within the digital world and may be beneficial in a collaborative setting. It could be used to show all of the interactions that have been performed on a visualization so that all collaborators have a point of reference to look at in order to understand why the visualization has changed if they were not paying attention to it at the moment the changes were made. This type of design strategy could be used for all of the intents mentioned in the design section: *Zooming, Panning, Deleting, Rearranging, Highlighting, Placeholdering, and Gaining more information about something*.

6 Chapter: Discussion and Conclusions

In this final chapter, we provide a summary of our results and design suggestions and provide possible explanations for our results as well as possible concerns surrounding our design strategies.

6.1 Summary of our study

We began this study with an interest in collaboration around large screens, as these types of displays can facilitate new kinds of collaborative applications with the space they afford. This is particularly true for data analysis, which is a type of collaborative work that we were interested in, specifically centered around parallel coordinates plots, which we were familiar with in our lab from a previous study (see Wilson, Brown, & Biddle, 2014). We wanted to create a scripted study where we would outline tasks and observe pairs of participants going through the motions of completing these tasks in order to determine the natural types of gestures they would perform. This was an important aspect of our study as we were drawing on the ideas of natural user interfaces (NUIs) and related concepts such as reality-based interaction (RBI; Jacob, et al., 2008) and Blended Interaction (Jetter, Reiterer, & Geyer, 2014).

While many studies focused on interaction techniques (which can be thought of as a combination of user gestures, system responses, and user intents), we were interested in focusing more specifically on the types of gestures people perform to achieve various intents while collaborating on a data graph. The main goal of our study was to determine common gesture-intent pairings that were observed as participants completed various tasks. We were additionally focused on going beyond these gesture-intent pairings to determine the types of interactions that could be designed from these results for collaborative settings. Our research questions were:

- 1) Which gestures are related to which intents?

- 2) Which gestures do people believe to be suitable for different intents?
- 3) How can knowledge of these gesture-intent relationships influence design ideas for interactive collaboration with a parallel coordinates graph on a large touch screen?

Our study resulted in the identification of 11 intents and 10 gestures that were observed frequently. There were two types of intents in this study: we labelled the first set as emergent intents, and the second set as prompted intents. Emergent intents were observed as participants completed analysis and comparison tasks directly related to the data (corresponding to our first research question), while prompted intents were observed after asking the participants to perform hypothetical tasks that included changing the graph in some way, such as panning to see data that is not currently on the screen (corresponding to our second research question).

After creating a co-occurrence table to see which gestures were paired with which intents, we found a total of 15 gesture-intent pairings with co-occurrences of at least 5. This showed that the gestures and intents matched up in most cases. These pairings (organized by gesture) were:

- 1) 2 person gestures to placeholder
- 2) 2 person gestures to find a certain point on the line
- 3) 2 person gestures to compare lines
- 4) Dragging items to rearrange them
- 5) Hand motions to examine things
- 6) Pointing to show something
- 7) Pointing to examine things
- 8) Spreading fingers to zoom in
- 9) Swiping/flicking to delete items
- 10) Swiping/flicking to pan

- 11) Touching a point of interest to highlight it
- 12) Touching a point of interest to gain more information about it
- 13) Tracing lines to find a certain point on the line
- 14) Tracing lines to compare them
- 15) Placing fingers on the screen to placeholder

6.2 Explaining the co-occurrences from our analysis

We believe that we found so many high co-occurrences because people are very similar in their gesture patterns when trying to achieve a specific intent. Referring back to Kendon (2004), gestures are made up of components called gesture units: preparation, stroke, post-stroke hold, and recovery. As an example, when waving a hand, preparation would be lifting the hand up, the stroke would be waving the hand, a post-stroke hold would be leaving the hand in the wave position shortly after the wave ends, and the recovery phase would be placing the hand back down. In our study, gestures were performed with the same components by participants in most cases. For example, when the intent was to find a certain point on a line, preparation involved raising the finger and placing it along the line, the stroke involved tracing the finger along the line, the post-stroke hold involved leaving the finger at the point of interest, and the recovery involved removing the finger from the graph back to a resting position. As another example, when panning, preparation involved lifting the hand and placing two or more fingers on the screen, the stroke involved swiping the screen with these fingers, and the recovery involved removing the fingers from the graph once the action was completed.

An important point to note in these findings is that we grouped similar gestures together, for example, swiping with two fingers and swiping with three fingers, which resulted in our high co-occurrence numbers. While these are the same gestures, but involve slight variance in

articulators (what Kendon calls the object(s) performing the gesture; in this case the fingers), we grouped them as the same gesture. This is the same idea that Wobbrock et al. (2009) used when determining their user defined gesture set, which they propose differs from gesture sets in the literature by allowing flexibility in the number of fingers that can be used, rather than binding specific numbers of fingers to specific actions. They state that gestures should not be distinguished by the number of fingers used, as people do not regard the number of fingers they use in the real world, except in skilled activities like playing the piano. We believe that this is an important idea and also helps to explain the reason behind why we were able to find such high co-occurrences in our study between gestures used and the intents they were performed to accomplish (while also allowing us to provide design suggestions that were not too broad using these grouped gesture categories).

While there were some gestures that were slightly different in movement, these gestures had a similar beginning and end phase and were therefore coded as the same gesture. For example, the gesture of dragging an item to rearrange the order (as mentioned in the design chapter) began with touching the item, dragging it to the desired position, and ended with removing the finger from the screen. Some participants chose to drag in a straight line, while others chose to drag in a circular motion around the other items, so as to not pass through them. We can think of the beginning phase of touching the item to be moved as the gesture registration phase (Wu et al., 2006). The *gesture registration* is the first phase of a gesture operation, and it sets the context for succeeding interactions. *Gesture relaxation* allows gestures to be performed with minimal constraints after it is registered (in the gesture registration phase). This places all of the physical tension in performing a gesture in the gesture registration phase, allowing the user to relax the shape and dynamics of gestures afterwards. From a design perspective, as long as the

gesture registration phase is performed in the same manner, and the final phase of the gesture is the same, there should be minimal constraints on what is done in between. This idea provides good reasoning for our grouping of these slightly different gestures as part of the same gesture category, “dragging items.”

In describing simple gestures, Lao et al. (2009) classify the types of movement(s) the hand makes while interacting with the screen. They describe three basic hand movements: pressing, tapping, and dragging. Pressing involves touching the surface and leaving the finger there, tapping involves touching the surface and lifting the finger right away, and dragging involves touching the surface and moving the finger along the surface. When looking at the 10 gestures we found in our study, it is evident that they can be categorized according to these three basic hand movement categories (excluding Hand motions and Pointing as these were not performed on the screen). *Table 9* displays this comparison.

Table 9. Categorization of our gestures with Lao et al.’s (2009) three basic hand movement types. Hand motions and Pointing are not included (from our gesture set) as they are not gestures performed on the screen. 2-person gestures fall into both “pressing” and “dragging” because this category usually involved a combination of placing fingers on the screen and tracing lines.

Hand movement type (Lao et al., 2009)	Gestures found in our study
Pressing	Placing finger(s) on the screen 2-person gestures on the screen
Tapping	Touching a point of interest
Dragging	Tracing lines Circling Dragging items 2-person gestures on the screen Spreading fingers apart Swiping or flicking motions

Perhaps this can help explain why we found very high co-occurrences between gestures and intents. It appears that to accomplish 11 types of intents, there are 3 simple hand movement (gesture) types that are used: pressing, tapping, and dragging. This is most likely due to the fact that these gestures are very basic and well known to us, as they draw from our preconceived notions of how touch systems operate (i.e. our experience with smartphones and tablets). In other words, these are gestures that are natural to users, as they draw on *concepts from the users' familiar reality*, including already well-established digital concepts such as spreading two fingers apart to zoom (Jetter et al., 2014). Therefore, because the basic building blocks of gestures are so well-known to users (everyone in our study had a lot of experience using touch screens), it is not surprising to find that most people will use very similar gestures to accomplish simple tasks/intents. Our findings might not have been as conclusive if we had asked participants to complete more complex intents, such as changing the color scheme of the graph for example.

6.3 Design Concerns

Norman and Nielsen (2010) have stated that gestural user interfaces (GUIs) don't provide the visual cues that show what the technology can afford (they don't easily depict what operations can be performed), which threatens the viability of these systems. In technology-enhanced collaborative environments, because people aren't always paying attention at the right time, it is not only imperative to make visible what you *can* do to the system, but it is additionally necessary to make visible what you *did* do to the system. In our design chapter, we outlined the issue of creating a common awareness of changes to the visualization as the major problem when dealing with collaborative visualizations on a large surface. We structured our design ideas according to the intents we found in our study, with a particular focus on Zooming,

Panning, Deleting, Rearranging, Highlighting (including Finding a point on the line and Comparing lines), Placeholder, and Gaining more information about something.

We discussed at least one way that the system could provide interaction techniques to accomplish each of these intents. In some cases, an intent could have more variability in the types of gestures used, or in the way the intent is achieved. For instance, for panning, all participants swiped their fingers across the screen to achieve this intent, for zooming, most participants spread their fingers apart, but there were other alternatives, such as creating a box around the area to be zoomed, as discussed in the previous chapter. Sometimes a participant would suggest more than one method for how he/she would like to complete an intent, which shows that perhaps a good option when designing interaction techniques on a screen is to incorporate more than one type of interaction to achieve an intent. The users shouldn't be restricted to completing an intent in one way if the system has the ability to allow for multiple types of interactions to complete the intent. This is portrayed well in smartphones, for instance, by providing the ability to get to the camera app in multiple ways, such as pressing the hard button on the outside of the phone, or using the touch interface to get to the app from the lock screen.

While the ability to perform different interactions to achieve the same intent has been effective for single-user products like cellphones, it is unclear how this type of design might affect collaborative systems. Returning to the main problem identified as providing all collaborators with awareness of what has occurred on the screen, we can speculate that an increased gesture set would require more design thinking on the issue. The ability to use multiple gestures could potentially increase the difficulty in making collaborators aware of what occurred because there are multiple ways to achieve the changes that were made to the visualization at any

point. However, this could also be beneficial in a collaborative setting as a larger gesture set could better accommodate the diversity within the way that the collaborators interact with the screen. Some people may prefer zooming by spreading their fingers, while others may prefer the box method as the primary method of zooming. We believe that an increased gesture set might have more benefits than not, as long as the awareness problem is taken into consideration when designing these interactions.

A similar concern to using multiple gestures to achieve the same intent, is using the same gestures to achieve different intents. This was seen a few times in our study, most notably, using a Swiping/flicking gesture to achieve both Panning and Deleting; and Touching a point of interest to achieve the intents Gaining more information about something, and Highlighting something. We believe that it is common for the same gesture to have multiple functions and systems should support this. The distinguishing feature would be where the gesture is performed. For example, on a parallel coordinates plot, when Touching a line, the line should be highlighted, while Touching a labelled category (such as an offense category in our graph) should provide more information about that category. Wobbrock, Morris, and Wilson (2009) had similar results where participants reused gestures for different intents. For example, they found that the gesture used to enlarge something, was also used for zooming in when performed on the background, and also used for opening a folder when performed on the folder. Designing for the ability of using the same gestures to complete different intents aligns with the concept of *gesture reuse* proposed by Wu, Shen, Ryall, Forlines, and Balakrishnan (2006). They state that reusing gesture primitives (the basic components that define a gesture) allows more gestures to be created without requiring more primitive gestures to be defined. This could decrease the time it takes to

learn how to interact fully with a new system, and we believe it is an important aspect that should also be included in collaborative design.

6.3.1 Concerns regarding our collaborative design strategies

We determined 5 design strategies that may help with the problem of collaborator's awareness.

- 1) Leaving a trace of the gesture on the screen: This design strategy could be used for the following intents: Zooming, Panning, Deleting, And Rearranging.
- 2) Showing the motion of the gesture: This design strategy could be used for the following intents: Zooming, Panning, Deleting, And Rearranging.
- 3) Glowing/blinking animations: This could be used for Highlighting, Placeholding, and possibly for Gaining more information about something as well.
- 4) Present an overview box: This could be used particularly for zooming and panning.
- 5) Using breadcrumbs to show what has been done on the screen: This type of design strategy could be used for all of the intents mentioned in the design section: Zooming, Panning, Deleting, Rearranging, Highlighting, Placeholding, and Gaining more information about something.

These design strategies are suggestions and should be tested to determine if they will enhance collaborator's awareness of the status of the visualization. We can pinpoint a number of possible issues that may arise. For the first three strategies (leaving a trace of the gesture, showing the motion of the gesture, and having a blinking animation), we can only imaginatively speculate if these would have any effect on awareness and understanding of changes on the screen. For instance, some users may look away from the screen for longer than the period that the gesture/animation lingers on the screen, and would therefore miss these visual cues entirely.

There is also an issue of how long the animation should remain on the screen, as once the visualization has changed from the implemented gesture, it might be a nuisance to still see the traces of fingers left behind or a blinking animation, and may distract the person(s) currently making the changes on the graph or the other collaborators.

Using breadcrumbs in addition to these animations might better create that sense of awareness and understanding for collaborators, however, there are also issues we can think of that may arise with this pattern. Many times when implemented (in web browsers for example), breadcrumbs allow the user to click on previous ‘crumbs’ to return to that page easily. If this type of interactive breadcrumbs bar was implemented in a collaborative visualization, users would be able to return to previous states of the visualization. This would be beneficial but introduces a new issue of being able to return to a state that occurred in the future of the state that the user previously returned to (going backwards and then forward). What if the user wants to make changes to a previous state? The system would have to change the future breadcrumbs as well, and it may become difficult for collaborators to keep track of what changes were made and where they were made (in what breadcrumb).

It could also be useful to delete some of the interactions within the breadcrumbs bar and is an additional interaction to consider. Perhaps the idea of *cooperative gesturing* (Morris, Huang, Paepcke, & Winograd, 2006) where “the system interprets the gestures of more than one user as contributing to a single, combined command” (p. 1201) could be implemented to delete content within the breadcrumbs bar, as this is an important action that should involve the ‘consent’ of all collaborators. This could also ensure that important interactions to the visualization are not removed. An undo feature would also be an acceptable and probably easier alternative, although not as involving of all members of the collaboration.

6.4 Limitations and future work

There are a number of limitations with our study. One involves the issue of generalizability of our results and design ideas. Since our research was solely focused on parallel coordinates graphs, we are not sure how relevant these findings would be when collaborating on other types of graphs or visualizations. For instance, one co-occurrence we saw frequently was tracing a line to find a certain point on the line, which would not be found in a type of graph that did not include lines or a similar method of displaying the data. Similarly, some of our design ideas might only be relevant to parallel coordinates graphs, such as highlighting a line.

Another limitation is that we did not assess the complexity of the intents and the effects that complexity might have on the gestures used. Wobbrock (2010) asked participants to rate intents on a complexity scale and used logistic regression to determine the effects that complexity had on the types of gestures used, and found that simpler commands more often resulted in physical gestures, while more complex commands resulted in metaphorical or symbolic gestures. In our study, we could have created groups of intents that differed in complexity and looked at how this affects gesture-intent mappings, but this is something that we may consider for future work.

We also did not look at reversible intents, such as zooming out (versus zooming in) and reversing a deletion for example. Adding these types of intents would have increased the number of intent categories we found, and provided us with more co-occurrences and leading to more design ideas. Future work should assess a variety of other intents that we did not address in our study, in order to obtain a larger user defined gesture-intent set.

Finally, our collaborative design patterns are only speculative at this point and due to time and resource limitations, could not be implemented and tested out to find out if they are actually effective in increasing collaborator's awareness and understanding of changes in the graph. Future work should assess these patterns and determine how they affect collaboration and how they can be better refined.

References

- Alexander, C., Ishikawa, S., & Silverstein, M. (1977). *A Pattern Language. Ch. Alexander*. New York: Oxford University Press. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:A+Pattern+Language#0>
- Alibali, M. (2005). Gesture in spatial cognition: Expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation*, 5(4), 307–331. doi:10.1207/s15427633scc0504
- Aslan, I., Uhl, A., Meschtscherjakov, A., & Tscheligi, M. (2014). Mid-air Authentication Gestures : An Exploration of Authentication Based on Palm and Finger Motions. *ICMI '14 Proceedings of the 16th International Conference on Multimodal Interaction, New York, NY, USA*, 311–318. doi:10.1145/2663204.2663246
- Atlas.ti Qualitative Data Analysis. (n.d.). Berlin. Retrieved from <http://atlasti.com/>
- Baudisch, P., Good, N., & Stewart, P. (2001). Focus plus context screens: Combining display technology with visualization techniques. *Proceedings of the 14th annual ACM Symposium on User Interface Software and Technology*, 31–40. doi: 10.1145/502348.502354
- Brown, J. M., Wilson, J., Gossage, S., Hack, C., & Biddle, R. (2013). *A Study of an Intelligence Analysis Team and their Collaborative Artifacts. Synthesis Lectures on Human-Centered Informatics*. Morgan & Claypool. doi:10.2200/S00492ED1V01Y201303HCI019
- Brown, J. M., Wilson, J., Gossage, S., Hack, C., & Biddle, R. (2013). *Surface Computing and Collaborative Analysis Work* (1st ed.). Morgan & Claypool Publishers.
- Buttler, T., Janeiro, J., Lukosch, S., & Briggs, R. O. (2011). Beyond GSS: Fitting collaboration technology to a given work practice. In M. R. S. B. Adriana S. Vivacqua, Carl Gutwin (Ed.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6969 LNCS, pp. 126–141). Springer. doi:10.1007/978-3-642-23801-7_11
- Buxton, B. (2007). *Multi-Touch Systems that I Have Known and Loved*. Retrieved from <http://www.billbuxton.com/multitouchOverview.html>
- Clark, H. H. (2003). Pointing and Placing. In S. Kita (Ed.), *Pointing Where language culture and cognition meet* (pp. 243–268). Hillsdale: Erlbaum. doi:10.4324/9781410607744
- Few, S. (2006). Multivariate Analysis Using Parallel Coordinates. *Perceptual Edge*, 1–9. Retrieved from http://www.perceptualedge.com/articles/b-eye/parallel_coordinates.pdf
- Grudin, J., & Poltrock, S. (2012). Taxonomy and Theory in Computer Supported Cooperative Work. *The Oxford Handbook of Organizational Psychology*, (2012), 1323–1348. doi:10.1093/oxfordhb/9780199928286.013.0040
- Han, J. Y. (2005). Low-cost multi-touch sensing through frustrated total internal reflection. *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology - UIST '05*, 115–118. doi:10.1145/1095034.1095054
- Hill, W. C., & Hollan, J. D. (1991). Deixis and the future of visualization excellence. *Proceeding Visualization '91*. doi:10.1109/VISUAL.1991.175820

- Hinrichs, U., & Carpendale, S. (2011). Gestures in the wild: studying multi-touch gesture sequences on interactive tabletop exhibits. *Proceeding CHI '11 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3023-3032. doi: 10.1145/1978942.1979391
- Hodges, S., Izadi, S., Butler, A., Rustemi, A., & Buxton, B. (2007). ThinSight: Versatile Multi-touch Sensing for Thin Form-factor Displays. *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology - UIST '07*, 259–268. doi:10.1145/1294211.1294258
- Inselberg, a. (1997). Multidimensional detective. *Proceedings of VIZ '97: Visualization Conference, Information Visualization Symposium and Parallel Rendering Symposium*, 1–8. doi:10.1109/INFVIS.1997.636793
- Isenberg, P., Elmqvist, N., Scholtz, J., Cernea, D., Ma, K.-L., & Hagen, H. (2011). Collaborative Visualization: Definition, Challenges, and Research Agenda. *Information Visualization*, 10(4), 310–326. doi:10.1177/1473871611412817
- Isenberg, P., Fisher, D., Paul, S. a., Morris, M. R., Inkpen, K., & Czerwinski, M. (2012). Co-located collaborative visual analytics around a tabletop display. *IEEE Transactions on Visualization and Computer Graphics*, 18(5), 689–702. doi:10.1109/TVCG.2011.287
- Jacob, R., Jacob, R., Girouard, a, Girouard, a, Hirshfield, L., Hirshfield, L., ... 2008. (2008). Reality-based interaction: a framework for post-WIMP interfaces. *Portal.Acm.Org*, 201–210. doi:10.1145/1357054.1357089
- Jetter, H. C., Reiterer, H., & Geyer, F. (2014). Blended Interaction: Understanding natural human-computer interaction in post-WIMP interactive spaces. *Personal and Ubiquitous Computing*, 18(5), 1139–1158. doi:10.1007/s00779-013-0725-4
- Johansen, R. (1991). Teams for tomorrow [groupware]. In *System Sciences, 1991. Proceedings of the Twenty-Fourth Annual Hawaii International Conference on*, 3, 521-534. doi: 10.1109/HICSS.1991.184183
- Jota, R., Nacenta, M. a, Jorge, J. a, Carpendale, S., & Greenberg, S. (2010). A Comparison of Ray Pointing Techniques for Very Large Displays. *GI '10 Proceedings of Graphics Interface 2010*, 269–276. Retrieved from <http://groupiab.cpsc.ualgary.ca/groupiab/uploads/Publications/Publications/2010-RayPointing.GI.pdf>
- Kendon, A. (2004). *Gesture: Visible action as utterance*. New York: Cambridge University Press.
- Kranstedt, A., Lücking, A., Pfeiffer, T., Rieser, H., & Wachsmuth, I. (2006). Deictic object reference in task-oriented dialogue. In G. Rickheit, I. Wachsmuth, & W. de Gruyter (Eds.), *Situated Communication* (pp. 155–207). berlin: Walter de Gruyter. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.69.2012>
- Lao, S., Heng, X., Zhang, G., Ling, Y., & Wang, P. (2009). A Gestural Interaction Design Model for Multi-touch Displays. *People and Computers*, 440–446. Retrieved from <http://dl.acm.org/citation.cfm?id=1671068>
- Levinson, S. (2008). Deixis. In *The Handbook of Pragmatics* (pp. 97–122). Oxford: Blackwell Publishing Ltd. doi:10.1002/9780470756959.ch5
- Lewins, A., & Silver, C. (2007). *Using software in qualitative research*. London: Sage Publications Ltd.
- Lozano, C., Jindrich, D., & Kahol, K. (2011). The impact on musculoskeletal system during multitouch tablet interactions. *Proceeding CHI '11 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 825–828. doi:10.1145/1978942.1979062

- Microsoft. (2012). Microsoft PixelSense. Retrieved May 1, 2015, from <http://www.microsoft.com/en-us/pixelsense/pixelsense.aspx>
- Microsoft. (2015). Microsoft Surface Hub. Retrieved May 1, 2015, from <https://www.microsoft.com/microsoft-surface-hub/en-ca>
- Morris, M. R., Huang, A., Paepcke, A., & Winograd, T. (2006). Cooperative Gestures: Multi-User Gestural Interactions for Co-Located Groupware. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1201–1210. doi:10.1145/1124772.1124952
- MultiTouch Ltd. (2013). *MULTITACTION Migration Guide: SUR40 / Pixelsense to MultiTaction*. Retrieved from http://multitouch.s3.amazonaws.com/multitaction/downloads/guides/MT_SUR40_MultiTaction_migration_guide_v1.0_letter.pdf
- Murugappan, S., Elmqvist, N., Ramani, K., & Lafayette, W. (2012). Extended Multitouch : Recovering Touch Posture , Handedness , and User Identity using a Depth Camera. *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology - UIST '12*, 487. doi:10.1145/2380116.2380177
- Norman, D. (1993). *Things That Make Us Smart: Defending Human Attributes In The Age Of The Machine*. Boston, MA: Addison-Wesley Longman Publishing Co., Inc.
- Norman, D. a, & Nielsen, J. (2010). Gestural Interfaces: A Step Backward In Usability. *Interactions*, 46–49. doi:10.1145/1836216.1836228
- Norman, D. a. (2010). The way I see it: Natural user interfaces are not natural. *Interactions*, 17(3), 6. doi:10.1145/1744161.1744163
- PQ Labs. (2015). PQ Labs: Innovation Beyond PC. Retrieved May 1, 2015, from <http://www.multitouch.com/product.html>
- Raje, R. R., Boyles, M., & Fang, S. (1998). CEV: collaborative environment for visualization using Java RMI. *Concurrency: Practice and Experience*, 10(11-13), 1079–1085. doi:10.1002/(SICI)1096-9128(199809/11)10:11/13<1079::AID-CPE397>3.3.CO;2-X
- Reetz, A., & Gutwin, C. (2014). Making Big Gestures: Effects of Gesture Size on Observability and Identification for Co-Located Group Awareness. *CHI 2014: Proceedings of the 32nd International Conference on Human Factors in Computing Systems, Toronto, Canada*. doi:10.1145/2556288.2557219
- Review Display Systems Ltd. (2012). Optical Bonding improves display performance when using projective capacitive touch screens. Retrieved July 1, 2015, from <http://www.newelectronics.co.uk/electronics-technology/projective-capacitive-touch-screens/41681/>
- Samaroo, R., Brown, J. M., Biddle, R., & Greenspan, S. (2013). The Day-in-the-life Scenario : A technique for capturing user experience in complex work environments. *Conference: Emerging Technologies for a Smarter World (CEWIT), 2013 10th International Conference and Expo on*. doi:10.1109/CEWIT.2013.6713761
- Suthers, D. D. (2006). Technology affordances for intersubjective meaning making: A research agenda for CSCL. *International Journal of Computer-Supported Collaborative Learning*, 1(3), 315–337. doi:10.1007/s11412-006-9660-y

- Tang, J. C. (1991). Findings from observational studies of collaborative work. *International Journal of Man-Machine Studies*, 34(2), 143–160. doi:10.1016/0020-7373(91)90039-A
- Tidwell, J. (2011). *Designing interfaces* (2nd ed.). Sebastopol, Calif: O'Reilly Media.
- Uniform Crime Reports. (2011). *Crime in the United States, 2011, Table 38 (online)*. Retrieved from <https://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2011/crime-in-the-u.s.-2011/tables/table-38>
- Van Dam, A. (1997). Post-WIMP user interfaces. *Communications of the ACM*, 40(2), 63–67. doi:10.1145/253671.253708
- Vatrapu, R. K., & Suthers, D. D. (2009). Technological intersubjectivity in computer supported intercultural collaboration. *Proceeding of the 2009 International Workshop on Intercultural Collaboration - IWIC '09*, 155–164. doi:10.1145/1499224.1499249
- Vetere, F., O'Hara, K., Paay, J., Ploderer, B., Harper, R., & Sellen, A. (2014). Social NUI : Social Perspectives In Natural User Interfaces. *Proceeding DIS Companion '14 Proceedings of the 2014 Companion Publication on Designing Interactive Systems*, 215–218. doi:10.1145/2598784.2598802
- Ware, C. (2013). *Information Visualization—Perception for Design* (3rd ed.). Waltham, MA: Morgan Kaufmann Publishers.
- Waszink, P. (1988). *Such Things Happen in the World': Deixis in Three Short Stories by N.V. Gogol*. Amsterdam: Rodopi.
- Wharton, C., Bradford, J., Jeffries, R., & Franzke, M. (1992). Applying cognitive walkthroughs to more complex user interfaces: Experiences, issues, and recommendations. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 381–388. doi:10.1145/142750.142864
- Wigdor, D., & Wixon, D. (2011). *Brave NUI World* (1st ed.). Burlington: Morgan Kaufmann.
- Wilson, J., Brown, J. M., & Biddle, R. (2014). ACH Walkthrough : A Distributed Multi-Device Tool for Collaborative Security Analysis. *In Proceedings of the 2014 ACM Workshop on Security Information Workers (SIW '14), ACM, New York, NY, USA*, 9-16.
- Wobbrock, J. O., Morris, M. R., & Wilson, A. D. (2009). User-defined gestures for surface computing. *Proceedings of the 27th International Conference on Human Factors in Computing Systems - CHI 09*, 1083. doi:10.1145/1518701.1518866
- Wu, M., Shen, C., Ryall, K., Forlines, C., & Balakrishnan, R. (2006). Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. *Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems, TABLETOP'06, 2006*(Figure 1), 185–192. doi:10.1109/TABLETOP.2006.19
- Wypych, T., Yamaoka, S., Ponto, K., & Kuester, F. (2011). System for inspection of large high-resolution radiography datasets. *IEEE Aerospace Conference Proceedings*. doi:10.1109/AERO.2011.5747534
- Yi, J., Kang, Y., Stasko, J., & Jacko, J. (2007). Toward a Deeper Understanding of the Role of Interaction in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6), 1224–1231. doi:10.1109/TVCG.2007.70515