

Wearable Tools for Affective Remote Collaboration

Kunal Gupta

University of Canterbury



Photo: Wearable Technology by Keoni Cabral

Wearable Tools for Affective Remote Collaboration

by

Kunal Gupta

Dissertation submitted in partial fulfillment for the degree of

Master of Human Interface Technology at the

Human Interface Technology Lab New Zealand,
College of Engineering,
University of Canterbury.

August 12, 2015

.

Supervisor: Dr. C. Bartneck, HIT Lab NZ
Co Supervisor: Prof. M. Billingham, University of South Australia
Dr. G. Lee, HIT Lab NZ

An electronic version of this thesis is available at

<http://library.canterbury.ac.nz/>.

Acknowledgements

My special gratitude to my Thesis supervisor Dr. Christoph Bartneck, for providing me an opportunity to do research in the lab under his guidance and expertise.

I would like to thank Prof Mark Billingham for sharing his valuable knowledge, experiences, expertise and accepting me as his disciple.

I would like to thank my co-supervisor Dr. Gun Lee for teaching experiment design, and its analysis and also for correcting if wrong by showing the examples.

I wish to thank Mr. Ken Beckman, for assisting throughout the thesis duration.

I am grateful to my parents, brother, family, friends, colleagues and staff at HITLabNZ for their constant support and motivation during this duration.

Wearable Tools
for
Affective Remote Collaboration

Abstract

Affective computing is the study and development of systems that can recognize human emotions and feelings. Emotions are always an interesting topic of research and these days researchers are trying to develop systems which can recognize, interpret and process emotions based on human physiological and neural changes for the development of well-being. As the market for wearable devices is expanding, it provides more opportunity of research in emotion sharing with remote person. This Master's thesis investigates the possibility of using wearable devices for affective remote collaboration. Previous research about affective computing, affective communication and remote collaboration using wearable devices is reviewed before starting the design process. Three wearable devices were developed, evaluated and discussed, two for emotion sharing between remote people, and the third for preliminary research to explore if eye gaze information can increase co-presence in remote collaboration. Conclusions and Future work are discussed based on the results from the research evaluation.

Declaration

I, Kunal Gupta, declare that the work presented in this thesis titled, 'Wearable Tools for Affective Remote Collaboration' is my own. I confirm that:

- This work was done wholly and mainly while in candidature for a research degree at the University.
- Where I have consulted the work that has already been published, is always clearly attributed.
- Where I have quoted from other's work, source is always given. With the exception of such quotations, this thesis is entirely my own work.

Kunal Gupta

Christchurch, August 12, 2015

Contents

- 1 Introduction 1**

- 2 Related Work 9**
 - 2.1 Affective Computing 9
 - 2.2 Emotion Representation 12
 - 2.2.1 Affect Visualization using Colors 13
 - 2.3 Affective Communication 15
 - 2.4 Affective Wearable Computing using haptic feedback 19
 - 2.5 Remote Collaboration using Wearable Systems . . 22
 - 2.6 Summary 26

- 3 CSense: Haptic band for Affective Communication 28**
 - 3.1 User Centered Design Process 29
 - 3.2 Initial Design 34
 - 3.2.1 Possible Design Solutions: User Centered Design Approach 34

3.3	Final Concept	43
3.3.1	Brainstorming	43
3.3.2	Implementation	44
3.3.3	Discussion	51
3.4	Summary	51
4	CoSense: Creating shared experiences	55
4.1	Research Opportunities	56
4.2	Initial Design.	57
4.3	Implementation	57
4.4	User Evaluation.	63
4.4.1	Experiment Design.	63
4.4.2	User Study	64
4.4.3	Results	64
4.5	Discussion and Future Work	65
4.5.1	Discussion	65
4.5.2	Problems Found.	65
4.5.3	Future Work	66
5	Gaze-Sense: An eye-tracking HMD for co-presence in remote collaboration	69
5.1	Research Questions	72

5.2	Design	72
5.2.1	Brainstorming	72
5.2.2	Stakeholders and Application Areas	74
5.3	Final Concept design	75
5.4	Implementation	77
5.4.1	Hardware	77
5.4.2	Software	81
5.4.3	Final Prototype	85
5.5	User Evaluation.	85
5.5.1	Evaluation Goal.	86
5.5.2	Experiment Design.	86
5.6	Results	95
5.6.1	Task Performance Time	95
5.6.2	Questionnaire: Quantitative Measure	96
5.6.3	Questionnaire: Qualitative Feedback.	135
5.7	Discussion	139
5.7.1	Implications of the Research.	143
5.7.2	Limitations	143
5.8	Conclusions	145
6	Conclusion and Future Work	147
6.1	Conclusion	147

6.2 Future Work	149
Bibliography	152
A Appendix A: Interview Questions	163
A.1 Emotion sharing in context of sports.	163
B Appendix B: Information sheet and consent form	165
C Appendix C: Poster for advertisement	169
D Appendix D: Ethics Approval	171
E Appendix E: Questionnaires	173
E.1 Pre-Task Questionnaire	173
E.2 Per-Task Questionnaire for Worker (Local User). .	176
E.3 Post-Task Questionnaire for Worker (Local User) .	179
E.4 Per-Task Questionnaire for Helper (Remote User). .	183
E.5 Post-Task Questionnaire for Helper (Remote User)	186

List of Figures

1.1	Wearable Devices	3
1.2	Sketch: Empathy. Sourced from [11]	3
1.3	Video Conferencing Modes	5
2.1	SHORE: Fraunhofer IIS face and object recognition. Sourced from [48]	11
2.2	Hernandez wearing SenseGlass system.Sourced from [47]	11
2.3	NeuroSky Mindwave. Sourced from [68]	12
2.4	Lumitouch. Sourced from [23]	16
2.5	AffectPhone. Sourced from [55]	17
2.6	inTouch. Sourced from [16]	18
2.7	The Hug. Sourced from [29]	18
2.8	Prototype by Salmiinen. Sourced from [44]	19
2.9	Participant using VIT prototype. Sourced from [10]	20
2.10	Tap Tap Prototype Sourced from [15]	21
2.11	Iam Feel. Sourced from [87]	21
2.12	Susan Fussell’s System	23

3.1	User Centered Design Process flow diagram. Sourced from [85]	30
3.2	Demographics of Users based on Relative Age, Gender, Type of Sports they enjoy.	30
3.3	Affinity Diagram from user interviews	32
3.4	Google Glass sketch, also used as SeeSense HMD. Source: [13]	35
3.5	Scenario 2	36
3.6	emoApp	37
3.7	emoApp:Sketch	38
3.8	emoApp's MIT App Inventor project screenshot	39
3.9	MIT App Inventor Block Editor Screenshot	40
3.10	Live broadcast of game DOTA2 by a player on the Twitch platform	41
3.11	Affinity Diagram	42
3.12	Model of emotion for CSense	45
3.13	Visual representation of idea (Designed by Charles Smart)	45
3.14	Circuit Diagram of CSense Hardware (Designed on Fritzing[37])	46
3.15	CSense Final Hardware Prototype	47
3.16	Brainstorming for the vibration patterns with respect to the emotional states)	48

3.17 Author using prototype controlling using Processing interface	50
3.18 Concept smartphone application prototype for CSense - I	53
3.19 Concept smartphone application prototype for CSense - II	54
4.1 Block Diagram of CoSense	58
4.2 Google Glass Screen interface with Spydroid	58
4.3 e-Health board and a user wearing sensors with e-Health board	59
4.4 CoSense Desktop Interface: Sketches	60
4.5 Raw Sensor data interface	61
4.6 Image graphic overlay	62
4.7 CoSense Desktop Interface	62
4.8 Experiment Conditions	68
5.1 Remote colleague sharing information using camera view sharing from a Google Glass display of the operating surgeon. Source from [3]	70
5.2 Block Diagram of the system	73
5.3 Philips Healthcare showing a proof of concept of using Google Glass (HMD) for medical surgery purposes.[46]	74
5.4 Gaze-Sense: Sketch	75

5.5 Remote Collaboration System: Sketch	76
5.6 Remote Collaboration System Interfaces: Wire- frame	76
5.7 Google Glass	78
5.8 Google Glass. Sourced from [8]	78
5.9 Google Glass. Sourced from [7]	79
5.10 Google Glass. Sourced from [5]	80
5.11 Eye Tracker for Google Glass with the help of [20]	80
5.12 Cameras for Eye Tracker	81
5.13 Eye Camera View using Pupil-Lab software[51] .	82
5.14 World Camera View on Display monitor using Pupil-Lab software[51]	83
5.15 Screenshot of multiple display system showing the working of Annotation Application	84
5.16 Author wearing the system and Eye Tracking Ap- plication running in the monitor behind	85
5.17 Example of task given to perform	88
5.18 Experiment setup layout	89
5.19 Countdown Timer to provide multiple focus points	90
5.20 Experiment setup: Participants performing the tasks	90
5.21 Participants performing practice task	93
5.22 Experiment in Progress	94

5.23 Task Time Performance: Interaction between the conditions	96
5.24 Connectedness: Interaction between the conditions (Worker)	100
5.25 Connectedness: Interaction between the conditions (Helper)	102
5.26 I was present: Interaction between the conditions (Worker)	103
5.27 I was present: Interaction between the conditions (Helper)	104
5.28 Partner sensed presence: Interaction between the conditions (Worker)	105
5.29 Partner sensed presence: Interaction between the conditions (Helper)	107
5.30 Partner knew when I needed assistance: Interaction between the conditions (Worker)	108
5.31 Partner needed assistance: Interaction between the conditions (Helper)	109
5.32 Enjoyed constructing the task: Interaction between the conditions (Worker)	110
5.33 Enjoyed constructing the task: Interaction between the conditions (Helper)	111
5.34 Focused on the task: Interaction between the conditions (Worker)	112
5.35 Worked together: Interaction between the conditions (Worker)	115

5.36 Worked together (Helper): A repeated measure two-way ANOVA followed by ART	117
5.37 Expressed Clearly: Interaction between the con- ditions (Worker)	118
5.38 Expressed Clearly: Interaction between the con- ditions (Helper)	119
5.39 Understood the partner: Interaction between the conditions (Worker)	120
5.40 Expressed Clearly: Interaction between the con- ditions (Helper)	122
5.41 Information from partner: Interaction between the conditions (Worker)	123
5.42 Information from partner: Interaction between the conditions (Helper)	124
5.43 Results of Ranking under different aspects . . .	125
5.44 Summary of qualitative result: Focus of atten- tion of partner	136
5.45 Summary of qualitative result: Task performance	137
5.46 Summary of qualitative result: Effect of Timer .	138
6.1 Idea concept image	150
B.1 Information sheet	166
B.2 Information sheet	167
B.3 Consent Form	168

List of Tables

2.1	Summary:Johann Wolfgang von Goethe’s Color theory. Sourced from [70]	14
2.2	Summary:Naz Kaya Color.Sourced from [70]	15
2.3	Summary:Claudia Cortes colorextraction. Sourced from [70]	15
2.4	Summary of research papers on remote collaboration with HMC or Eye Tracker.	25
3.1	Summary:Vibration Patterns	49
3.2	Summary:Vibration and tightening patterns w.r.t emotion	49
5.1	Four conditions with two independent variables.	91
5.2	Task Performance Time	96
5.3	Summary of the results of inferential statistics for the Likert Questionnaire	99
5.4	Mean connectedness (out of 7) for different conditions (Helper)	100
5.5	Mean connectedness (out of 7) for different conditions (Helper)	101

5.6 Mean presence (out of 7) for different conditions (Worker)	102
5.7 Mean presence (out of 7) for different conditions (Helper)	104
5.8 Mean partner sensed my presence (out of 7) for different conditions (Worker)	105
5.9 Mean presence sensed by partner (out of 7) for different conditions (Helper)	106
5.10 Mean: Partner knew when I needed assistance (out of 7) for different conditions (Worker)	107
5.11 Mean: Partner needed assistance (out of 7) for different conditions (Helper)	109
5.12 Mean: Enjoyment (out of 7) for different condi- tions (Worker)	110
5.13 Mean: Enjoyment (out of 7) for different condi- tions (Helper)	111
5.14 Mean: Focus (out of 7) for different conditions (Worker)	112
5.15 Mean: Focus (out of 7) for different conditions (Helper)	113
5.16 Mean: Task completion confidence (out of 7) for different conditions (Worker)	114
5.17 Mean: Task completion confidence (out of 7) for different conditions (Helper)	114
5.18 Mean: Worked together (out of 7) for different conditions (Worker)	115

5.19 Mean: Worked together confidence (out of 7) for different conditions (Helper)	116
5.20 Mean: Expressed clearly (out of 7) for different conditions (Worker)	118
5.21 Mean: Expressed clearly (out of 7) for different conditions (Helper)	119
5.22 Mean: Understood Partner (out of 7) for different conditions (Worker)	120
5.23 Mean: Expressed clearly (out of 7) for different conditions (Helper)	121
5.24 Mean: information from partner (out of 7) for different conditions (Worker)	122
5.25 Mean: information from partner (out of 7) for different conditions (Helper)	124
5.26 Ranking Questions	125
5.27 Mean: Task Enjoyment for different conditions (Worker)	126
5.28 Mean: Task Enjoyment for different conditions (Helper)	126
5.29 Mean: Felt connected with partner for different conditions (Worker)	127
5.30 Mean: Felt connected with partner for different conditions (Helper)	128
5.31 Mean: Stay focused on assembling the model for different conditions (Worker)	129

5.32 Mean: Stay focused on assembling the model for different conditions (Helper)	129
5.33 Mean: You were present with your partner for different conditions (Worker)	130
5.34 Mean: You were present with your partner for different conditions (Helper)	131
5.35 Mean: Needed assistance while performing tasks for different conditions (Worker)	132
5.36 Mean: Needed assistance while performing tasks for different conditions (Helper)	132
5.37 Mean: Understood Partner's communication for different conditions (Worker)	133
5.38 Mean: Understood Partner's communication for different conditions (Helper)	134

“see with the eyes of another, listen with the ears of another, and feel with the heart of another” For the time being, this seems to me an admissible definition of what we call social feeling

Alfred Adler

1

Introduction

The above quoted lines are from the famous psychotherapist, Alfred Adler, who is recognized for his work and theories emphasizing the importance of feelings of an individual. Emotions have a very important role in influencing everyday human activities that involve experience, social communication, learning and decision-making. The study and development of the systems that can recognize human emotions and feelings is known as “Affective Computing” [73].

Recently, research on empathic or affective computer interfaces has become popular. Consequentially there are a lot of multimodal systems that can recognize emotions with the help of facial expressions, speech analysis [19], and physiological sensor data [74] such as skin conductance, skin resistance and other galvanic skin responses, heart rate and blood oxygen level, and EEG sensor data. However, there has been very little research on how these emotional states can be

shared with another person along with the audio/visual cues to increase empathy. The goal of this thesis is to explore how technology can be used to allow people to better share their emotional state with each other.

Sharing an emotion with other person is very important for creating shared emotional experiences [77], and is known as affective communication. It can be done with the help of audio, visual, and haptic cues. The social sharing of emotion is elicited by emotional experiences, and the sharing of emotion is directly proportional with the intensity of the experience, irrespective of the type of emotion, whether it is a positive or negative [76].

Wearable systems are the body-borne electronic devices that can be worn by the user in the form of textiles, accessories, or any other form attached under, over or on top of clothing [62]. These are useful as they are attached to body and are capable of constant interaction with user, whereas normal computers, laptops, and smartphones cannot be carried or used all the time. Wearable devices can be used in various applications such as in the health care field, education, entertainment etc.

A simple and easy to use example of a wearable device is the Nike+, which is a small sensor which can be placed in the shoes, allowing a person to track her time, distance travelled, and calories burnt while walking or running. Another example of wearable devices is Google Glass [43], a head mounted computer with a small display, a camera and a touch pad for input, providing features of a normal smartphone like checking mail, texting, receiveing call, Google search, Google map etc.



Figure 1.1: Wearable Devices

Wearable devices can be used for Affective Computing, by using “on body” sensors (EEG, GSR, heart rate etc.) that can be used to collect and process physiological data from the user, and recognize different emotional states.

Empathy [28] has been defined:

”the psychological identification with or vicarious experiencing of the feelings, thoughts, or attitudes of another.”

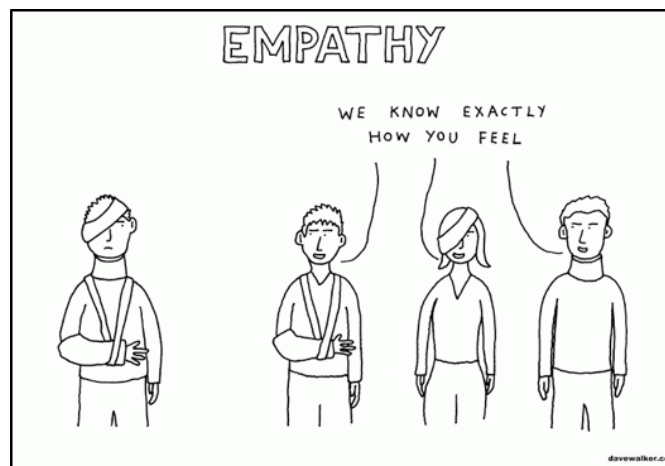


Figure 1.2: Sketch: Empathy. Sourced from [11]

From past research in the field of psychology it has been found that team processes, performance and team satisfaction can be influenced by emotion, and team member’s empathy [54]. The awareness between collaborators affects the

efficiency of collaboration and improve the relationship between participants [30]. For example Fabien [78] found that good performance could be achieved on the automatic emotion recognition if gaze behavior is provided.

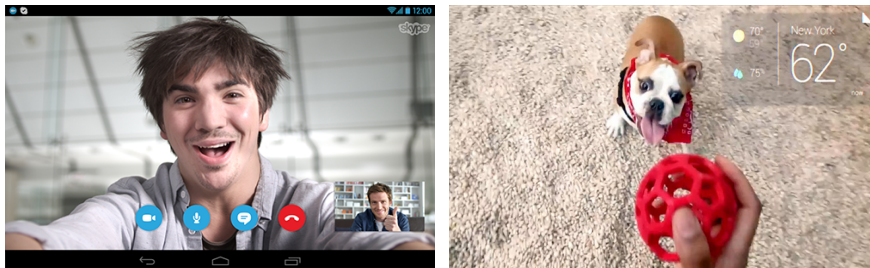
In this research we explore how wearable devices can be used for affective communication, and empathizing increasing presence in remote collaboration. The thesis is divided in three parts, one for each prototype developed:

- C-Sense: A haptic wristband for affective communication,
- Co-Sense: Creating shared experiences,
- Gaze-Sense: An eye-tracking HMD for co-presence in remote collaboration.

In first part, we present C-Sense (pronounced as See-Sense), a haptic wristband with a vibrator and servomotor attached to it in such a way that it will provide tightening and loosening effects to the user. Previous research has shown that haptic cues can be used to communicate emotional state on a desktop computer [86] or wearable system [22]. However, there has been little research on using wearable haptic devices for emotion sharing, and no one has tried to provide haptic, audio and video cues in wearable devices. In our research we developed a wearable device with haptic cues, along with showing visual and audio cues on a Head Mounted Display (Google Glass). In order to prototype C-Sense, we did some user requirement analysis in the context of sharing of emotions, then followed the Interaction Design process by interviewing potential users, creating affinity diagrams, personas and scenarios,

and performed ideation of possible solutions. Next we developed low fidelity prototypes of the ideas and conducted pilot tests, and finally we developed a high fidelity prototype.

In the second part we present our prototype “CoSense”. CoSense is a wearable system that shares a user’s first person view along with their current emotional state, in order to create a shared emotional experience with a remote user. Although there is an abundant availability of technology for remote communication, there is still a need of efficient affective communication. Current technologies, such as Skype or Google hang-out, communicate with visual and audio cues in a “talking head” conferencing mode [see figure 1.3a]. However wearable computers such as Google Glass have a head mounted camera that can share the user’s view of their environment, providing a “task space” view of their environment [see figure 1.3b].



(a) Skype: Talking Head Video, Sourced from [83] (b) Google Glass: Point of View, Sourced from [4]

Figure 1.3: Video Conferencing Modes

In the CoSense interface we used physiological sensors to capture what user is feeling, and the wearable camera/ microphone on Google Glass to record what they are seeing and hearing. Then we transmitted these feelings, sights and sounds to a remote user to create a shared emotional experience.

This chapter is broken into several parts. CoSense is an-

other solution of the same research problem explored the C-Sense prototype, creating shared emotional experiences, so the “Design Process” section of the CoSense chapter builds on the lessons learned from the design process of C-Sense. How ideas from design process section were transformed into a physical form is discussed in the “Implementation and Prototype” section. The finalized hardware and software of system components are also reviewed in this section. An overview of the experiment design, procedures and user evaluation of CoSense prototype is discussed in next section “Evaluation”. The quantitative and, qualitative analysis, results of the experiment are described in the “Results” section. Finally, we finish this chapter with an in-depth discussion about the results and the future work in “Discussions and Future Work”.

The final part of the thesis explores how eye-tracking could be used to increase the co-presence of remote collaborators. From Co-Sense, we found that the sharing of an emotional state with a remote person can be enhanced if we use a wearable HMD to share a first person view along with the emotions detected from wearable physiological sensors. In this chapter we explore what would happen if a person could share their focus of attention with their remote collaborator. Eye tracking is one of the best techniques for detecting a user’s focus of attention. So, if eye-tracking information is combined with the emotional state of the person, this could provide an ideal method for creating a shared attention and emotional state. So, we explore how head mounted eye tracking can be used to enhance remote collaboration between a local worker and remote expert.

This chapter will also be divided into sections similar to

the previous chapters. We will discuss the real world applications of using the above mentioned technology and conduct a needs analysis. Investigation, limitations and research opportunities explained in this section will lead to the “Implementation and Prototype” section, where the hardware and software technologies for the prototype will be reviewed. We will discuss the step-by-step approach of the prototyping phase including the various alternatives we thought of developing for this problem. After brainstorming about real life applications, we came up with an experiment design that will be discussed in the “Evaluation” section. We will also elaborate the task and the procedure of the user experiment. Once we are done with the experiment task, we will analyze the quantitative and qualitative data gathered from that along with the results and publish it in the “Results and Conclusion” section. Finally, in the last section of this part “Discussion”, we will summarize the whole part along with the discussion on the findings, limitations, and suggestions for improvements.

The last chapter of this thesis would be the final “Conclusion and Future Work” in which we will be talking about our complete research and future opportunities in the field of empathic computing.

To summarize this thesis, the problem statement of this research is “How can wearable devices be used to share emotional experiences between users and so create a deeper sense of empathy and understanding?”. To address this, we will be exploring different cues in various wearable devices to investigate the level of empathy between remote users in affective remote collaborations to provide rich and immersing human-human experience.

This thesis makes the following contributions:

- investigates the language of haptic feedback for affective communication in wearable devices.
- describes the user study with recognition and sharing of emotional states using a HMD to a remote person in order to augment empathy.
- describes one of the first user studies with head worn eye tracking with live annotations for remote collaboration.

2

Related Work

The research in this thesis is built on previous work in Affective Computing, Remote Collaboration and Wearable Computing. In this section we review key related research in each of these areas.

2.1. Affective Computing

Affective Computing, is computing of human emotions, its influences, developing systems to provide ability to recognize and express emotions while interacting with people [73]. It is related to human affects and emotional states presented according to the theory of emotions. Theories of Emotion mainly focus on human emotions and questions like what emotion is, how and why we feel a emotion and how these emotions are produced. For example, James-Lange [21] developed one of the first theories of emotion over 80 years ago, although this is

not widely accepted today. More recently Clynes [24] proposed a term called sentics, describing bodily components that carry emotions. Ekman researched how changes in the human motor system in response to different emotional states can produce different facial expressions [31], and there has been work on how emotions can cause speech modulation [69].

With the advancements of research and development in affective computing, researchers are trying to develop intelligent systems, which automatically recognize and respond to user's emotions. For example, Jocelyn Scheirer developed the expression-glasses [79] which sensed facial muscle movements and recognized expressions using a pattern recognition system. The integration of facial movements and speech recognition can be used to identify affect [27]. In a similar way, Busso [19] used visual markers on the face to track facial motions and combined this with speech recognition to recognize four user emotions; Sadness, anger, happiness and neutral.

An open source affect and emotion recognition engine named openEAR [34] uses acoustic data from the user to perform emotion recognition. Companies such as Emotient [32] and Fraunhofer [48] have also developed computer vision techniques for measuring emotion from face expression (See figure 2.1).

Picard's research group at the MIT Media Lab [58] demonstrated that physiological sensors and wearable computers could be used to recognize a wide range of emotional states. For example, their first wearable affective computing device was introduced in 2000 [45] and later they developed the Galvactivator [74], a glove that senses and communicates skin conductivity, which is an indicator of physiological arousal



Figure 2.1: SHORE: Fraunhofer IIS face and object recognition. Sourced from [48]

and valence emotional states. Recently they released the Q-sensor (2009) and a new clinical-quality wearable sensor called the E3 sensor (2014) [33] commercialized by Empatica. The E3 captures several types of physiological information such as PPG (heart rate and heart rate variability data), Electro dermal Activity (EDA), temperature and 3-axis accelerometer data. It measures physiological data from both the sympathetic and parasympathetic branches of autonomic nervous system.

Wearable computers can be used to provide emotional feedback to the wearer. For example, Javier Hernandez developed SenseGlass [47] that visualize the emotional states in a form of a meter (as can be seen in figure 2.2) shown on the Google Glass display. Emotional states were recognized by using the wearable Q-Sensor.



Figure 2.2: Hernandez wearing SenseGlass system. Sourced from [47]

Finally, research in Brain Computer Interaction has explored if brain activity can be associated with emotions [60]. For example, Neurosky's Mind Wave [68] is a portable wearable device that recognizes emotions while doing some activity using brain EEG Activity.



Figure 2.3: NeuroSky Mindwave. Sourced from [68]

As can be seen from this research, there are a number of ways that user emotion can be measured. However, most affective computing research has focused on systems that recognize an individual's emotional state, and not systems that help people understand the emotional state of someone else. In the next section we review research on how to develop systems that can convey affect between people.

2.2. Emotion Representation

In the above section, we discussed about the previous research done in the field of emotion recognition. On sharing the recognized emotional state, it is important to represent it in such a way that it can be detected by the remote person and enhances empathy. The parameters through which shared emotions can be visualized to the remote person will be discussed in this section.

2.2.1. Affect Visualization using Colors

One way to represent affective or emotional state is through the use of colour, however there is, there is no completely intuitive specific encoding of color that can be used to visualize for affect. There are also cultural, gender, age-group, situation and race components for color's emotional connotation. There is past work in which in which researchers tried to provide affect-color schemes. For example, in the color theory of Johann Wolfgang von Goethe [90], colors were categorized in positive and negative parts along with the emotions. Colors with positive parts such as yellow, orange (red-yellow), vermeil (yellow-red) were used for arousing, lively and ambitious states. Whereas colors with negative parts such as blue, blye-red and red-blue stand for restless, yielding and yearning. Instead of using an exact interpretation of each color, he described the colors with examples. Table 2.1 lists each color with its meaning [70].

Emotion	Positive Trait	Negative Trait	Color
happiness	passive		red-yellow
joy	pleasant	unpleasant	yellow
sadness	comfort	cold	blue
discomfort	active	restless	red-blue
powerful	energetic	irritating	yellow-red
same as red-blue, but more negative	more active	more restless	blue-red

Emotion	Positive Trait	Negative Trait	Color
faith	seriousness		red
calm	calm		green

Table 2.1: Summary:Johann Wolfgang von Goethe’s Color theory.

Sourced from [70]

Since it is very difficult to precisely identify an emotion based on the displayed color, Naz Kaya [65] represented frequency of the color for each emotion in the results that specifies how many users from a group of college students referred to a color for an emotion. Table 2.2 will shows the overall score on emotion-color combination (Using Munsell Color Space [64]) [70] .

Emotion	Color with Munsell notation [64]
happy	Yellow (7.5Y 9/10)
calm	Blue (10B 6/10)
anger	Red (5R 5/10)
Comfortable	Green (2.5G 5/10)
Tired	Purple (5P 5/10)
annoyed	Blue-Green (5BG 7/8)
loved/no emotion	Red-Purple (10RP 4/12)
Empty	White (n/9)
Disgust	Green-Yellow (2.5GY 8/10)
depressed	Black (n/1)
Excited	Yellow-Red (5YR 7/12)
Bored	Gray (n/5)

Emotion	Color with Munsell notation [64]
Powerful	Purple-Blue (7.5PB 5/12)

Table 2.2: Summary:Naz Kaya Color.Sourced from [70]

Same as Goethe, Claudia [26] also suggested the positive and negative traits of colors in terms of emotions. Extraction of meaning in a color is very difficult but she tried to create a model that gives an influenced positive and negative emotion meaning. A brief summary can be seen in the Table 2.3 [70] however more information can be on found her website [26].

Emotion	Positive Trait	Negative Trait	Color
Happiness	Lively	Cautious	Yellow
Joy	Ambition	Tiring	Orange
Sadness	Faithful	Depressed	Blue
Anger	Active	Offensive	Red
Introspective	Leadership	Arrogant	Purple
Faith	Calm	Greedy	Green

Table 2.3: Summary:Claudia Cortes colorextraction. Sourced from [70]

2.3. Affective Communication

Affective communication [36] is the interpersonal expression process of emotional states and feelings about things, themselves, and others. Sharing emotional state is an important part of day-to-day life. Humans interact with each other or

with their surrounding by using facial or speech cues to express their emotions. In general emotional communication can be categorized into two types, verbal and non-verbal. Use of voice, change in voice quality, rate, speaking style, change in tone, are verbal cues, while touch, distance, physical appearance, posture etc. are considered as a part of non-verbal communication.

Previous research has shown that emotions can be shared using various cues e.g. audio, visual, touch, change in temperature etc. For example, Chang [23] prototyped a system called “LumiTouch” which shared emotions using change in color (See figure 2.4). This system consists of two picture frames connected with each other using the Internet. If one person touched his or her picture frame, the other frame lit up, signifying that the first person is feeling some emotions for the second person. The intensity of light from the frame changes if the person is squeezing the frame indicating stronger feelings, and the other person can also respond to these emotions similarly.



Figure 2.4: Lumitouch. Sourced from [23]

The AffectPhone [50] is a mobile phone that detects user’s emotional state using GSR sensors on a phone, and transmits it to another phone using change in temperature to show

emotional state (See figure 2.5).

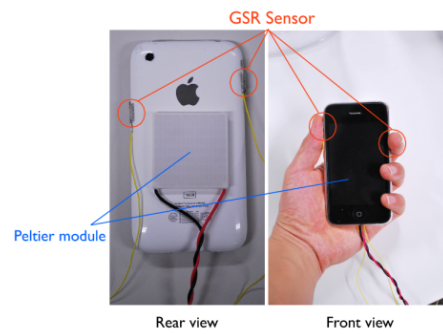


Figure 2.5: AffectPhone. Sourced from [55]

The sense of touch in human-human communication is an effective cue for expressing subtleties of emotional states especially when you have a rhythm in the pattern of that touch. For example, a person’s fingers tap on the table varies along with the shift in his emotional state [67]. There are a canny number of affective haptic devices, which are the systems that help in eliciting or representing the emotional state of a human using their sense of touch. Strong emotional experiences can be elicited with even a mere interpersonal touch [42] [35]

In 1997, Scott Brave [16] showcased “inTouch: A Medium for Haptic Interpersonal Communication (See figure 2.6)”. The main idea was to create a shared experience of touch at a remote distance. So one person will feel as if he is with other person touching the same physical object. He prototyped a mechanical system with rollers and connected it to another system with same dimension and rollers at the same location. Moving the rollers of one system would cause the same movement in the rollers of the other the system, which could be felt by the other user, creating an illusion of copresence.

“The Hug” [29] (See figure 2.7) is a haptic pillow that facilitates intimate communication over long distance. It is a

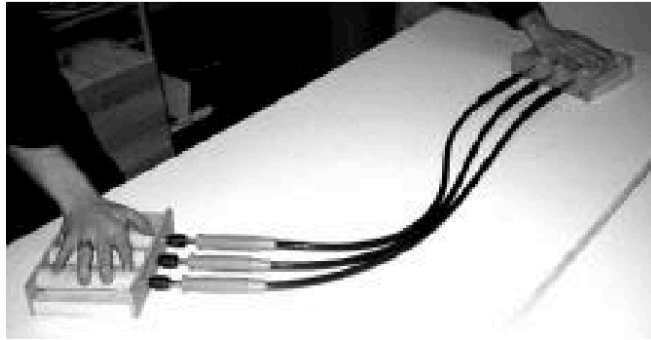


Figure 2.6: inTouch. Sourced from [16]

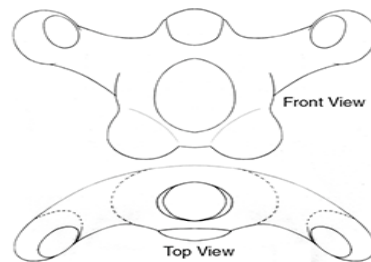
pillow shaped robotic device with sensors, and vibrator motors and sensors that can record pressing or rubbing of it's back, creates unique vibration patterns and sounds on the device at remote end. It has some pre-recorded messages like “goodbye”, “incoming hug” and “no one at home” which will be played with certain gestures.



(a) The Hug prototype



(b) Example use case



(c) Gestures and Sketches of the hug

Figure 2.7: The Hug. Sourced from [29]

Salmiinen [44] made a prototype to check whether emotional experiences could be perceived by a simple haptic stimulator. This was just one roller in the form of a rotating finger-

tip stimulator for 12 different stimuli with an average length of 500ms. He found that it is possible to share emotional information, which include arousal, dominance and pleasantness using it.

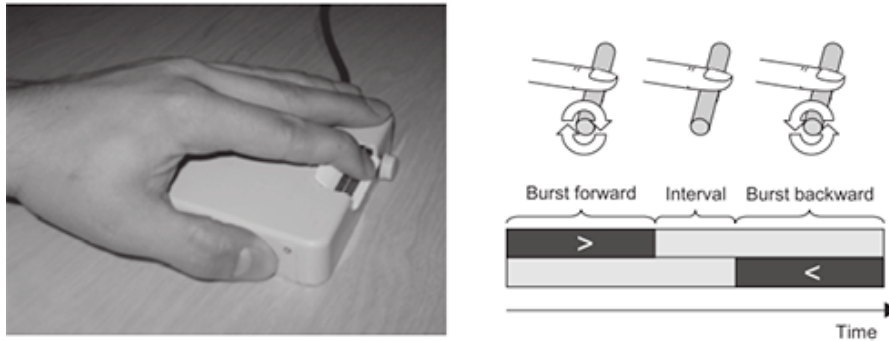


Figure 2.8: Prototype by Salmiinen. Sourced from [44]

The Virtual Interpersonal Touch (VIT) [10], is a force feedback haptic device for a collaborative virtual environment where a person can touch another remote person to increase the effectiveness of the collaboration (See figure 2.9). A 2DOF force feedback joystick was used to express seven emotions pre-decided by the researcher. Then participants were asked to perform a gesture with the joystick for each emotion. These gestures were recorded by the researcher to play at the remote end where the remote participant was asked to express the type of emotion they feel while touching the joystick. Results were positive for VIT but not as accurate as when people were expressing emotions through non-mediated handshakes.

2.4. Affective Wearable Computing using haptic feedback

The devices reviewed so far have mostly been handheld or connected to desktop computers. In recent years there has



Figure 2.9: Participant using VIT prototype. Sourced from [10]

also been research on how wearable computers can be used to provide haptic affective experiences. The main difference between wearable computers, compared to portable computers or smartphones, is that since they are constantly in close proximity to the body, they can provide intimate physical contact. Now since touch increases the trust, wearable devices with haptic feedback can provide a good option for affective communication.

For example, Leonardo Bonanni [15] introduced a haptic wearable modular scarf, tap tap (See figure 2.10) that can record, broadcast, and playback human touch. It had sensors, which asynchronously record the tactile information of the human body to convey person's affection can also play back that touch pattern to her lover, family member or a doctor.

Recently, Tsetserukou [87] developed a novel wearable humanoid robot that reinforces it's own feelings and simulates the emotions felt by the partner in a videotext online communication system. The algorithm recognizes nine emotional states from the text sent/received during the communication and simulates that emotion on the wearable haptic device at-



Figure 2.10: Tap Tap Prototype Sourced from [15]

tached at the various parts of the body in the form of vibrations that gives a sense of copresence. For example, if the system detects excitement it can generate the sound of a heart beating faster, or create a “butterflies in the stomach” effect to stimulate joy, etc.

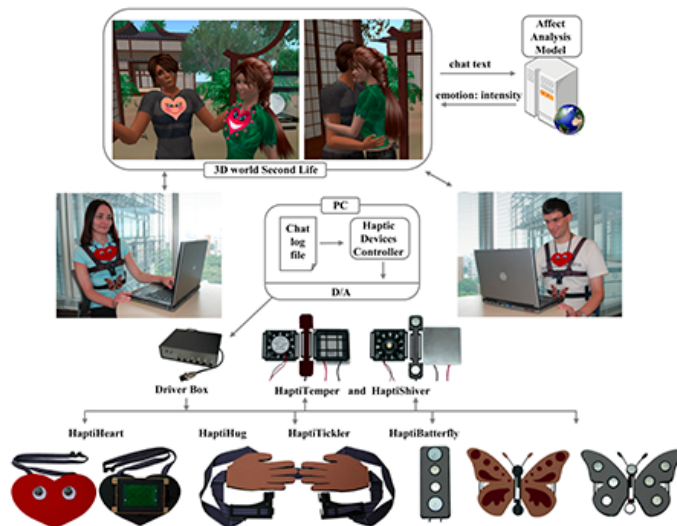


Figure 2.11: Iam Feel. Sourced from [87]

This research shows there have been a number of systems developed to convey affect remotely and that it is possible to use haptic, audio and visual cues to convey remote emotion. In our work we will build on this, but we are also interested in how to use technology to see what a remote person is seeing and so have a deeper emotional connection. In the next

section we review wearable systems that have been used for remote communication.

2.5. Remote Collaboration using Wearable Systems

As discussed in the introduction, one of the important elements of empathy is being able to see what someone is seeing. Video conferencing technology can be used to allow people to see each other, but software such as Skype is typically focused on providing a “Talking Head” experience that allows you to see the remote persons face, but not what they are seeing and doing. In recent years people have begun to explore how head mounted cameras and displays can be used to provide first person video from a remote person, and so give a “Task Space” video experience.

For example, Armstrong [3] used Google Glass to provide remote views of diabetic limb salvage surgery. The surgeon in the operating theatre shared a first person live view of the operation with remote surgical colleagues using the Google Hangout video sharing application in Google Glass. The remote colleagues were able to use real time diagrams and MRI images to provide expert assistance.

Susan Fussell [38] investigated the virtual physical co-presence between remote collaborators in a bicycle repair task (Physical Task) by creating shared a visual space using a head mounted camera worn by the worker sharing live video feed to helper.

These systems share a remote user’s view, but they don’t

show exactly where the person is looking in this view. Research on using eye gaze information for remote collaboration is a comparatively unexplored area of research. One of the few systems that explores this is the work of Fussell et al [40] [39][41]. They developed a system with a Head Mounted Camera (HMC), and an Eye Tracking Camera attached to a head mounted display (HMD) for local person (worker), which was sending a real time workspace video feed along with the user’s eye details to the monitor display of a remote helper (See figure 2.12). The remote helper was asked to assist worker by using the video from his monitor and providing verbal feedback. The worker was not wearing a head mounted display, so the remote helper was not able to provide visual cues.



(a) Worker wearing Head Mounted Camera and Eye Tracker, Sourced from [40]

(b) Helper’s Monitor view. Sourced from [40]

Figure 2.12: Susan Fussell’s System

Coordination of remotely situated users in a complex collaborative task using shared gaze was studied by Neider et al [66]. A pseudo-realistic city scene was developed in which one sniper target was popping up randomly and two remote users with eye gaze trackers had to locate and reach the target using a shared eye gaze and shared voice cues.

This research has shown that wearable computers and head-mounted displays and cameras can be used to share remote

views and help someone understand what a remote person is doing. However there has been little research on systems for first person video that include gaze cues, especially systems that have an eye-tracker, head mounted camera and head mounted display all integrated together.

There is little research available that has investigated eye gaze tracking in remote collaboration. The table below summarizes some of the relevant research papers. HMC = head mounted camera, HMD = head mounted display, ET = eye-tracker, R = Remote, FtF = Face to Face.

Paper	HMC	HMD	ET	R	FtF	Notes
[40]	X		X	X	X	Compared scene camera to HMC in remote collaboration to audio only and face to face conditions
[39]			X		X	Same setup as [40], Face to Face collaboration only - also include results from [40]
[66]			X	X		Both subjects had gaze tracking and looked at monitors for task. Compared shared video + speech only, to shared gaze only, to speech + gaze in visual search task
[72]	X		X	X		Used desktop display, HMC view not shared

Paper	HMC	HMD	ET	R	FtF	Notes
[71]	X		X	X		Modelling Focus of Attention using HMMs - same setup as [72]
[81]			X	X		Asynchronous collaboration, working on desktop screens
[41]			X		X	Face to Face collaboration on robot assembly task
[82]			X	X		Using eye tracking in remote desktop conferencing - desk mounted ET
[96]	X		X	X		One helper with ET collaborating with two remote workers. Note gaze information was not shared with remote workers - used for analysis
[17]			X	X		Remote users wearing ET and looking at desktop screens
[12]	X	X		X		Remote User was pointing at the Local person's scene on his monitor using mouse and it was shown on Local User's display.

Table 2.4: Summary of research papers on remote collaboration with HMC or Eye Tracker.

Previous research has shown that eye gaze is important for emotion recognition and empathy [28]. So, in our research, we will try to see effect of gaze behavior on the copresence between two people in a remote assistance collaboration condition. As can be seen from table 2.4 there is no example of an earlier system that combines a HMC, HMD, and eye tracking in a wearable system. .

2.6. Summary

In this literature review section, we discussed the research conducted in the field of affective computing, affective communication and remote collaboration. We have seen that there have been a number of technologies developed for recognizing emotion and some systems for conveying remote emotion. However, there has been little research conducted on using haptic cues in wearable devices for sharing emotional states. Thus our first prototype (CSense) will explore how haptic feedback in a wearable device can be used to increase emotional understanding and empathy between remote people.

We also showed how wearable Head Mount Display (HMD) devices such as Google Glass can have a first person view camera attached to the HMD, so a person can share what she is seeing and hearing. Use of head worn cameras have been shown to be effective for remote collaboration, however there has been little research on how emotional cues can also be provided on a first person view from a wearable computer. Thus our second prototype (CoSense) will explore how a wearable computer and head worn camera can be combined with emotional sensing to convey a person's emotional state.

Finally, we although researchers have used eye trackers attached to head mounted devices and cameras, we couldn't find any research on remote collaboration techniques using a complete wearable system with an eye tracker, camera and a head mounted display for worker to see the instructions from helper. Such a system would allow the remote user to understand what the local user was looking at and give visual feedback to help them in their task. The final prototype (gaze-Sense) is a wearable system that combines a camera, display and eye tracking. Overall the three prototypes developed explore areas in remote collaboration that haven't been extensively studied before, will an overall focus on improving emotional understanding.

3

CSense: Haptic band for Affective Communication

CSense is a system consisting of two components, a wristband with a vibration motor to simulate touch, and a servomotor with a small band to simulate the handgrip of another person. This idea was inspired from an incident that happened with the author. Imagine a situation where one person is stressed and wants to express her situation with her close friend who lives in a different part of the world. Now in this situation, even if they both talk, it is very hard for her friend to empathize. However if the person has a system that can simulate her emotional state in a haptic form, then this could augment the empathizing capacity between both friends.

Wearable computing devices are getting a lot of attention from the fitness, entertainment, education and medical domains due to their ability to be always on and close to a per-

son's body. We tried to use these same features for our research since our motive was to provide a personal experience for empathizing between two or more users. Since we are interested in using wearable technology for affective communication, we are mostly interested in various output modalities for sharing emotional states.

3.1. User Centered Design Process

We followed a User Centric Design Process in developing the CSense prototype (shown in figure 3.1). This process starts with a User Requirements Analysis in which the challenges faced by our target user in sharing emotions with remote person, and their expectations from the solution will be explored. This was then followed with a Design and Prototyping Phase where we came up with the possible solutions which meet the user's requirement and developed low and high-fidelity prototypes of these solutions. Finally we conducted pilot testing with users and came up with a final prototype that could be a possible solution. This final solution is described in the "Implementation" section of this chapter.

User Need Analysis

Our research task was to provide a wearable system for sharing emotional states with other people. As a part of the research, we chose people who play sport as our stakeholders for the requirements analysis. So any person who is interested in sharing his or her emotional states before, during or after playing or watching any sport, is our target user.

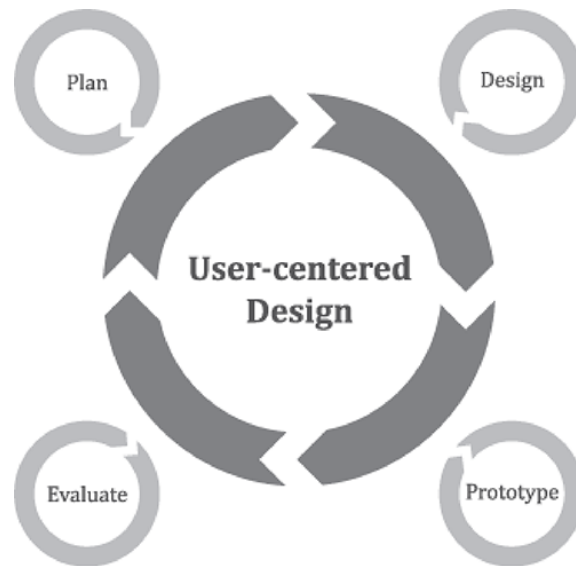


Figure 3.1: User Centered Design Process flow diagram. Sourced from [85]

Stakeholders

The main aim of this analysis was to explore the way people share their emotions, the extent to which they share, their experiences with sharing emotions, their comfort level of the emotional state they are sharing, and with whom they are sharing. Information was collect by having one-to-one interviews with users and observing people.

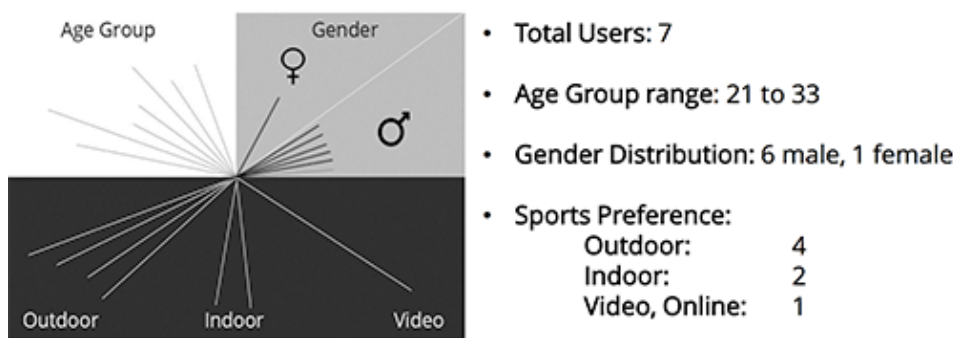


Figure 3.2: Demographics of Users based on Relative Age, Gender, Type of Sports they enjoy.

We conducted an interview session with 7 stakeholders (6 male, 1 female), ranging in age from 21 to 33 years (See fig-

ure 3.2, during which we asked questions about what kind of sport they played most often, what sort of emotions comes before, during and after playing, how do they share it and with whom, and their views on sharing emotions whether it is good or not necessary.

We observed their responses and found that they didn't want to share their emotions with everyone if they lost a game but will be excited to share if they win. However, they would want to share their defeat with a close person as it could give them relief. The people interviewed currently use voice calling, text, or social networking to share their emotions with remote friends.

Participant : "If you share, you will get better understanding of what is happening"

Participant : "I don't like to share too many feelings, so sharing to a particular level is good and with specific persons"

Participant : "You feel happy when you share the win of the game with your loved ones but you will not share when you loose because of the embarrassment."

To sort out the responses and prioritizing the basic requirements of the stakeholders, we used the K-J Method (Affinity Diagram) [52]. We created an affinity diagram (See figure 3.3) by sorting and categorizing the data as the positive or negative aspects of the emotions sharing on different colored sticky-notes. For example, one participant said his "Coach

shows a motivational movie to boost confidence of the team” which was playing basketball and before the game starts, so the statement was written on a green colored (outdoor sport) sticky-note and was grouped with other positive answers by the participants of things done before playing any sport. Similarly, another participant said that he gets nervous before the game as lots of people are watching which affects his performance. So, this was categorized in a separate class along with other negative aspects before playing sport.

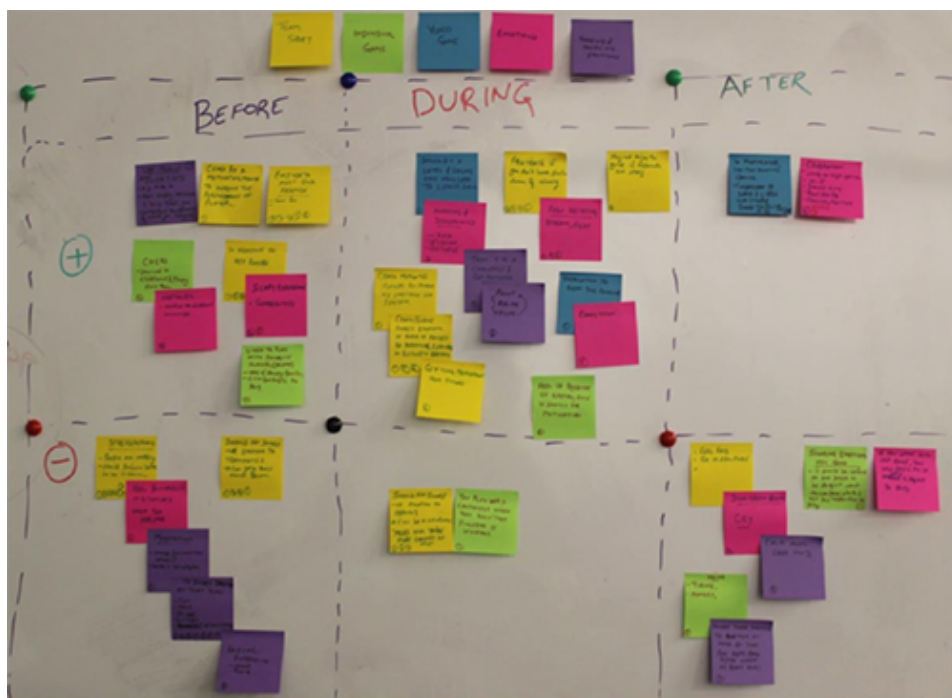


Figure 3.3: Affinity Diagram from user interviews

User Requirements

After completing stakeholder’s interviews and field observations, we developed the following set of non-functional and functional system requirements for the prototype. Our main functional requirement is that the system should be able to share the user’s emotions whereas the non-functional requirement is that the system can be placed close to the user so that

he or she can use it anytime just by touch or voice or gesture. Based on the functional requirement, it is necessary for the system to have a continuous flow of data, which is the emotional state of the user on a real time basis. Since the interview was focused on the context of playing or watching sports, the system will need to be able to be used when the user is running, talking and busy with playing. If he or she wins, they would be busy in celebration or if they lose, could be sad, and using the system alone.

Other Usability Requirements for the system are as follows:

- *Unobtrusive*: The system should be designed so that the user doesn't notice that it is being used. It has to be small and part of the daily routine.
- *Fun while Using*: The system should be easy to use which will encourage the user to use it most of the time.
- *Clear communication*: The interface should have accessible content that will not confuse the remote user with whom our main user is sharing their emotional state.
- *Two-way communication*: The system should be a two-way system where the user who is sharing emotional state with another user can get a response from them that will augment the empathizing process.
- *Sharing Interface*: The system interface should provide the option of single or multi person sharing, plus sharing only particular emotions etc.
- *Proximity and mobility*: The system should be in a close proximity of the user and should be portable so that user can carry it at all times.

- *Intuitive Interaction*: The system could be used by anyone so it should not be age group specific. Interaction should be natural, intuitive and easy to use.

3.2. Initial Design

In this section we discuss a few possible designs based on the user requirements. We rapid-prototyped the solutions with low-fidelity and high fidelity techniques and through pilot testing we came up with one best solution out of those few. We completed one further design iteration based on the user feedback, and developed a final prototypes which will be explained in the next section.

3.2.1. Possible Design Solutions: User Centered Design Approach

Idea 1 SeeSense: Real time sharing of emotions using display, sensors and T-Shirt.

The first idea explored was real time sharing of a user's emotion using a head mounted display, physiological sensors and a T-shirt. These components are described in more detail below:

Components:

SeeSense Glasses: A Head Mounted Display device with an integrated camera capturing the user's first person view of the environment.

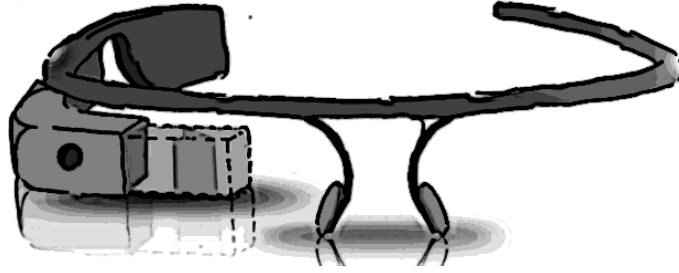


Figure 3.4: Google Glass sketch, also used as SeeSense HMD. Source: [13]

SeeSense Wear: A T-shirt having Sensors at various points that can sense temperature and also has heating elements that can imitate the temperature of a remote person that is using it. The shirt will have a heartbeat physiological sensor to compute the emotional state and a wi-fi module that will transmit all the data to the server along with the camera live feed. It will also have a contracting and expanding mechanism that simulates the emotional state with the help of a Soft Pneumatic Exoskeleton [93].

Usage Scenarios

Scenario 1: Person A will either be playing or sitting in the audience in a sporting event. She is using the SeeSense glasses to share a live first person view of the game that she is watching, which will show the area where she is looking at. She is also wearing the “SeeSense Wear” shirt that is sharing her body temperature, emotional state and heartbeat to the person with whom she is sharing.

On the other end, Person B is using the SeeSense System to experience a live view of the sporting event on her SeeSense glass and feeling a similar temperature via the shirt, which is also contracting or expanding as per the emotional state of

person A.

Scenario 2:



Bart Going to Play Soccer with his friends just for fun.



He meets his old friends over there, but his best friend was not there.



So he calls him and ask what is the problem?



His friend says that he is not feeling well so can't come over there. But he really wants to see the game.
⊗



So, Bart bought the SeeSense and started to share his view and feelings to his friend. And by this motivation that his friend is watching he won the game.



Figure 3.5: Scenario 2

Bart goes to play football in a match with his team but his best friend is not feeling well, so he cannot come to motivate and support him. However, by using the SeeSense System he shares his view and feelings with his friends that gives him support as he knows that his friend can feel what he is feeling and with this support they won the match.

Idea 2: emoApp

EmoApp is a proof of concept android application for sharing emotional states, recording messages and listening to the acknowledgements from other people.

Based on all the user requirements and user's feedback about how they currently share their emotions we came up with this android application idea. It is a basic application that provides the option to select the person with whom user wants to share the emotional state. This was because most users in the initial user requirement interviews were not comfortable with sharing their emotions with everyone. There is also a feature to record the message along with the emotional state that can be adjusted by the user by using slider bars on the screen. In the future we can automate the emotion setting feature by using a separate emotion recognizing system.

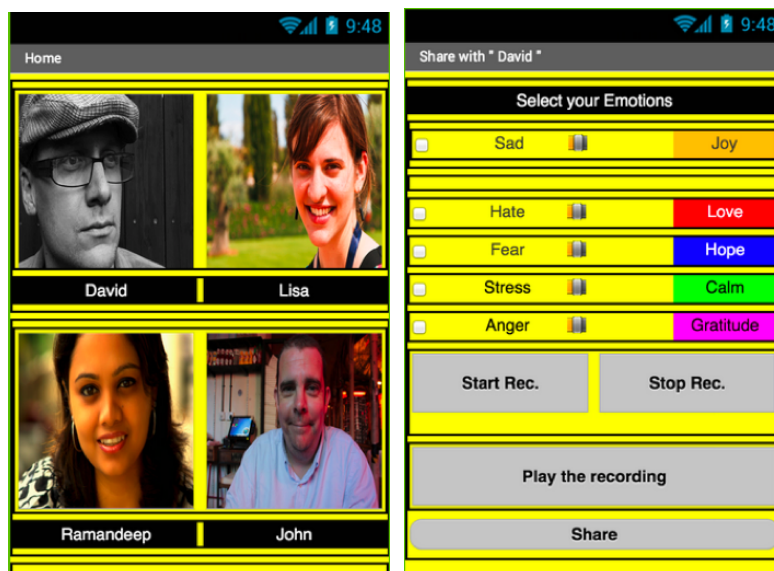


Figure 3.6: emoApp

On the Home Screen, you will get options of the people with whom you can select and share, and can add new friends

as well. Once you selected the person, a new screen will be opened for that person. You can then select the emotions by ticking the emotion desired and adjusting the slider to say how strongly you are feeling that emotion. You will get an option of recording a message for the person, e.g. if you are stressed out and you need someone to talk to you, you can just record a message and send it to the selected person.

The remote person can respond to the shared emotion with a pre-recorded voice message if they are busy or can reply back by sending their voice message. In this way, the user can be motivated or relaxed at that instant and the remote person don't have to be disturbed if they are in the middle of important work.

Prototyping

Low Fidelity Prototyping:

We started with sketching the basic layout of the *emoapp* application to check the necessary functionality (See figure 3.7).

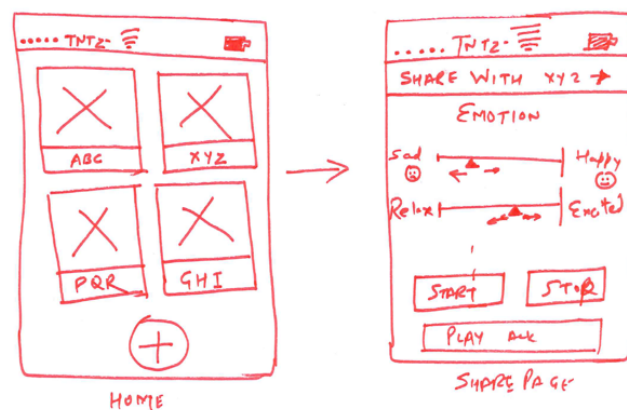


Figure 3.7: emoApp:Sketch

High Fidelity Prototyping:

After sketching, we developed a high fidelity prototype using the MIT App Inventor application [63], which is a tool for developing Android applications by simple drag and drop visual programming (See figures 3.8). It is very useful for non-developers, since it provides almost complete the functionality of a normal Android application and there is no coding required. It works using a drag-and-drop visual blocks graphical interface where each block is defined for a particular task.

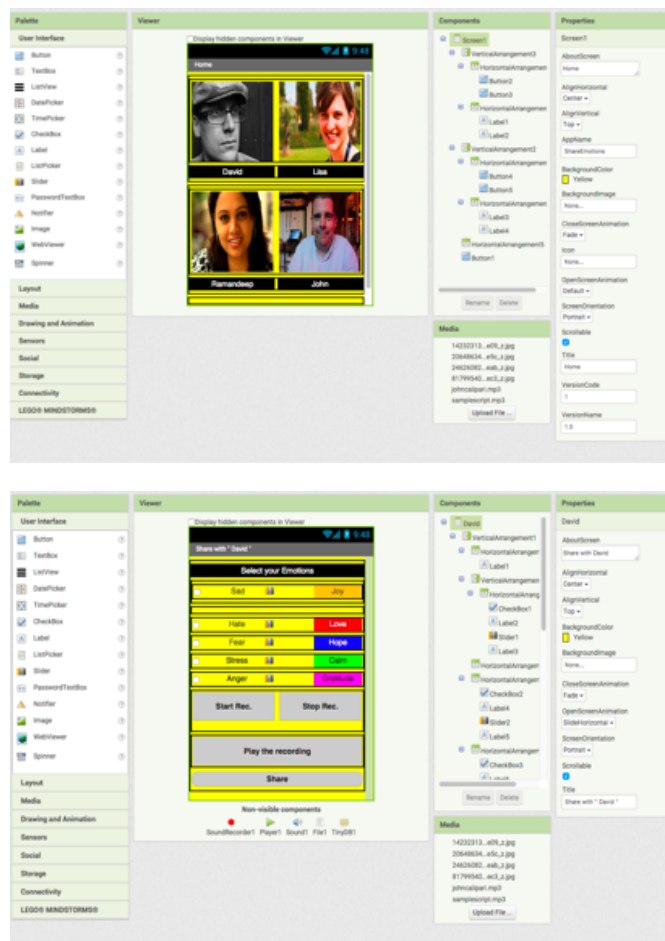


Figure 3.8: emoApp’s MIT App Inventor project screenshot

Idea 3: emoGlove

EmoGlove is a glove that shares a haptic experience while

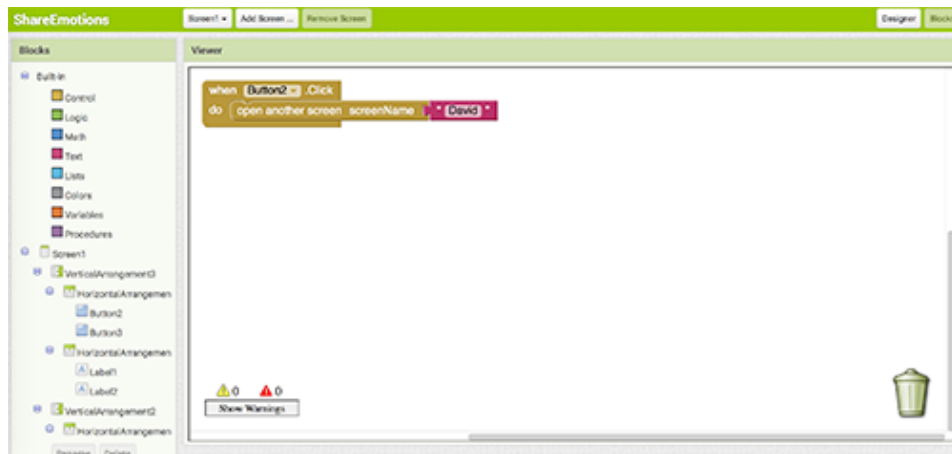


Figure 3.9: MIT App Inventor Block Editor Screenshot

playing online or computer games like with the PS3, Xbox etc. Let's imagine a scenario where a person is playing a video game either on Television or computer using a remote control system. Existing game console controls often have a vibration mechanism where the player will feel vibrations while playing. This generally vibrates when the player health in the game is being lost, which may result in a change in the player emotions. The idea of emoGlove is to share these emotional states with a remote person along with the first person player view of the player, with the hope that this may enhance the experience of remote person.

There are existing companies like "Twitch"[88], which provide a platform for players to share their gameplay remotely, so emoGlove could be an enhancement of this existing service. There will be sensors attached to the glove on various points where on holding remote controller the user usually gets vibrations and these sensors will record the pattern of the vibrations and the same vibration pattern will be replicated on the remote user's glove. So a live view using HMD and haptic feedback in gloves may help in enhancing the experience.



Figure 3.10: Live broadcast of game DOTA2 by a player on the Twitch platform

Pilot Testing

We conducted a small Thinking Aloud[61] pilot test in the form of user interview in which we focused on the feelings of the user after using the prototypes, whether they really want to share their emotions using the modes used in these prototypes, if they felt that the interface design was intuitive, and if they felt any discomfort. To use prototypes, we used a Wizard of OZ [53] and asked participants to tell their thoughts loudly while using the systems. At the end we also asked them about which idea that they liked the most and the idea that they didn't like, and the reasons for each choice.

Based on the responses from the users, we created an affinity diagram (See figure 3.11) to categorize the data about the useful aspects of the prototypes, the useless aspects of the prototypes, about not solving the problem completely, applications other than sports, and their reaction after using the prototypes.

From this interview, we understood that the stakeholders didn't fully like any single prototype, which was expected. According to them the first prototype (SeeSense) was best but

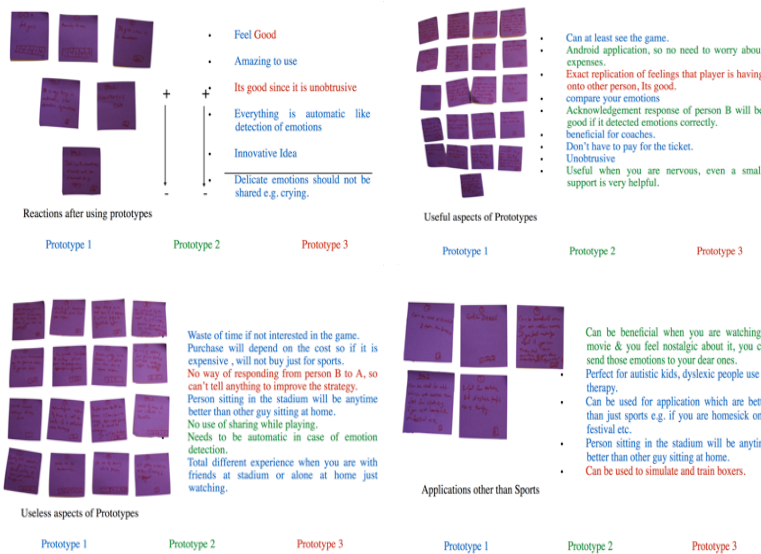
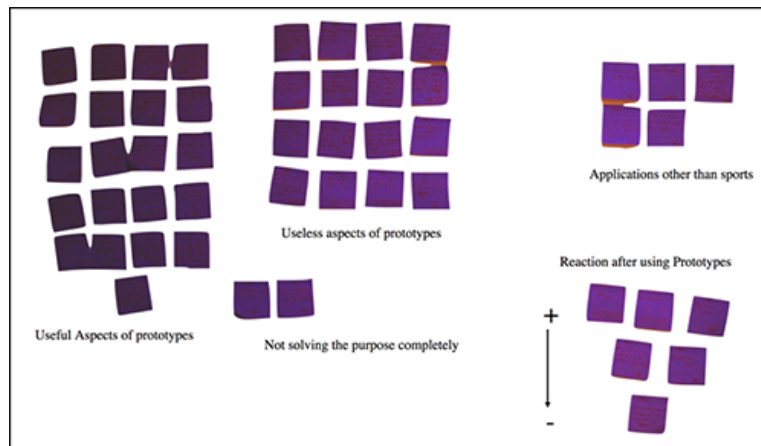


Figure 3.11: Affinity Diagram

could be improved by thinking of something which included a wearable device with haptic responses based on physiological data (emotional states). However they liked the idea of a smartphone application or any input mode to select the person with whom the user wants to share their emotional state, instead of sending it publically since the user might be a bit

cautious and might not feel comfortable in using the system.

3.3. Final Concept

3.3.1. Brainstorming

We brainstormed on the idea of sharing emotion based on the previous responses and tried to expand the scope of stakeholder from the person who watches or plays sports to a general person who wishes to share her emotions with their friends or publically. The author worked with his colleague Charles Smart on the brainstorming, ideation, prototyping and pilot testing of the final prototype.

Together, we redefined our objective of the research as “how the sense of touch could be used to communicate emotions remotely in a variety of contexts, using haptic technology”. Following brainstorming based on the previous study’s responses and user requirements, we came up with the following initial ideas:

- A T-shirt that could create the sense of an arm on the shoulder or a pat on the back,
- A shirt that constricts, to receive a virtual hug from someone,
- A glove that creates the feeling of holding hands,
- An air jet that can create a tickling sensations, for playful communication
- A wristband that sends/receives emotional state readings and represents these as vibration patterns

We decided to focus our project on:

A wristband that shares emotions through haptic sensations on the wrist to intuitively communicate emotional states and be controlled by a smartphone application.

This aligned with our previous study and the user requirements, according to which the system should be wearable, unobtrusive, support two-way communication, have a sharing input mode, and provide intuitive interaction.

3.3.2. Implementation

By keeping the above-mentioned idea in mind, we started implementing our prototype. We will discuss the hardware and software used for implementation separately.

Emotion Model Classification

The first tasks before starting prototyping was to finalize how many emotional states we were going to share using our prototype. We chose four quadrants of a simple two-dimensional emotional space: Angry in negative-active, Sad in negative-passive, Relaxed in positive-passive and Excited in positive-active states (as can be seen in image 3.12) [94].

Hardware

Our prototype design focuses on a wearable device with vibration and tightening and loosening capability. To provide these features we used the following components (See figure 3.13:

1. Wristband

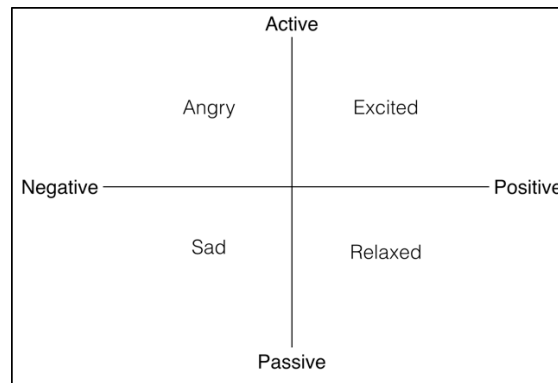


Figure 3.12: Model of emotion for CSense

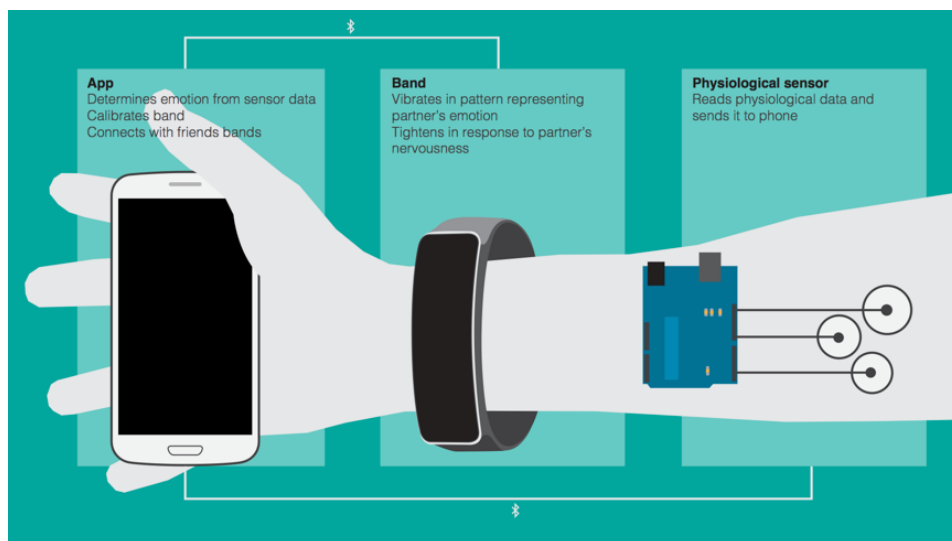


Figure 3.13: Visual representation of idea (Designed by Charles Smart)

2. Vibrator motor
3. Servomotor
4. Arduino Controller

We choose a wristband as an accessory to embed our prototype in since the wrist is one of the most easily accessible body parts that can be quickly brought in front of face. A wristband can also be in the form of sweatband which is very commonly used by sportsman, and so would be intuitive.

Vibration motors were used to generate haptic vibration tickling sensations. We used one 5V coin shaped vibrator mo-

tor, which was very easy to fit in the wristband. One Servomotor was used in an attachment with a thin band arranged over the wristband in such a way that they could provide a loosening and tightening mechanism.

Finally, we connected all the electronics components to an Arduino controller [2], which is a programmable microcontroller development board.

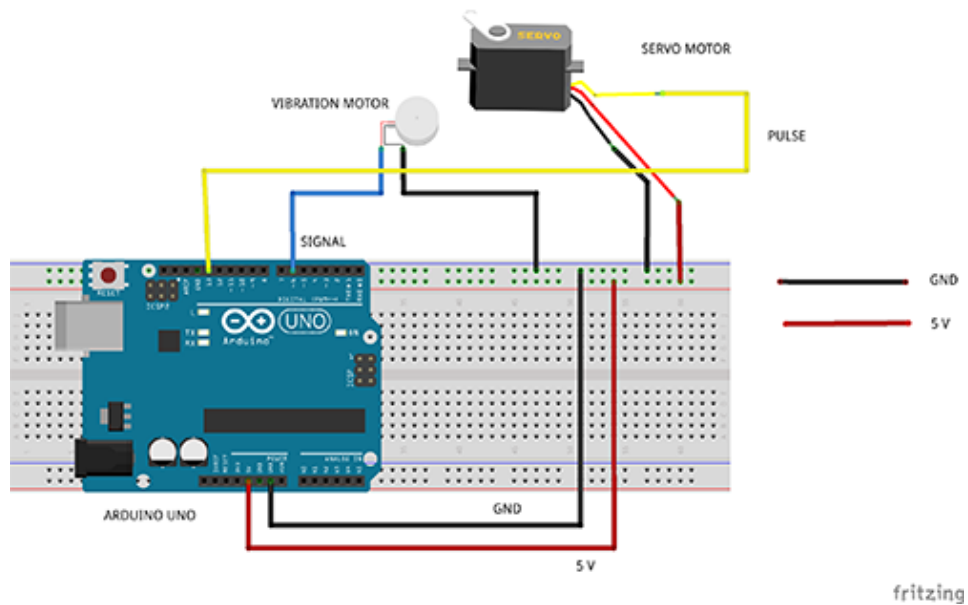


Figure 3.14: Circuit Diagram of CSense Hardware (Designed on Fritzing[37])

After arranging all the hardware components together, we finished the hardware part of our prototype. See figure 3.15 for a picture of the final hardware.

Software

Arduino provides its own library of functions and a developer environment to program the microcontroller. Initially, we developed code to control the vibrator motor and servomotor separately. To control the vibrator motor, we used PWM pin 6 of Arduino that is programmed to provide a range of analog

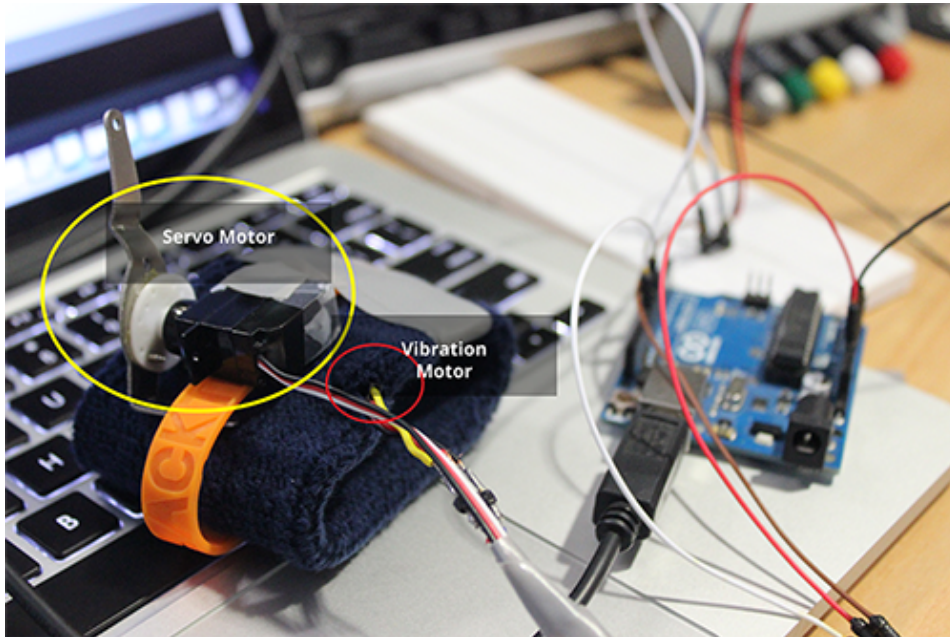


Figure 3.15: CSense Final Hardware Prototype

values from 0 to 1023 that will increase or decrease the intensity of vibration. However, to control the servomotor we used the “Servo” library provided by Arduino that has inbuilt functions to switch on, rotate anti-clockwise, rotate clockwise and switch off the motor, which was connected to pin 13.

Initial brainstorming was done regarding the pattern of vibrations that we would be using to represent emotional states. We tried vibrations matching the beats of music that was categorized into different emotions. However, pilot tests suggested that vibrations with the tempo of music were not able to represent emotional states properly. Next we tried patterns of heart rate as patterns of vibration motor that was comparatively recognizable in pilot tests, but this also didn’t work as most of the participants said that may be the wearer is running, not sure about angry or excited since they were relating the heart beat patterns with their own. Finally, since we had to create four different patterns for four basic emotions, we tried variations in the frequency and intensity of the vibration

motor along with the tightening and loosening of wristband with the help of the thin band attached to the servomotor.

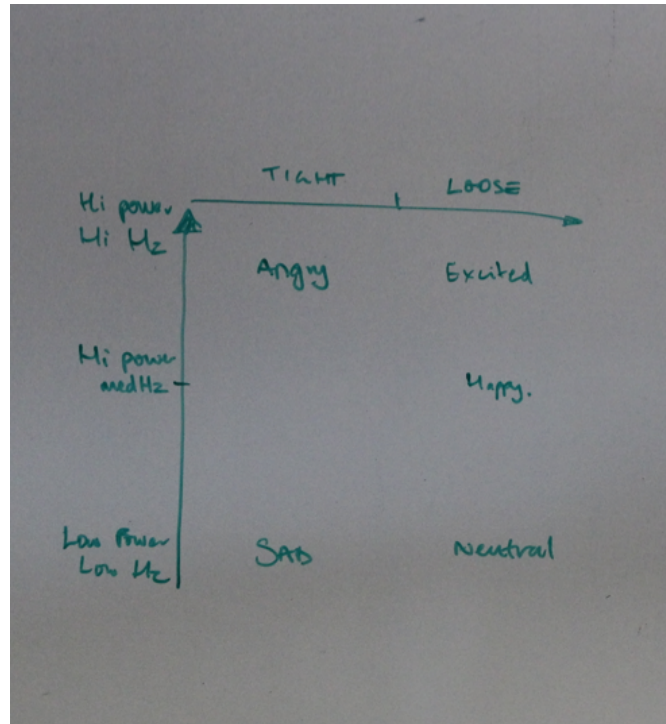


Figure 3.16: Brainstorming for the vibration patterns with respect to the emotional states)

Vibrations patterns were created with a combination of intensity of heartbeat and frequency of heartbeat where one heartbeat simulates the heartbeat of a normal human being i.e. 72 bpm with an intensity of 100 power out of 255 (analog voltage). These patterns are described in Table 3.1. We also categorized emotions with respect to vibration and tight/loose patterns, summarized in Table 3.2.

Pattern	Technical Description
High Frequency	110 bpm
Low Frequency	55 bpm
High Intensity	240 power
Low Intensity	50 power

Pattern	Technical Description
----------------	------------------------------

Table 3.1: Summary:Vibration Patterns

Emotion	Vibration Pattern	Tightening Pattern
Excited	High Intensity and Medium Heart Rate frequency	Loose
Angry	High Intensity and High Heart Rate frequency	Tight
Sad	Low Intensity and Low Heart rate frequency	Tight
Relax	Low Intensity and Medium Heart rate frequency	Loose

Table 3.2: Summary:Vibration and tightening patterns w.r.t emotion

We created an interface using Processing[75] to control the hardware devices. Processing is a rapid prototyping tool with GUI capability and an easy interfacing capability with Arduino.

For pilot testing, we asked a few colleagues to experience all the patterns at once, and after that we played all four patterns in random order and asked them to recognize the emotions based on the patterns and their own instincts. This test suggested that these patterns were better compared to the previous patterns.

For the smartphone application, we created a mockup and then interactive prototype of the app using the Invision[49] rapid prototyping tool.

- This application starts with a “Home” page (as can be

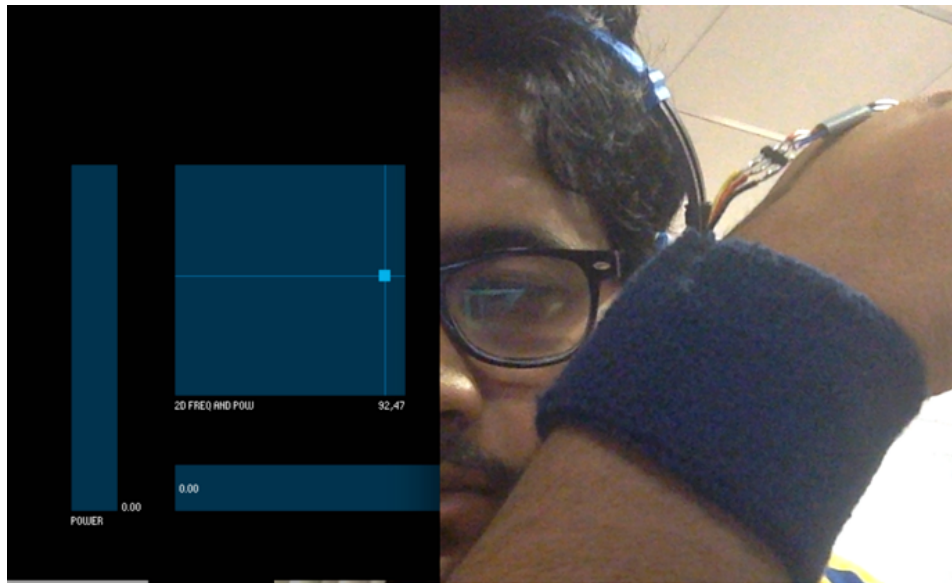


Figure 3.17: Author using prototype controlling using Processing interface

seen in image 3.18a) where an existing user can sign in or a new user can create an account (See figure 3.18b) or sign up using an existing Google or Facebook account.

- After signing in, the user will see a one-time Application Manual(See figure 3.18c) that shows all the features and flow of the application.
- In next screen (See figure 3.18d), user can see the available CSense devices and on tapping the device name, user can select the device .
- If CSense device is connected for the first time, next screen (See figure3.19a) will be to setup the device by placing the colored dots in the emotion's quadrant that will vary the vibration patterns and once done the user will select the vibration patterns for that emotion.
- After setting up the device, the user will see the emoZone screen (See figure 3.19b) where the user can categorize her friends as per her comfort zone whether that particu-

lar emotion can be shared with that particular friend. For example, the user can select her friend “Amily” and place her in the “Sad” and “Happy” zone that will allow the band to share sad or happy emotion with “Amily” from then.

- In the next screens, you can select a specific friend (See figure 3.19c) and share (See figure 3.19d) that particular emotional instance with that friend.

3.3.3. Discussion

In pilot testing with the prototype we found that users could easily distinguish between the four vibration patterns representing the four emotional states. They could easily differentiate between active (Angry, Excited) and passive (Sad, Neutral) emotions, and they found the mobile interface easy to use for setting the emotions. Finally, they reported that they found the heartbeat pulse the most compelling experience as they felt that the remote user’s pulse was almost their own. This interface shows that vibration and constriction haptic feedback could be an effective cue for conveying limited amounts of emotion.

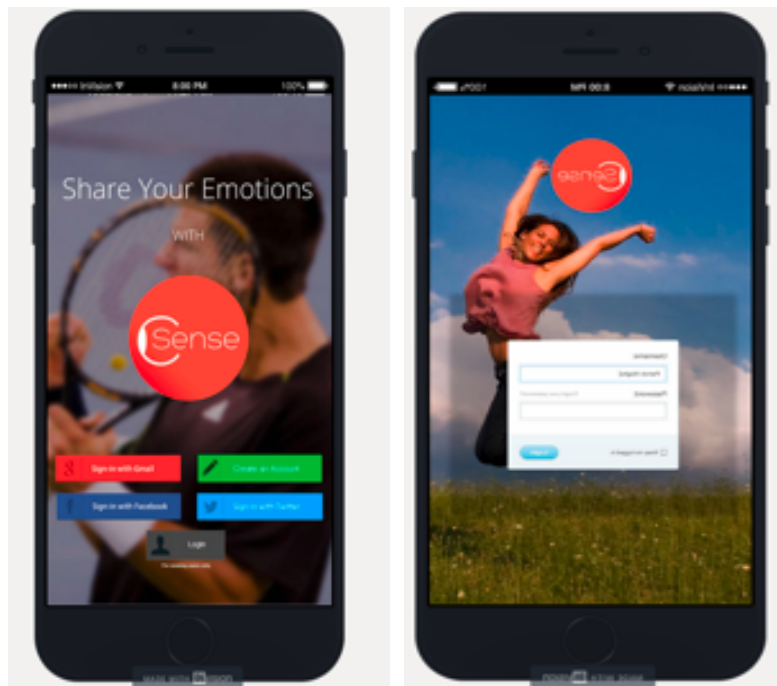
3.4. Summary

In this research, we mainly focused on representing emotional states in the form of haptic feedback. Our results suggested that this cue is efficient but still there are refinements like creating better patterns of vibrations to simulate states.

We created a wristband with a vibrator motor that vibrates and a servo motor that regulates tightening and loosening

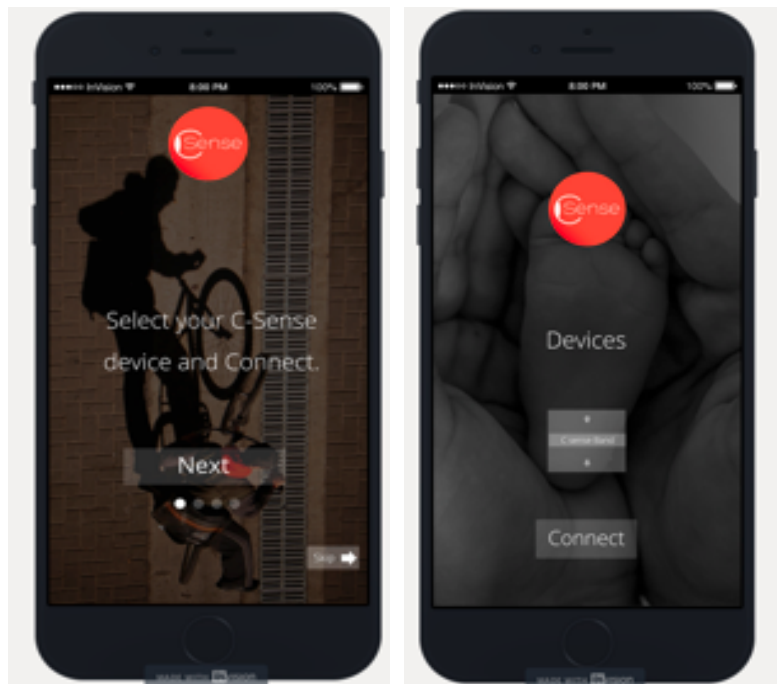
of the band with respect to emotional states. Through pilot tests, we argued that vibration patterns similar to the heart-beat pulse with varying intensity of vibration and frequency of pulse along with tightening and loosening of band helps in differentiating active and passive emotions.

In this chapter we have focused on a wearable device that doesn't provide many cues about what the user is seeing or hearing. However with head worn computers like Google Glass it is relatively easy to stream live audio and video to a remote user. In the next chapter we present a prototype of a wearable computer that allows a remote user to see and hear what a user is doing, as well as viewing cues about their emotional state.



(a)

(b)



(c)

(d)

Figure 3.18: Concept smartphone application prototype for CSense - I



(a)

(b)



(c)

(d)

Figure 3.19: Concept smartphone application prototype for CSense - II

4

CoSense: Creating shared experiences

Wearable devices such as Google Glass have cameras and microphones in them that enable video and audio to be streamed to a remote person. This allows the remote person to hear and see with the ears and eyes of the Google Glass user. However there has been relatively little research on using wearable technology like this to enable people to share feelings as well.

The goal of this research is to explore if sharing physiological sensor data in real time between people can be used to increase shared emotional experiences and create more empathy. This is part of a broader aim to develop wearable systems that will enable a user to share what they are seeing, hearing and feeling with another person. The problem statement is: “How can wearable devices be used to share emotional experiences between users and so create a deeper sense of empathy

and understanding?”

The author worked with Mr. Sudhanshu Ayyagari, an Electrical Engineering PhD candidate at University of Canterbury, who was developing EEG hardware and software. He contributed to the development of the prototype, the experimental design and the user testing.

4.1. Research Opportunities

In the CSense research, we mainly explored passive monitoring of emotions, where a person continuously monitors their emotional levels during their everyday activities and makes the data available to a close friend or family member. For example a daughter may check on her elderly mother’s heart rate from time to time to make sure that she is doing okay.

In the CoSense research, we will restrict our scope to active collaboration of emotional state, where a person is engaged in a short period of activity and wants to have a remote person share the experience with them. For example, going for a roller coaster ride for a few minutes. For this purpose, we wanted to use physiological sensors to capture what a user is feeling, and wearable cameras/microphones to record what they are seeing and hearing. Then we want to be able to transmit these feelings, sights and sounds to a remote user to create a shared emotional experience.

There are several key differences between this and the earlier related work:

1. Emotions will be automatically detected and shared from physiological sensor data, rather than being explicitly in-

put by the user,

2. Users will share video and audio from their surroundings, augmented by emotional signals in visual,
3. The focus is on helping one person have a shared emotional experience with another and so increasing understanding and empathy for their situation.

4.2. Initial Design

The main components for the “CoSense” system are:

1. A wearable computer such as Google Glass that will stream video and audio of the user,
2. A sensor system to compute emotions using physiological data.
3. A desktop interface for the remote user to view the images and emotional cues being sent from the wearable user.

We made a rough block diagram for the whole idea shown in figure [4.1](#)

4.3. Implementation

As per the user requirements discussed in section [3.1](#) and initial design concepts, the wearable user (Sender) used a Google Glass display running the Spydroid software [1] that supports video streaming to a remote desktop. The open source Spydroid software was modified to provide a Glass interface for showing the user’s emotional state, so the Sender could see

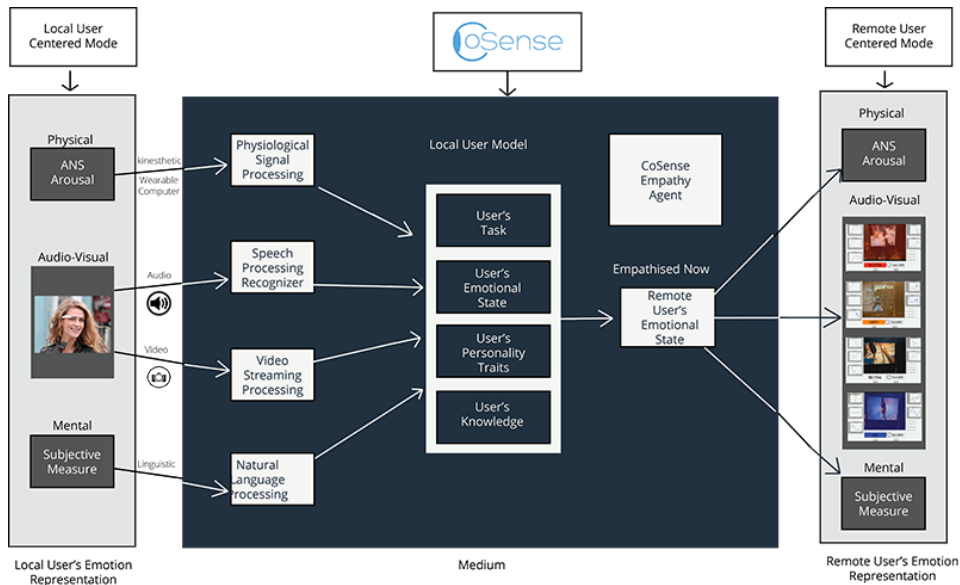


Figure 4.1: Block Diagram of CoSense

which emotions were being broadcast to the remote collaborator. Figure 4.2 shows the Google Glass interface screens. In order to switch between just video, video with emotional state and video with heart rate, we provided a swipe gesture on the Glass touchpad.

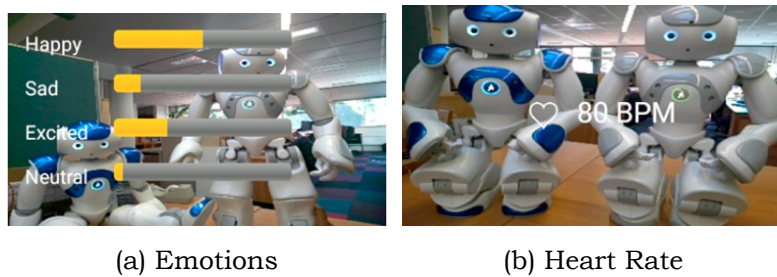


Figure 4.2: Google Glass Screen interface with Spydroid

We also developed several systems for detecting emotional cues based on previous research and available technology. The first was based on the e-Health hardware platform [25], which is an Arduino compatible hardware that has connections for a wide range of biosensors. We chose to use the GSR, pulse oxygen, and ECG sensors connected to the board, providing a wide range of physiological data. Figure 4.3 shows

a user wearing the sensors and the e-Health board collecting data.

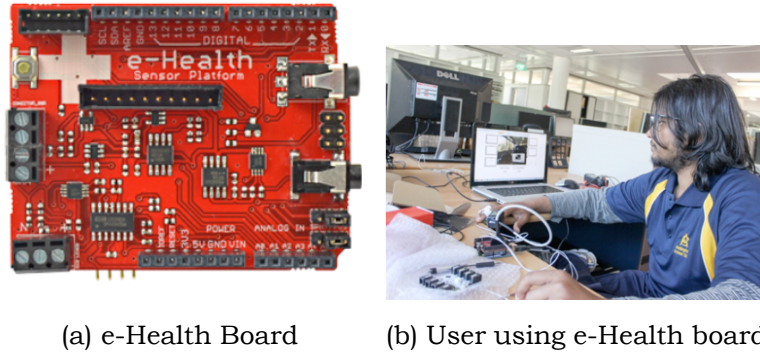


Figure 4.3: e-Health board and a user wearing sensors with e-Health board

We also explored the use of the Bitalino platform [14] and bio-sensors available from SEED studio [84]. The Bitalino had the advantage that it could be integrated into Android platform and so was very easy to include into the mobile interfaces, but neither of these systems provided the same level of performance as the e-Health system.

In addition to hardware we researched a variety of software libraries that could be used for emotion detection from raw sensor input. We finally decided to use the SSI framework [92] for some of our interfaces. This provides the ability to record and analyze human behavior in real time from variety of data sources, including the e-Health sensor, and others. In particular, we explored the use of the EmoVoice [89] component of the SSI Framework. This performs real time pitch tracking and audio processing of speech input to perform emotion recognition from acoustical properties.

In the final version of prototype, a python application was written on the PC to which e-Health board was connected, that could read in the raw data from the sensor data and then cal-

culate the user’s emotional state. In the application, we just use a simple threshold function that assigns emotion based on the signals coming from the GSR, Pulse Oxygen and ECG sensors. For example if all the three sensor readings are high then we assume that the user must be aroused and in an excited emotional state. In order for this to work reliably, the system must be calibrated for each user. When they are connected a user is asked to relax while baseline readings from the sensors are taken. These are then used to determine when the user is entering different emotional states. Just as with the CSense prototype, the CoSense system is designed to recognize the four emotions: Excited, Sad, Happy and Neutral.

For the PC application interface used by the remote person (the Receiver), we started with sketching an interface that would enhance empathizing and the level of understanding (See figure 4.4).

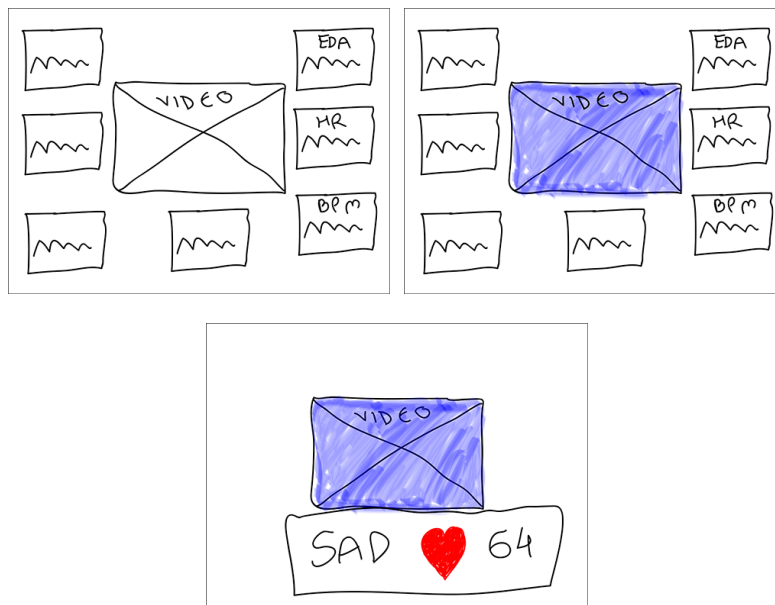


Figure 4.4: CoSense Desktop Interface: Sketches

One of the research challenges is how to represent the emotional cues in the interface. We explored three different ways

of showing the user's emotions;

1. Raw sensor data,
2. Emotion Labels,
3. Image graphic overlay.

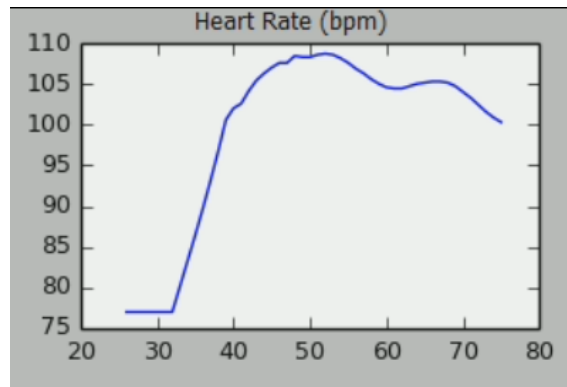


Figure 4.5: Raw Sensor data interface

Raw sensor data interface consists of graph plots from the data collected using the sensor (See figure 4.5). Of these different ways, the raw data is probably the most difficult for an untrained user to understand. In contrast, the emotion label (e.g. “Sad”), is shown by the user's heart rate and is very simple to understand.

The image graphic overlay is a transparent slide added on top of the main video, tinting the view of the user's environment. For example, the live video was tinted with an orange color if the emotional state is Happy that might change to red if the emotional state changes to Excited, Blue if the state is Sad and no change in color in case of Neutral state (See figure 4.6).

The Receiver was able to see and hear what the Sender was doing as the Google Glass video was streamed to her webpage

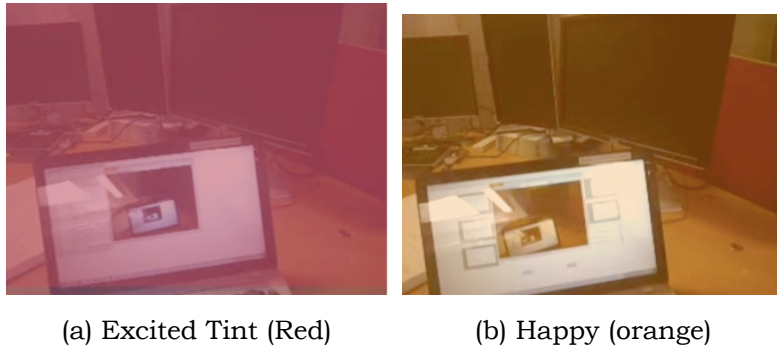


Figure 4.6: Image graphic overlay

using the Spydroid software. For the final prototype, sets of python scripts were written to display the sensor information (such as Sender’s current ECG data etc) around the Receiver’s screen. In addition, heart rate information and the users current emotional state as can be seen in the figure 4.7. We also provided an on-screen button to toggle all of the different interface elements.

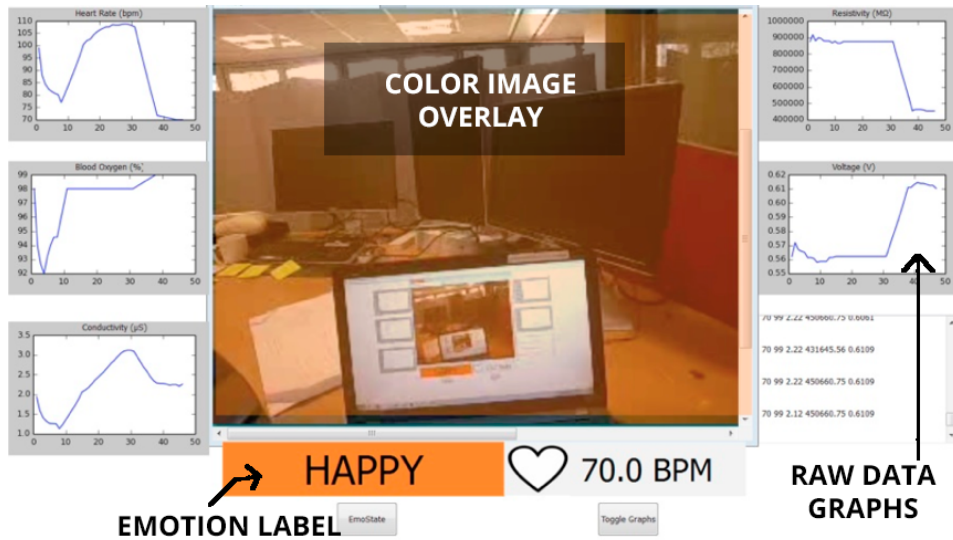


Figure 4.7: CoSense Desktop Interface

4.4. User Evaluation

4.4.1. Experiment Design

We conducted a user test, in which we compared the following four interface conditions for the Receiver:

1. No Cues: Just Video,
2. Emotion Cue: emotion label, heart rate and video,
3. Raw Graphs: video with the raw data graphs, and
4. All cues: Video, raw data graphs, heart rate and emotion tag.

This experiment required two participants out of which one participant was randomly assigned as Sender (Local Participant) and other was assigned as Receiver (Remote Participant).

Three tasks were designed in order to elicit the emotions in the Sender:

1. A guessing game with a Nao robot (Duration: 20 minutes),
2. Playing a first person shooting video game (The Evil Within) (Duration: 20 minutes), and
3. Watching a clip from the movie Butterfly Effect (Duration: 20 Minutes).

4.4.2. User Study

A within subject experiment study was done with 14 participants i.e. 7 pairs (9 Male and 5 Female within a range of age of 16-41, Mean age was 28). The effectiveness of each condition's interface for sharing of emotion was explored using subjective and objective measures after each trial. Both users were asked to complete a questionnaire before stating the experiment right after signing the consent form, then a questionnaire after every condition and at the end a questionnaire after all the tasks to rank the conditions.

4.4.3. Results

From the ranking questionnaire, we found that *C2(Emotion Cues)* was significantly better than *C1 (No Cues)*, *C3 (Raw Graph Cues)* and *C4 (All cues)* for *Q1 (How strongly do you feel the emotion ?)* and *Q2 (How well do you think you understood how your partner was feeling?)*. Whereas *C1 (No Cues)* was significantly better than *C2, C3, C4* for *Q3 (How easy was it to understand the interface?)*.

For a complete explanation of the results see the Work in Progress paper published at CHI 2015 [9] , Seoul, South Korea

4.5. Discussion and Future Work

4.5.1. Discussion

The goal of this experiment was to explore which combination of the interface cues would be better at conveying emotion to the remote user.

From the results, we observed that the system we developed created an awareness of the Sender's emotional state in the Receiver. These tests suggested that the Receiver could perceive a deeper understanding of the Sender's emotional state if they were provided some emotional representation in a visual form along with audio and video of the Sender's environment. Condition 2 (live video color tinted with the user's heart rate and an emotion state label), was felt to be more helpful by Receivers than the interfaces showing the raw sensor data (Condition 3) and even Condition 4 in which both Conditions 2 and 3 were mixed. Users said that this was because the interface provided simple visual cues without excessive information. The additional cues were useless for most of them as they were not able to interpret emotional state by looking at the graphs.

4.5.2. Problems Found

There were a number of problems encountered during the research that needed to be overcome. One of the first was the lack of experience with the hardware for emotion capture and limited availability of hardware. We needed a Signal Processing expert for recognizing emotions from a trained data-set using a robust hardware sensor system that can give filtered

and consistent data, but researchers working on this project were not that experienced with signal processing and due to lack of funding, we couldn't procure an efficient physiological sensor system.

Another problem was the difficulty of reliable emotion measurement and tracking of changes of emotion over time. It quickly became evident that different people produce different physiological signals for the same emotion and so it is very difficult to build a system that people can just put on and use. We partially solved this by using calibration and establishing a baseline sensor performance for each user, but this could be improved in future work.

Finally, visualizing emotional states is a difficult problem, since the color representation of emotion may vary for each individual. For example, one person can find orange for happiness whereas another person might feel it exciting.

4.5.3. Future Work

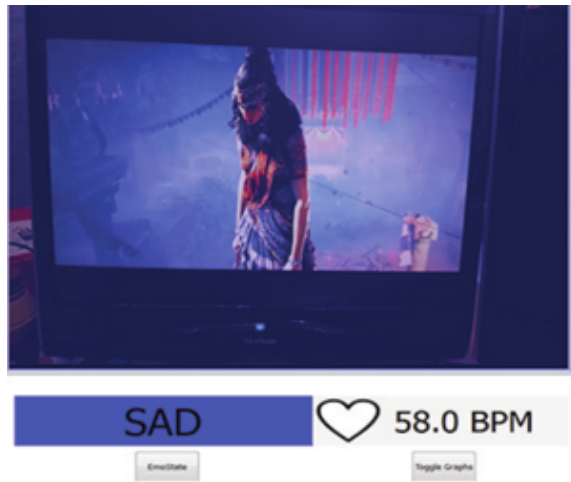
The current interfaces developed have used simple cues such as text, icons and graphics to represent emotion. However there are a wider range of emotion representations that can be explored, including the use of color, emoticons, icons and rich sound effects.

In this research, the physiological sensors we used, could be developed in form of a wearable system for example, physiological sensors attached on a glove or t-shirt etc. In future, we would like to explore the possibility of using these sensors as a wearable system.

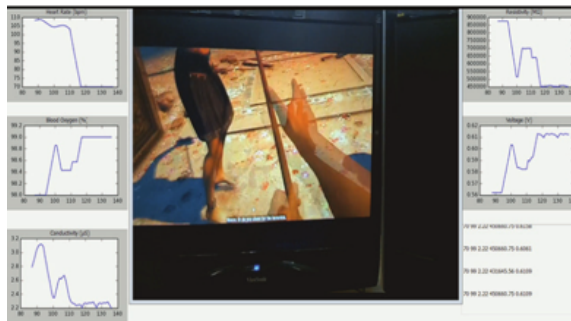
There was only one-way visual communication available in the current system, whereas empathy can be increased if we provide two-way visual communication. So, two-way visual communication technique will be explored in future.



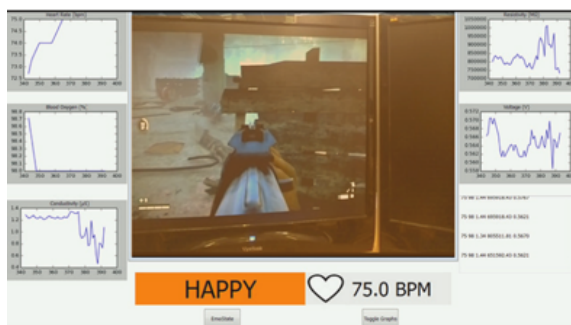
(a) C1: No Cues



(b) C2: Emotion Cues



(c) C3: Raw Graphs Cues



(d) C4: All Cues

Figure 4.8: Experiment Conditions
68

5

Gaze-Sense: An eye-tracking HMD for co-presence in remote collaboration

In the two prototypes described so far we were focusing on the sharing of emotional states with a remote person in order to augment empathy between both. Remote collaboration is an important application for this type of system; if we share emotional states of a Worker with a remote Helper, it might improve their level of collaboration as they feel more connected to each other.

There are many examples of how remote collaboration may help a person perform a real world task better. For an example, a surgeon is performing an operation but a specialist surgeon is not able to be present in the operating theatre, so technology could be used to allow the specialist to participate

remotely and instruct the operating room staff and surgeon in order to complete the surgery successfully. Specialist could see the operation table and staff on a video conferencing monitor through cameras mounted in the operation theatre, which might help him or her feel present in that room.

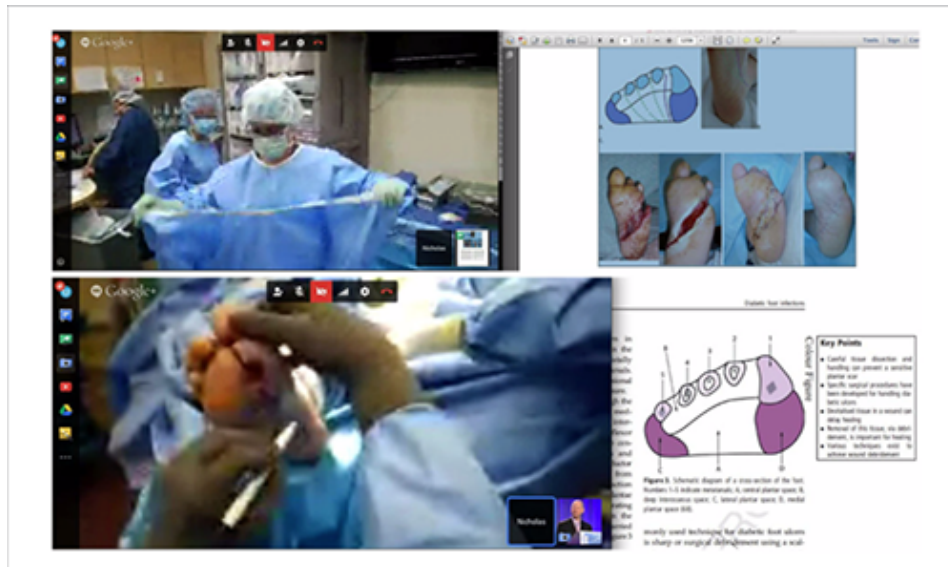


Figure 5.1: Remote colleague sharing information using camera view sharing from a Google Glass display of the operating surgeon. Source from [3]

However participating remotely is not the same as being there in person. In particular, it may be difficult for a remote person to know the exact focus of attention of the local Worker. Returning to our earlier example, if the specialist knows the exact focus of attention of the operating staff, this will simulate side by side collaboration in which a person can observe focus of attention of collocated persons. In this chapter we describe research that we conducted on using gaze tracking to convey focus cues between a local Worker and remote Helper.

Human Eye gaze can be used for an enormous number of applications such as sharing information, showing intimate expressions, maintaining social control and regulating inter-

actions. In a face-to-face conversation, eye gaze is an important factor in understanding each other properly as it provides information about where the person is directing his or her focus of attention. Langton [59] suggested that social attention can be achieved by mutual gaze and head movements. Jiazhi Ou et al [71] has shown that focus of attention can be predicted from intention (eye gaze) in remote collaboration tasks. So in this research we wanted to investigate the possibility of using an eye tracker on a HMD to mediate the focus of attention of the Worker, and measure the impact on connectedness and co-presence [57] while sharing the real task space view with a remote person. There has been some research on using eye-tracking for “talking head” based video conferencing, but little or no research on using eye-tracking in “task-space” video conferencing applications.

In order to do this research, we started by reviewing previous research in the field of remote collaboration using HMD’s or Eye tracking (Table 2.4 on page 25) . Based on those findings, we followed the design process to provide a basic layout of the system. After that we developed a prototype system consisting of an Eye tracker attached to Head Mount Display with a head mount camera fixed on it in such a way that it will share the First person view. To evaluate the prototype, we designed and conducted a user experiment with a remote collaborate setup. Finally, we conclude our research by discussing the results and directions for future work.

5.1. Research Questions

In this research, we will be exploring the following research questions:

- Can sharing of the Focus of Attention (FoA) of a Worker to a Helper, using eye-tracking, make an impact on the connectedness and co-presence of the remote collaboration?
- Can virtual co-presence be increased by combining the Worker's Focus of Attention and the Helper's Annotation in the shared visual space?
- Can sharing of Worker's Focus of Attention and the Helper's annotations increase task performance?

5.2. Design

5.2.1. Brainstorming

The system that we were developing combined the following key elements to support collaboration between a local user and remote expert (can be seen in image [5.2](#) on page [73](#));

1. a head mounted eye-tracker,
2. head mounted camera,
3. head mounted display, and
4. remote viewing software that allows a remote expert to annotate on the local user's view.

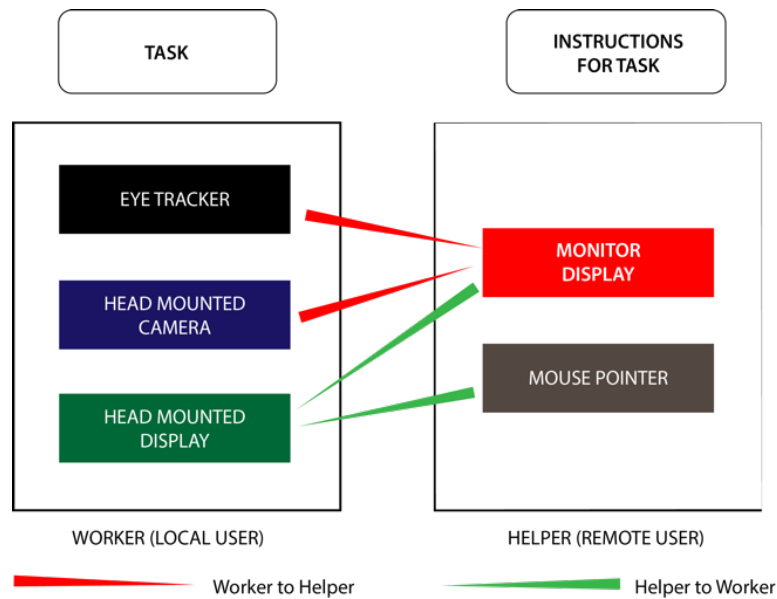


Figure 5.2: Block Diagram of the system

We brainstormed with colleagues who had experience of research in remote collaboration and people who had experience of using video conferencing about the features that such a system should have. Based on their suggestions and our observations in remote collaboration situations, we found that the main features of the system at Worker's should be:

1. Hands-free,
2. Easy to use,
3. Intuitive,
4. Display in visible range,
5. Reliable Cues for focus of attention, and
6. Robust.

For this research we were focusing on physical tasks for remote collaboration and we assumed that the Helper will have a physical instruction manual to help the local user. To design the user interface for Helper, we kept these things in mind:

1. It should be simple,
2. The interface should be clean and minimal,
3. It should convey information to the user properly, and
4. It should be intuitive to use.

5.2.2. Stakeholders and Application Areas

For this research we are targeting remote assistance based applications in which one expert will assist a naïve Worker. There are enormous number of potential applications for this kind of system, e.g.

- *Education:* Online learning where it is important to know the attention of the student.
- *Medicine:* We have already mentioned the use of this system in surgery to enhance collaboration between operating staff and remote expert surgeon assisting them.



Figure 5.3: Philips Healthcare showing a proof of concept of using Google Glass (HMD) for medical surgery purposes.[46]

- *Military:* Using HMDs to provide maximum information from a command and control room during missions.

- *Industry:* A Worker repairs complex systems on a remote location could use this kind of system in which remote Helper could assist him.

5.3. Final Concept design

We started with sketching the layout of the system for the Worker, since we already knew that we will need one eye tracker to track the eye pupil, a camera mounted facing towards the world that will capture the task space, and one small HMD that will be used to show virtual cues from the Helper super-imposed over the task space video.



Figure 5.4: Gaze-Sense: Sketch

The Worker and Helper both had slightly different interfaces since the Worker was wearing a HMD while the Helper looked at a desktop display. For the overall interaction we chose a simple video conferencing model in which both people (Helper and Worker) were able to see the same workspace using the head mounted camera and were able to share audio + visual cues. For the Helper's interface, the information that we showed was the video shared from Worker via the Head Mounted Camera (HMC) with the eye gaze information shown

on top of it. To provide the annotated input from Helper to Worker, a mouse click pointer seemed to be an easy and natural way of interaction as almost all the people who use computers are familiar with the mouse and its functionalities. Figure 5.5 shows a sketch of the Worker and Helper interfaces.

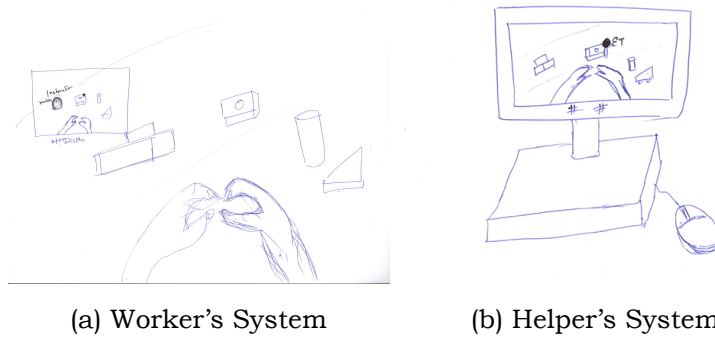


Figure 5.5: Remote Collaboration System: Sketch

We created wireframes for the interfaces to be shown at Worker's display and Helper's display (see figure 5.6). The Helper's interface shows a marker for the eye gaze information on top of the video of the task space from Worker's HMC, and the same video will be displayed on the Worker's HMD with an instruction pointer added from the Helper's side.

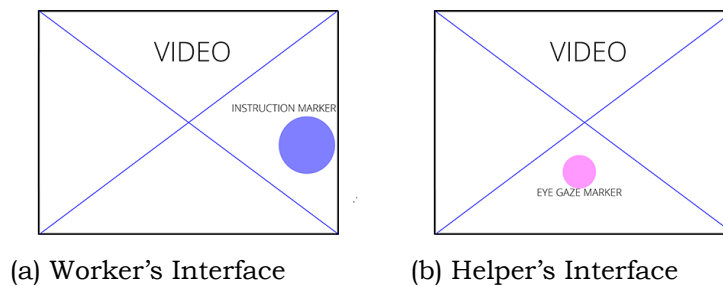


Figure 5.6: Remote Collaboration System Interfaces: Wireframe

5.4. Implementation

From the brainstorming and final concepts, we developed a working prototype of the system for user evaluation. This section will describe the hardware and software part of the system in detail. The experiment design and user evaluation will be discussed in the next section.

5.4.1. Hardware

For our prototype we needed a HMD with eye tracker capability and a HMC that could share the view to a remote user. The functionality needed from the prototype was:

- Head Mounted Camera (HMC),
- Eye Tracker (ET),
- Head Mounted Display (HMD), and
- Sharing to remote user

To fulfill all these requirements, several different display devices were evaluated to see how suitable they were, including Google Glass, the Vuzix Wrap 1200VR, and the Brother AirScouter. In this section we discuss each of these in turn.

Google Glass

Google Glass an optical see through HMD and integrated wearable computer with the Android operating system on it. For input, it uses touch gestures on a touchpad that is located on the side of the glass or voice commands like “OK GLASS”.



Figure 5.7: Google Glass

There is a 5-megapixel camera, with a capability of recording 720p HD video that is facing outwards. The display is an LED illuminated, 640 X 360 Himax HX7309 LCoS (Liquid Crystal on Silicon), and field-sequential color system. It provides a good display quality due to the smaller size and precise distance from the eye that makes the display quality sharp. We tried to use to the video camera for live streaming using Spydroid[1] but due to the extreme video time lag, it turned out to be a terrible experience. For this reason we had to explore other HMD options.



Figure 5.8: Google Glass. Sourced from [8]

Vuzix Wrap 1200 VR

Next we tried the Vuzix Wrap 1200 VR [91] which has a high resolution 1280 X 720 pixels LCD display in a stereo video arrangement, and can be attached to a computer using standard VGA and USB cables. This display is not an optical see-through headset, however by attaching a camera to it, we can use this device as a video see-through display. In this case

we can show video from the camera on the screens allowing the user to see a real time view of the world. However, the main disadvantage of using the 1200VR was that there was no space to attach the eye tracker on the HMD. So, we could not use this device either.



Figure 5.9: Google Glass. Sourced from [7]

Brother AirScouter

The Brother AirScouter [18] is a high quality optical see-through monocular display. It is similar to Google Glass except with a higher 800 by 600 pixel resolution. The monocular display is actually a small projector that can be attached to either side of the glasses based on the eye-dominance of the user. It also has a front-back, right-left and up-down adjustable feature that can be connected to the computer via USB port. It provides a 22.4 degree Field of View, is lightweight, and easy and comfortable to wear.

The only disadvantage of this system was the compatibility with graphics card and operating system while installing. It uses its special drivers that can only be used with NVidia graphics cards and on Windows XP or Windows 7. We found the solution to this problem by using 64-bit, Windows 7 operating system, 8.00 GB RAM, 3.60 GHz processor, NVIDIA

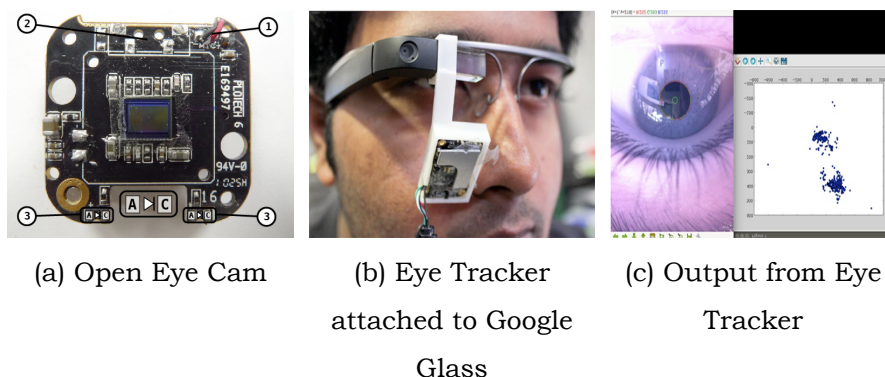


Figure 5.10: Google Glass. Sourced from [5]

GeForce GTX 970 graphics card. So we decided to build the final prototype around the Air Scouter display.

Eye Tracker and Head Mounted Camera

A key element of the prototype is the need to be able to track the users gaze. None of the head mounted displays had an integrated eye-tracker, so we needed to build one ourselves. We did this based on the open source eye-tracker developed by WearScript [20] To do this we hacked the Microsoft lifeCam HD 5000 camera by breaking its case, switching out the low power eye safe IR LED with a blue LED and removing the IR Filter from it. Once it was done successfully, we designed enclosure for new webcam using SolidWorks [80] (see figure 5.11).



(a) Open Eye Cam

(b) Eye Tracker attached to Google Glass

(c) Output from Eye Tracker

Figure 5.11: Eye Tracker for Google Glass with the help of [20]

In addition to creating a custom eye-tracker we needed to add a second camera to capture the user’s view of the real world. We chose the Logitech’s C920 World Camera due to compatibility with the eye tracking software we used.



(a) Eye Camera ((Microsoft LifeCam HD 5000) Source: Microsoft)



(b) World Camera ((Logitech C920) Source: Logitech)

Figure 5.12: Cameras for Eye Tracker

5.4.2. Software

The Software part of the prototype consisted of two main components. One was the tracking of eye pupil and overlaying gaze marker on the real world video captured from World Camera (Eye Tracking Application). The other part consisted of sharing the mouse pointer over the video in the display of the HMDs (Annotation Application).

The WearScript eye-tracker was designed to work with Google Glass, but since we were using the Brother Air Scouter display, we couldn’t use the same eye-tracking software provided by WearScript. Instead we used the open source eye tracking system developed by Pupil Labs [51]. This software tracks the eye pupil in the video stream from the Microsoft LifeCam HD 6000 webcam and maps it over the video from the World Cam-

era capturing the real world.

The Pupil Labs software is developed using the Python programming language because of its quick performance and ease of use. However, all the high performance media compression code, custom functions, computer vision and display libraries are written in C and accessed via Python using ctypes that glue all the pieces together.

When this application starts, it initiates two following processes simultaneously:

- *Eye Process*: This process is responsible for tracking the eye pupil and broadcasting its position in the eye camera space. It starts by grabbing images from the video stream of eye cam, then applies computer vision algorithms to detect the pupil position from the image, and streams the position of the detected eye pupil (see figure 5.13). The Pupil Lab software is designed to work with the Microsoft LifeCam HD-6000 whereas we used the LifeCam HD-5000, so we had to modify the code slightly .

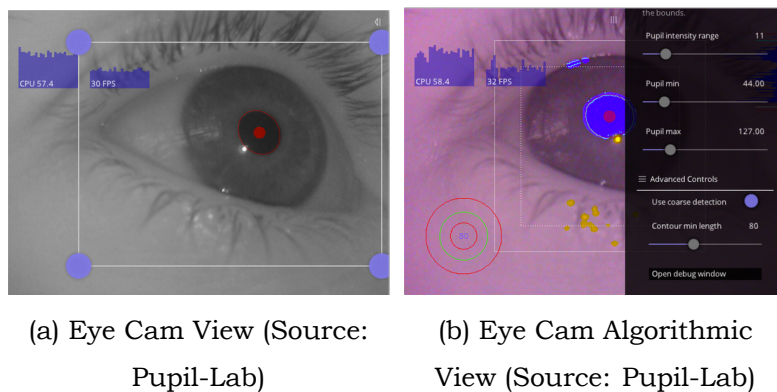


Figure 5.13: Eye Camera View using Pupil-Lab software[51]

- *World Process*: This process grabs images from the video stream of the World Camera, receives the broadcasted

pupil positions from Eye Process, and maps the pupil position in the eye camera space to a gaze position in the world camera space. This mapping depends on the scaling factors that are calculated after calibrating the system. Figure 5.14 shows the outcome, with the green spot showing the user's eye gaze position.

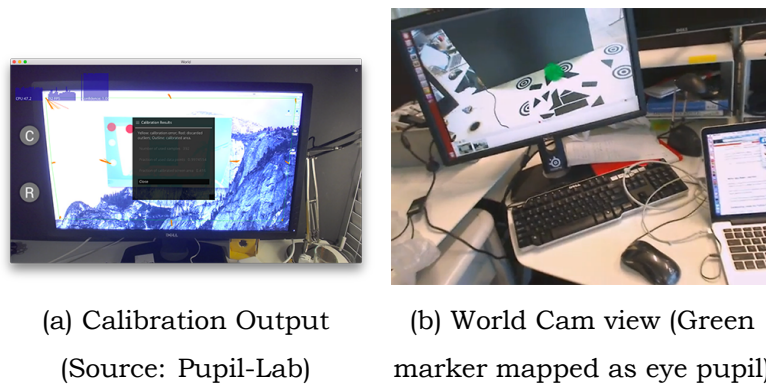


Figure 5.14: World Camera View on Display monitor using Pupil-Lab software[51]

For the remote expert, we wanted to show the local user's view with their eye gaze indicated on top of it. Pupil Lab's default eye tracking viewer application provides this functionality except it adds some additional icon buttons for calibration and recording, and graphs showing CPU processing, the FPS and pupil recognition confidence etc. So to develop the viewing application we just needed to modify the Pupil Lab viewer code to remove those additional components, leaving only the video and eye gaze marker part.

In addition to the application for the remote expert, we needed to create an application for the local user and the head mounted display. To do this we developed an application with the help of a colleague Seungwon Kim (PhD candidate at HIT-

Lab NZ), that replicated the Helper's display except for the mouse pointer movement which was activated in form of a red marker when the Helper clicks on his monitor. The Helper is watching the video stream shared from Worker's HMC along with the eye gaze marker with the freedom to click on the display whenever they wanted to instruct the Worker to perform a task, e.g. "Pick that Object" etc. and that whole video with the instruction marker was displayed back in the HMD. So the local user sees exactly what the remote Helper is seeing on their display (live video of the Helper's environment), except for seeing an additional visual cue showing the remote Helper's mouse pointer (see figure 5.15).

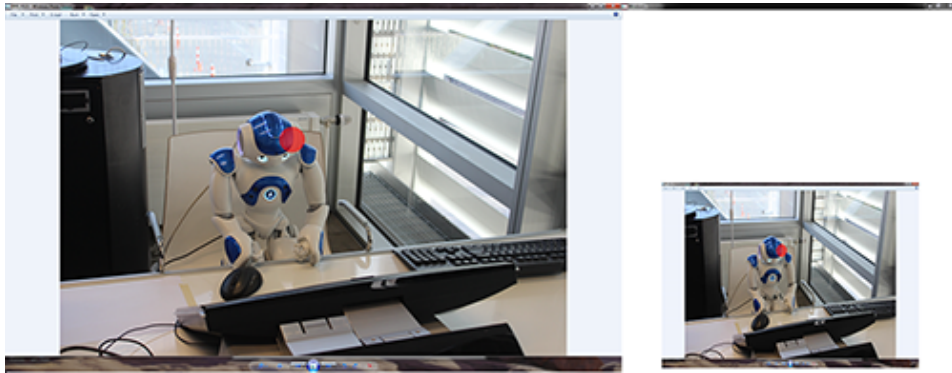


Figure 5.15: Screenshot of multiple display system showing the working of Annotation Application

The above image demonstrates the working of the remote Helper application, where the left display is the main display monitor that was in front of Helper and the right small monitor was the extended monitor view that was later used for the Brother AirScouter HMD. The red dot in both views is the mouse pointer location of the remote Helper.

5.4.3. Final Prototype

The final prototype system used the Eye Cam (Microsoft Life-Cam HD 5000), and World Cam (Logitech C920) mounted on the Brother AirScouter HMD, connected to the computer system via USB ports (see figure 5.16). The Eye Tracker Application was initially developed for Ubuntu, but the driver for AirScouter HMD was available only for Windows 7, so we had to port the entire application code to Windows.



Figure 5.16: Author wearing the system and Eye Tracking Application running in the monitor behind

5.5. User Evaluation

In this section, we report on a user experiment designed to evaluate the system that we designed in a comparative study between various interfaces conditions. This section is divided in further subsections explaining the goal of the evaluation, experimental design, results and analysis.

5.5.1. Evaluation Goal

The main objective of this evaluation was to compare the co-presence level in a traditional HMD based remote collaboration system and a system that provides additional attention information. In order to achieve it, we designed an experiment comparing various interfaces for Helper and Worker with respect to the time taken to complete a particular task and by answering questionnaires that explore various aspects of the collaborative experience including co-presence.

5.5.2. Experiment Design

Hypothesis

The main hypotheses of the experiment were:

- H1: There is a significant difference in co-presence between traditional video conferencing remote collaboration and providing additional cues (i.e. Worker's attention information or Helper's instruction marker) along with traditional video conferencing remote collaboration.
- H2: There is a significant difference in time performance to complete a task between traditional video conferencing remote collaboration and providing additional cues (i.e. Worker's attention information or Helper's instruction marker) along with traditional video conferencing remote collaboration.

The remote collaboration system using pointer for annotation by Helper and eye gaze information of Worker have not

been explored before. From the previous research, we know that pointer increases the connectedness between Worker and Helper [56]. This led to the formulation of H1.

As Neider [66] has suggested that shared gaze condition is twice as fast and efficient than solitary search in a time-critical, coordinating parallel activity spatial task. It also suggested that only shared gaze search is even better than shared gaze-plus-voice search. In our research we want to explore the effect of pointer and eye gaze information from Worker as compared to normal video only task. This led to the formulation of H2.

Experimental Setup

The experimental setup was designed in such a way that it could provide a remote collaboration experience. In order to reproduce remote video conferencing experience in a controlled experiment environment, we setup the whole system in a room with Worker on one side and Helper on the other separated by a large white board. The HMD hardware was at the Worker's side, connected to the computer with the monitor on the Helper. We arranged the whole system to reflect a remote collaboration experience so that both participants were not able to see each other, but can see the shared task space video. The audio communication between the participants was through normal speaking. All the interface conditions were using the same setup except for additional cues provided on the shared video.

The experimental task was to construct structures using LEGO Duplo pieces manipulated by the Worker with assistance from the Helper. We created four different structures with 17 pieces in each task in order to keep the task at a

constant difficulty level. In preliminary tests the time taken to complete each structure was around 90 seconds, and the time difference between structures less than 20 seconds, so the tasks were of similar difficulty level.

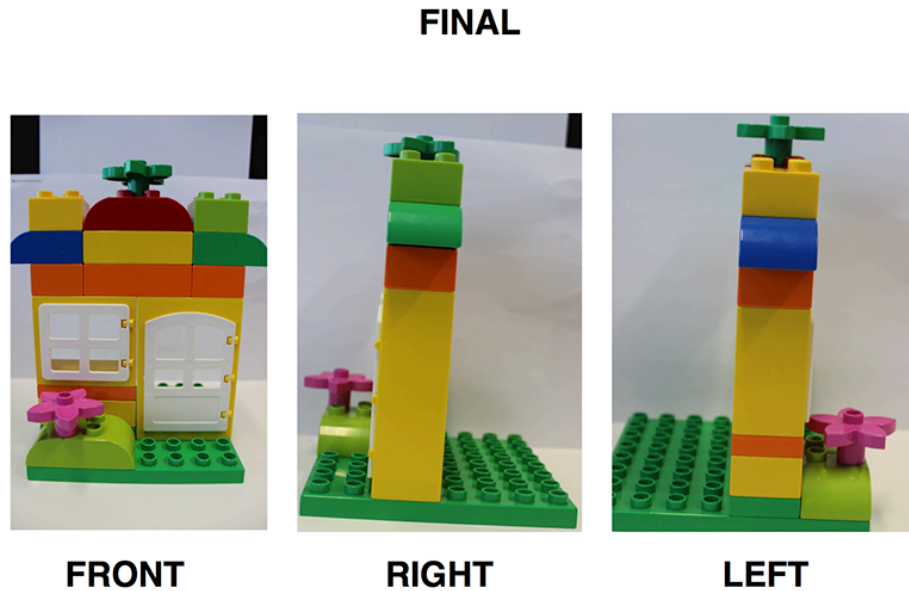


Figure 5.17: Example of task given to perform

To show the benefit of sharing attention information, we added two constraints to the LEGO construction tasks. First, we used two tables arranged in a shape of letter 'L', using one table to keep the LEGO blocks and asking participants to construct the structure on the other. Also, the Workers was allowed to take only one piece at a time from the block table to the main workspace and had to use that block before taking other one. With this configuration participants needed to turn their head from one table to another.

Another constraint was that we introduced an additional divided attention subtask. For this, we introduced a count-down timer which the participant had to pay attention to while constructing the LEGO structure. The timer was created using a four digit seven segment LED display and an Arduino

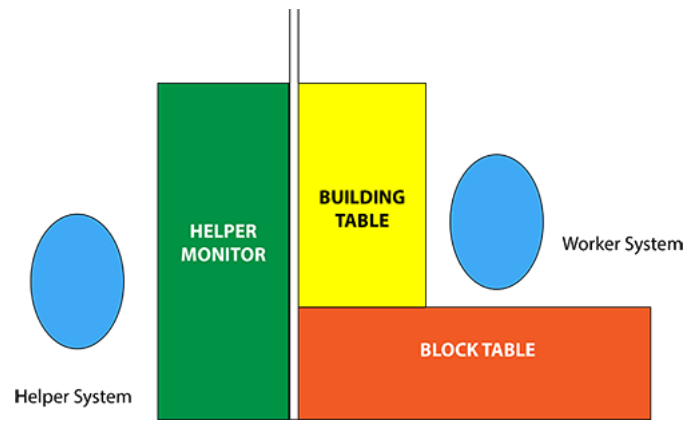


Figure 5.18: Experiment setup layout

micro-controller. It starts its countdown from 40 seconds and reduces the timer towards 0 and continues to negative unless a small tactile button is pressed and resetting the timer to a random number. We placed this system on the table where the participant built the LEGO structure, close enough to the building platform so that it will be visible within the shared camera view. We asked participants (both remote and local) to keep track of it and press the button before it reaches 0. To make sure the participants cannot avoid paying attention to the timer, the reset button was active only when between 0 to 4 seconds were left on the timer. We planned to explore the situation where both users had to keep track of another task while working on the main task, which is common in a real life remote collaboration. For example, while cooking with the help of a remotely assisting expert of that particular dish, the chef has to keep track of numerous factors, such as checking the oven while frying some stuff etc. In our experiment, we asked subjects to keep track of the countdown timer and press the button before it reaches 0, while constructing the LEGO structure.

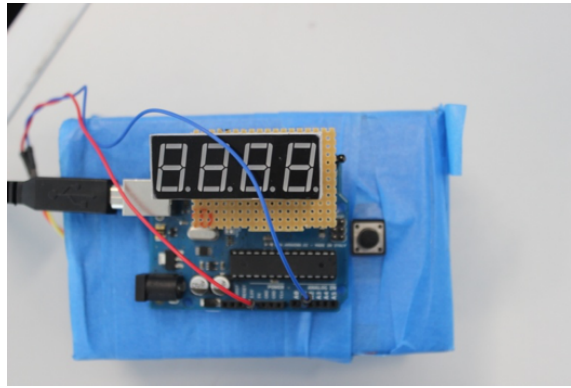


Figure 5.19: Countdown Timer to provide multiple focus points



Figure 5.20: Experiment setup: Participants performing the tasks

Experimental Procedures

We evaluated different interfaces and interaction techniques of the system using a within-subject study of 2 by 2 (i.e. four) conditions. We had two independent variables, POINTER and EYETRACKING, the first representing if there was a instruction point marker from Helper to Worker shown on the HMD and the second representing if there was an eye gaze marker from the Worker to Helper shown on his display. Table 5.1 summarizes the interfaces used in the four conditions

	EYETRACKING: NO	EYETRACKING: YES
POINTER: NO	NO CUE (NONE)	Eye Tracker Cue (ET)

	EYETRACKING: NO	EYETRACKING: YES
POINTER: YES	Pointer Cue (P)	ET and P Cues (BOTH)

Table 5.1: Four conditions with two independent variables.

1. **No Cue (NONE):** In this condition, only Audio and Video cues were shared as a part of video conferencing model for remote assistance collaboration.
2. **Pointer Cue (P):** Pointer refers to the instruction point marker used by the Helper to instruct the Worker. When the Helper clicks on the shared video on his display, the marker appears at the clicked position, which is also visible on the Worker's HMD. In this condition, pointer is provided as an additional cue in addition to the other cues in the NONE condition.
3. **Eye Tracker Cue (ET):** Eye Tracker cue refers to the Eye Gaze marker displayed on Helper's Monitor as a cue to show focus of attention (FoA) of the Worker in his task space. In this condition, Eye Tracker was provided as an additional cue along with the other cues in NONE condition.
4. **ET and P Cues (BOTH):** In this condition, both Eye tracker and Pointer Cues were provided to both participants in addition to the cues provided in NONE condition.

The order of conditions tried by the participants was changed using a Balanced Latin Square design in order to counterbalance the carryover effects between the conditions.

Participants were recruited from the university by posting advertisements (can be found in Appendix C) on the notice board and by sending an email to the HITLab's mailing list. Participants were randomly assigned as a Worker or a Helper. Then they were asked to read an information sheet and sign the consent form for participating in the experiment. The information consent form with a copy of questionnaire for both Helper (remote user) and Worker (local user) participants can be found in Appendix. We gave them a pre-task questionnaire asking for demographic information of the participants, including their previous experience with remote collaboration using video conferencing and using LEGO blocks. After answering this questionnaire, they were told about the main objective of this research, and the experimenter demonstrated the interfaces by explaining the cues provided with each interface.

Most of the participants had never worked with each other to make anything with LEGO pieces even in a face-to-face collaboration. Before the experimental sessions, we provided them a practice face-to-face collaborative task in which the Helper was provided with step-by-step instruction manual of the structure and asked to assist the Worker who just had the access to the LEGO blocks but completely unaware about the final structure. Through this practice task, both participants were familiarized with each other's communication skills and had experience of constructing a LEGO structure together remotely.

After the practice task, participants were separated to sit at their desks and perform the experimental tasks using the provided interface in each condition. In each condition, in or-



Figure 5.21: Participants performing practice task

der to let the Helper get familiarized with the task before giving instruction to the Worker, we let the Helper create the structure by himself first at his workspace following the instructions provided. Then the same LEGO blocks were provided to the Worker to perform the experimental task following the instructions given by the Helper. After each condition, both the Helper and Worker were asked to complete a questionnaire with Likert scale questions and the time taken for completing the task was recorded. After trying all the conditions, participants were asked to fill a post-experiment questionnaire consisting of questions where they were asked to rank the interfaces in based on various aspects of their experiences as stated in each question.

Participants

In order to tweak the system and debug it before the real experiment, we conducted pilot tests with two pairs of participants (Worker and Helper). The results from these pilot tests are not counted in the final results of the experiment.

We invited 15 pairs (30 participants) to participate in the real experiment, but the data of only 13 was used as 1 pair quit the experiment early, and Helper of one pair rated all the



(a) Worker performing the task



(b) Helper assisting the Worker in condition ET

Figure 5.22: Experiment in Progress

conditions as 7 on the likert scale because he had to go somewhere, so that pair was counted as an outlier. The participants had an average experience of remote collaboration using video conference once a year, however none of the participants had previous experience of constructing LEGO structures over video conferencing.

All of the participants were university students within aged 21 to 33, out of which 23 (76.67%) were male and 7 (23.33%) were female and 18 (60%) were non-native English speakers,

although all of them had a good understanding and speaking level of English.

5.6. Results

This section reports on the analyzed results from the experiment. First, we report on the results of the data recorded from the task such as time performance in performing a task. Next, we evaluate the quantitative data gathered from the participants in the form of questionnaire. Finally we summarize the qualitative feedback collected through open questions in the questionnaire where participants wrote down their thoughts about the system.

5.6.1. Task Performance Time

The task completion time (measured in seconds) is analyzed using two-way repeated measures ANOVA test ($\alpha=.05$).

A repeated measure two-way ANOVA revealed that there was a significant main effect of both POINTER ($F(1, 12)=4.908, p=.047$) and EYETRACKING ($F(1, 12)=5.811, p=.033$) on the time taken to complete a task using these interfaces. No significant interaction was found between POINTER and EYETRACKING ($F(1, 12) = 0.566, p=0.466$).

Descriptive statistics (see Figure 5.23 and Table 5.2) shows that the participants took less time to complete the task in ET ($Mean=245.7, Std. dev.=61.9$) and P ($M=234.5, SD=74.6$) conditions compared to the baseline NONE condition ($M=258.3, SD=70.8$). The overall performance of participants was fastest while using the interface in BOTH condition ($M=200.5, SD=50.7$).

Condition	Mean	Std. Deviation	N
NONE	258.338	70.8307	13
ET	245.692	61.8821	13
P	234.462	74.5739	13
BOTH	200.538	50.7028	13

Table 5.2: Task Performance Time

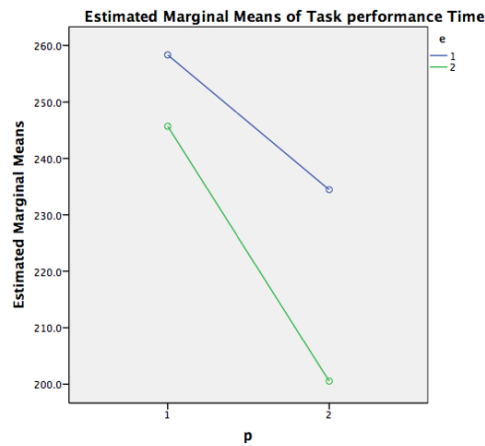


Figure 5.23: Task Time Performance: Interaction between the conditions

5.6.2. Questionnaire: Quantitative Measure

After finishing the task in each condition, participants were asked to answer a questionnaire that included 11 questions on various aspects of the collaborative experience by rating on a Likert scale (Range 1 to 7 where 1 was strongly disagree and 7 was strongly agree). After finishing all four conditions, participants were given a questionnaire to rank (from BEST to WORST) the conditions based on their experiences with respect to the question statements asked. We mapped the BEST, SECOND BEST, THIRD BEST and WORST responses as 4, 3, 2 and 1 respectively for the ease of evaluation. We designed the questionnaire by referring to an existing ques-

tionnaire used in previous research by Seungwon Kim [56] on improving co-presence in video conferencing. Since we had an important factor, “focus of attention of Worker” in our research, so we modified the questionnaire by adding few questions asking about understanding the focus of attention of the partner.

To analyzing the results of both Likert scale rating and ranking responses, we decided to use the Aligned Rank Transform (ART) for non-parametric factorial analyses using ANOVA procedures ($\alpha=.05$) proposed by Wobbrock et al. [95]. Compared to the Friedman test, this method allows factorial analysis of the results so that we can preserve the 2x2 factorial design of the experiment when analyzing the results in ordinal measures.

Likert Scale Rating Questionnaire

Here we report on the results of the Likert scale rating questions. Table 5.3 lists the 11 question about various aspects of the collaborative experience. The table also summarizes the results of inferential statistics showing the significance of the main effects of each factor POINTER and EYETRACKING and their interaction for Local (Worker) and Remote (Helper). Overall for the local user, POINTER had a significant main effect on Q1, Q2, Q3, Q4, Q9, Q10 and Q11, and EYETRACKING had a significant main effect on Q1, Q2, Q3, Q4, Q6, Q8, Q9, Q10, and Q11. However, for remote user, POINTER had a significant main effect on Q1, Q2, Q3, Q4, Q8, Q9, Q10 and Q11, and EYETRACKING had a significant main effect on Q1, Q2, Q5, Q9, Q10, and Q11. In the rest of this section we report on further details of the analysis.

User	Q No.	Question	p-values of effects		
			P	E	P X E
Local	Q1	I felt connected with my partner	0.009	0.014	0.014
	Q2	I was present with my partner	0.044	0.003	0.659
	Q3	my partner was able to sense that I was present with him	0.022	0.010	0.298
	Q4	Partner could tell when I needed assistance	0.013	0.001	0.037
	Q5	Enjoyed the task	0.917	0.063	0.521
	Q6	Focused on task	0.581	0.040	0.619
	Q7	completed the task	0.695	0.062	0.081
	Q8	We worked together	0.546	0.019	0.233
	Q9	I was able to express myself clearly	0.027	0.001	0.095
	Q10	Understood partner's response	0.002	0.002	0.261
	Q11	info from partner was useful	0.005	0.008	0.063
Remote	Q1	I felt connected with my partner	0.002	0.001	0.030
	Q2	I was present with my partner	0.010	0.016	0.033

User	Q No.	Question	p-values of effects		
			P	E	P X E
Remote	Q3	my partner was able to sense that I was present with him	0.000	0.052	0.014
	Q4	I could tell when my partner needed assistance	0.009	0.089	0.012
	Q5	Enjoyed the task	0.123	0.036	0.640
	Q6	Focused on task	0.065	0.256	0.458
	Q7	completed the task	0.134	0.140	0.128
	Q8	We worked together	0.042	0.395	0.011
	Q9	I was able to express myself clearly	0.004	0.002	0.813
	Q10	Understood partner's response	0.010	0.001	0.241
	Q11	info from partner was useful	0.018	0.001	0.189

Table 5.3: Summary of the results of inferential statistics for the Likert Questionnaire

Q1. I felt connected with my partner

Local User: Worker

A repeated measure two-way ANOVA with ART revealed that for the local users there was a significant main effect of both POINTER ($F(1, 12)=9.763, p=.009$) and EYETRACK-

ING ($F(1, 12)=8.220, p=.014$) on how connected they felt to the remote user. There was a significant interaction between POINTER and EYETRACKING ($F(1, 12) = 8.291, p=.014$). As shown in Figure 5.24 and Table 5.4, compared to the NONE condition ($Mean=4.62, Std. Dev.$

$=1.19$) participants for both ET ($M=5.69, SD=0.85$) and P ($M=6.00, SD=0.71$) conditions significantly higher. The significant interaction between the two factors and the rating for the BOTH condition ($Mean = 6.15, std. dev. = 0.89$) being marginally higher than ET and P conditions can be explained by the ceiling effect where the rating is getting saturated as the BOTH condition rated close to the highest possible value of 7.

Condition	Mean	Std. Deviation	N
NONE	4.615	1.1929	13
ET	5.692	.8549	13
P	6.000	.7071	13
BOTH	6.154	.8987	13

Table 5.4: Mean connectedness (out of 7) for different conditions (Helper)

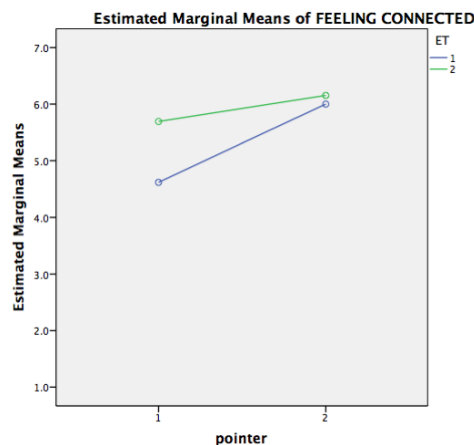


Figure 5.24: Connectedness: Interaction between the conditions (Worker)

Remote User: Helper

A repeated measure two-way ANOVA with ART revealed that for the remote users there was a significant main effect of both POINTER ($F(1, 12)=15.096, p=.002$) and EYETRACKING ($F(1, 12)=17.153, p=.001$) on the level of feeling connected. There was a significant interaction between POINTER and EYETRACKING ($F(1, 12) = 6.052, p=.030$). Descriptive statistics (see Table 5.5 and Figure 5.25) show that P ($M=5.46, SD=0.66$) and ET ($M=5.62, SD=1.04$) conditions are rated higher compared to NONE condition ($M=4.23, SD=1.17$), and the BOTH condition ($M=6.08, SD=1.04$) is rated higher than ET and P conditions. The significant interaction between the two factors can be explained by the ceiling effect of the rating for BOTH condition being saturated as reaching the highest possible value.

Condition	Mean	Std. Deviation	N
NONE	4.231	1.1658	13
ET	5.615	1.0439	13
P	5.462	.6602	13
BOTH	6.077	1.0377	13

Table 5.5: Mean connectedness (out of 7) for different conditions (Helper)

Q2. I felt that I was present with my partner on the same workspace while performing the task.

Local User: Worker

A repeated measure two-way ANOVA with ART revealed that for the local users (Worker) there was a significant main

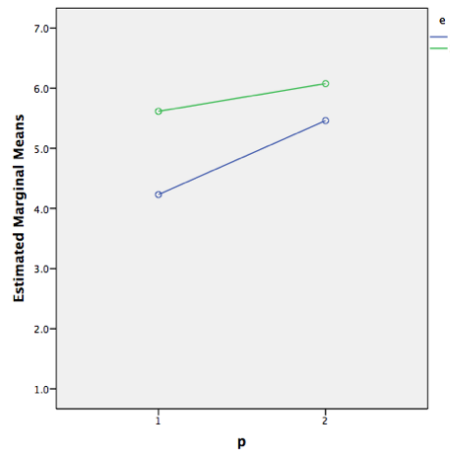


Figure 5.25: Connectedness: Interaction between the conditions (Helper) effect of both *POINTER* ($F(1, 12)=5.086, p=.044$) and *EYETRACKING* ($F(1, 12)=14.153, p=.003$) on the participant’s feeling of co-presence with the remote helper while using these interfaces. No significant interaction was found between *POINTER* and *EYETRACKING* ($F(1,12) = 0.205, p=0.659$). Descriptive statistics (see *Table 5.6* and *Figure 5.26*) show that the NONE condition ($Mean=5.077, std. dev.=1.1152$) is comparatively rated lower than P condition ($Mean=5.538, std. dev.=0.6602$) and ET condition ($Mean=6.000, std. dev.=0.7071$). And the BOTH condition ($Mean = 6.462, std. dev. = 0.6602$) is rated comparatively higher than the other three conditions

Condition	Mean	Std. Deviation	N
NONE	5.077	1.1152	13
ET	6.000	0.7071	13
P	5.538	0.6602	13
BOTH	6.462	0.6602	13

Table 5.6: Mean presence (out of 7) for different conditions (Worker)

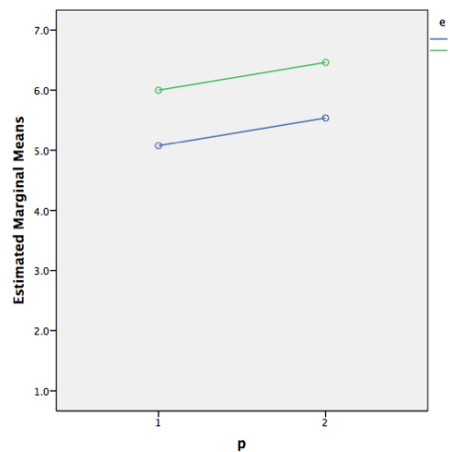


Figure 5.26: I was present: Interaction between the conditions (Worker)

Remote User: Helper

A repeated measure two-way ANOVA with ART revealed that for the remote users there was a significant main effect of both POINTER ($F(1, 12)=9.412, p=.010$) and EYETRACKING ($F(1, 12)=7.926, p=.016$) on the participant's subjective sense of presence with the partner while using these interfaces. There was a significant interaction between POINTER and EYETRACKING ($F(1, 12) = 5.781, p=.033$). Descriptive statistics (see Table 5.7 on 104 and Figure 5.27 on 104) show that the NONE condition (Mean=4.231, std. dev.=1.4806) is rated comparatively lower than ET condition (Mean=5.385, std. dev.=0.9608) and P condition (Mean=5.769, std. dev.=0.8321). The significant interaction between the two variables can be explained by as the rating for BOTH condition (Mean=5.923, std. dev.=1.1875) reaching the upper bound of the rating range, the effect of the two factors are being saturated.

Condition	Mean	Std. Deviation	N
NONE	4.231	1.4806	13

Condition	Mean	Std. Deviation	N
ET	5.385	.9608	13
P	5.769	.8321	13
BOTH	5.923	1.1875	13

Table 5.7: Mean presence (out of 7) for different conditions (Helper)

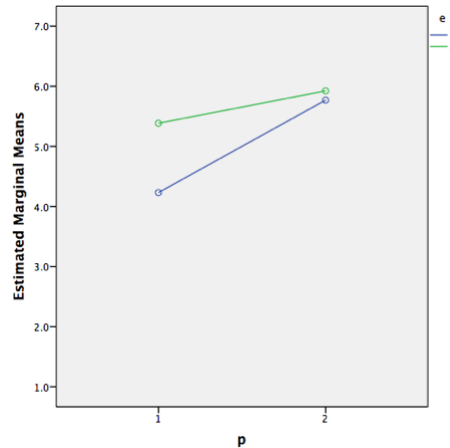


Figure 5.27: I was present: Interaction between the conditions (Helper)

Q3. I think my partner was able to sense that I was present with him on the same workspace while performing the task.

Local User: Worker

A repeated measure two-way ANOVA with ART revealed that for the local users (Worker) there was a significant main effect of both POINTER ($F(1, 12)=6.858, p=.022$) and EYETRACKING ($F(1, 12)=9.179, p=.010$) on feeling that her partner can sense her presence. No significant interaction was found between POINTER and EYETRACKING ($F(1, 12) = 1.185, p=0.298$). Descriptive statistics (see Figure 5.28 and Table 5.8) show that the NONE condition ($Mean=4.846, std. dev.=0.9871$) is

rated comparatively lower than P condition ($Mean=5.462$, $std. dev.=0.6887$) and ET condition ($Mean=5.846$, $std. dev.=0.6887$). The BOTH condition ($Mean = 6.077$, $std. dev. = 0.7596$) is rated higher than the other three conditions.

Condition	Mean	Std. Deviation	N
NONE	4.846	.9871	13
ET	5.846	.6887	13
P	5.462	.7763	13
BOTH	6.077	.7596	13

Table 5.8: Mean partner sensed my presence (out of 7) for different conditions (Worker)

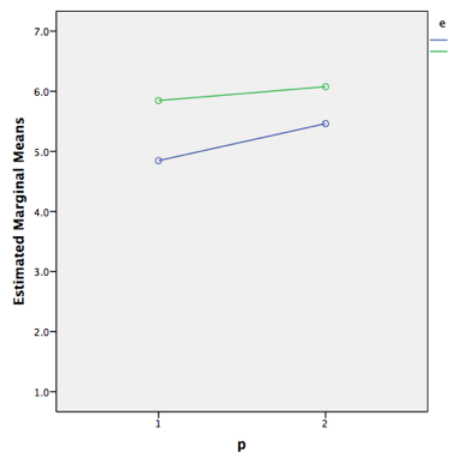


Figure 5.28: Partner sensed presence: Interaction between the conditions (Worker)

Remote User: Helper

A repeated measure two-way ANOVA with ART revealed that for the remote users there was a significant main effect of POINTER ($F(1, 12)=22.511$, $p=.000$) but no significant main effect of EYETRACKING ($F(1, 12)=4.6481$, $p=.052$) on the feeling

that her partner can sense her presence while using these conditions. However, there was a significant interaction between POINTER and EYETRACKING ($F(1, 12) = 8.359, p = .014$). Descriptive statistics (see Table 5.9 on 106 and Figure 5.29 on 107) shows that the NONE condition (Mean=4.154, std. dev.=1.1435) is rated lower than the other conditions: ET (Mean=5.385, std. dev.=0.8697), P (Mean=6.154, std. dev.=0.6887), and BOTH (M=6.1, SD=0.9) conditions. While the BOTH condition is rated lower than P condition, the difference was not statistically significant based on Wilcoxon Signed Rank test ($Z = -0.276, p = 0.783$), whereas NONE and ET is significantly different based on Wilcoxon Signed Rank test ($Z = -2.654, p = 0.008$). The interaction between the two factors appears to be due to the effects of the two factors being saturated as they are combined together.

Condition	Mean	Std. Deviation	N
NONE	4.154	1.1435	13
ET	5.385	.8697	13
P	6.154	.6887	13
BOTH	6.077	.8623	13

Table 5.9: Mean presence sensed by partner (out of 7) for different conditions (Helper)

Q4. Partner knew when I needed assistance

Local User: Worker

A repeated measure two-way ANOVA with ART revealed that for the local users there was a significant main effect of both POINTER ($F(1, 12) = 8.372, p = .013$) and EYETRACKING

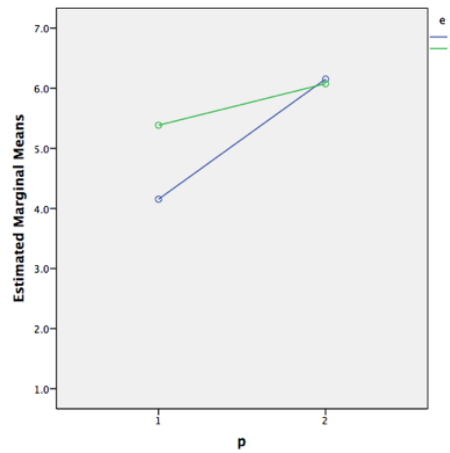


Figure 5.29: Partner sensed presence: Interaction between the conditions (Helper)

($F(1, 12)=19.761, p=.001$) on the participant's subjective rating on whether the partner could tell when the local user needed assistance while using these interfaces. There was a significant interaction between POINTER and EYETRACKING ($F(1, 12)=5.501, p=.037$). Descriptive statistics (see Figure 5.30 and Table 5.10) shows that the NONE condition ($Mean=4.462, std. dev.=1.4806$) is rated comparatively lower than ET condition ($Mean=5.846, std. dev.=0.8006$) and P condition ($Mean=5.846, std. dev. = 0.8987$). The significant interaction between the two variables suggests that the mean for the BOTH condition ($Mean=6.538, std. dev.=0.5189$) having ceiling effect and the effect of the two variables being saturated as the eye tracker and pointer were used together.

Condition	Mean	Std. Deviation	N
NONE	4.462	1.4500	13
ET	5.846	.8006	13
P	5.846	.8987	13
BOTH	6.538	.5189	13

Table 5.10: Mean: Partner knew when I needed assistance (out of 7) for different conditions (Worker)

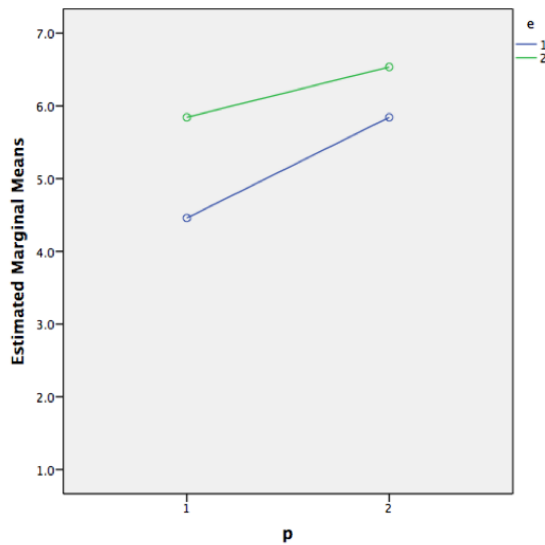


Figure 5.30: Partner knew when I needed assistance: Interaction between the conditions (Worker)

Remote User: Helper

A repeated measure two-way ANOVA with ART revealed that for the remote users there was a significant main effect of POINTER ($F(1, 12)=9.828, p=.009$) but no significant main effect of EYETRACKING ($F(1, 12)=3.434, p=.089$) on whether partner could tell when she needed assistance while using these interfaces. However, there was a significant interaction between POINTER and EYETRACKING ($F(1, 12) = 8.715, p=.012$). Descriptive statistics (see Figure 5.31 and Table 5.11) show that the NONE condition ($Mean=4.692, std. dev.=1.4936$) is rated comparatively lower than the other three conditions: P condition ($Mean=5.615, std. dev.=0.9608$), ET ($Mean=5.692, std. dev.=1.0316$), and BOTH ($Mean = 6.462, std. dev. = 0.6602$).

Condition	Mean	Std. Deviation	N
NONE	4.692	1.4936	13

Condition	Mean	Std. Deviation	N
ET	5.692	1.0316	13
P	5.615	.9608	13
BOTH	5.923	1.0377	13

Table 5.11: Mean: Partner needed assistance (out of 7) for different conditions (Helper)

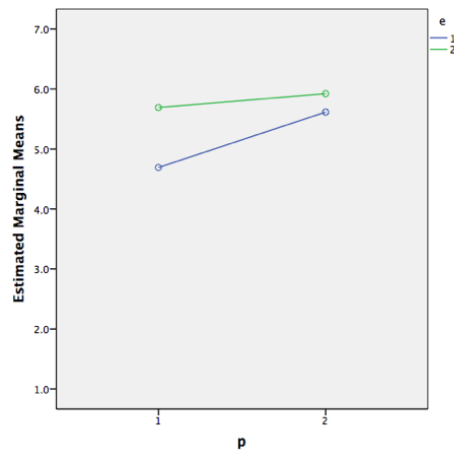


Figure 5.31: Partner needed assistance: Interaction between the conditions (Helper)

Q5. I enjoyed the experience

Local User: Worker

A repeated measure two-way ANOVA with ART revealed that for the local users there was no significant main effect of EYETRACKING ($F(1, 12)=4.195, p=.063$) and POINTER ($F(1, 12)=0.011, p=.917$) on the perceived level of enjoyment in constructing the task while using these conditions. There was also no significant interaction between POINTER and EYETRACKING ($F(1, 12) = 0.436, p = 0.521$). Descriptive statistics are shown in Table 5.12 and Figure 5.32.

Condition	Mean	Std. Deviation	N
NONE	5.538	.9674	13
ET	6.077	.6405	13
P	5.692	.7511	13
BOTH	5.846	.8987	13

Table 5.12: Mean: Enjoyment (out of 7) for different conditions (Worker)

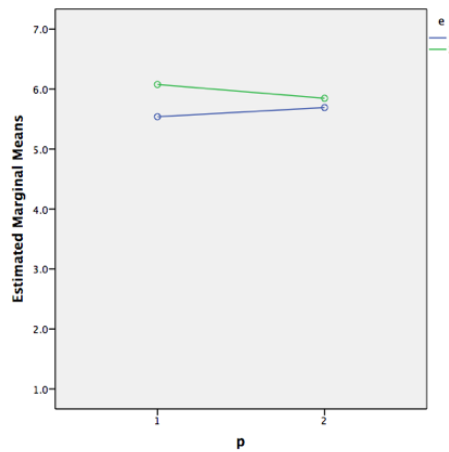


Figure 5.32: Enjoyed constructing the task: Interaction between the conditions (Worker)

Remote User: Helper

A repeated measure two-way ANOVA with ART revealed that for the remote users there was a significant main effect of EYETRACKING ($F(1, 12)=5.589, p=.036$) but no significant main effect of POINTER ($F(1, 12)=2.753, p=.123$) on the level of enjoyment in constructing the task while using these conditions. There was also no significant interaction between POINTER and EYETRACKING ($F(1,12) = 0.230, p = 0.640$). Descriptive statistics (see [Figure 5.33](#) and [Table 5.13](#)) shows that the conditions with EYETRACKING (ET and BOTH combined

Mean = 5.735, Std. Dev = 1.167) was rated higher than those conditions without EYETRACKING (NONE and P combined, Mean = 5.269, Std. Dev. = 1.016).

Condition	Mean	Std. Deviation	N
NONE	5.000	1.1547	13
ET	5.615	1.1209	13
P	5.538	.8771	13
BOTH	5.846	1.2142	13

Table 5.13: Mean: Enjoyment (out of 7) for different conditions (Helper)

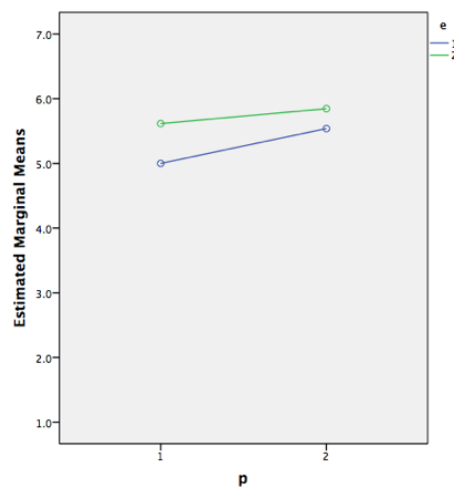


Figure 5.33: Enjoyed constructing the task: Interaction between the conditions (Helper)

Q6. I was able to focus on the task activity

Local User: Worker

A repeated measure two-way ANOVA with ART revealed that for the local users there was a significant main effect of EYETRACKING ($F(1, 12)=5.334, p=.040$) but no significant main effect of POINTER ($F(1, 12)=0.321, p=.581$) level of fo-

cus on the task while using these conditions. There was also no significant interaction between POINTER and EYETRACKING ($F(1,12) = 0.260, p = 0.619$). Descriptive statistics (see Table 5.14 and Figure 5.34) shows that the conditions with EYETRACKING (ET and BOTH combined, $Mean = 5.9615, Std. Dev = 0.7285$) were rated higher than those conditions without EYETRACKING (NONE and P combined, $Mean = 5.423, Std. Dev=1.074$).

Condition	Mean	Std. Deviation	N
NONE	5.308	1.1821	13
ET	5.923	.6405	13
P	5.538	.9674	13
BOTH	6.000	.8165	13

Table 5.14: Mean: Focus (out of 7) for different conditions (Worker)

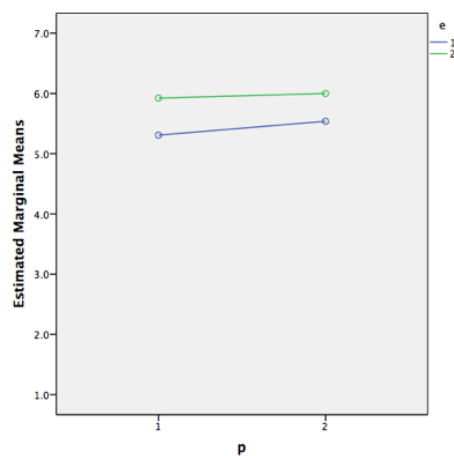


Figure 5.34: Focused on the task: Interaction between the conditions (Worker)

Remote User: Helper

A repeated measure two-way ANOVA with ART revealed

that for the remote users there was no significant main effect of EYETRACKING ($F(1, 12)=5.334, p=.040$) and POINTER ($F(1, 12)=0.321, p=.581$) on the perceived level of focus on the task while using these conditions. There was also no significant interaction between POINTER and EYETRACKING ($F(1, 12) = 0.260, p = 0.619$). Table 5.15 show the descriptive statistics.

Condition	Mean	Std. Deviation	N
NONE	5.385	1.0439	13
ET	5.615	1.1209	13
P	5.923	.6405	13
BOTH	5.846	1.1435	13

Table 5.15: Mean: Focus (out of 7) for different conditions (Helper)

Q7. I am confident that we completed the task correctly.

Local User: Worker

A repeated measure two-way ANOVA with ART revealed that for the local users there was no significant main effect of EYETRACKING ($F(1, 12)=4.248, p=.062$) and POINTER ($F(1, 12)=0.161, p=.695$) on the confidence of completing the task while using these conditions. There was also no significant interaction between POINTER and EYETRACKING ($F(1, 12) = 3.636, p = 0.081$). Descriptive statistics are shown in Table 5.16.

Condition	Mean	Std. Deviation	N
NONE	6.154	.9871	13
ET	6.308	.6304	13

Condition	Mean	Std. Deviation	N
P	6.000	1.0000	13
BOTH	6.615	.6504	13

Table 5.16: Mean: Task completion confidence (out of 7) for different conditions (Worker)

Remote User: Helper

A repeated measure two-way ANOVA with ART revealed that for the remote users there was no significant main effect of EYETRACKING ($F(1, 12)=2.500, p=.140$) and POINTER ($F(1, 12)=2.585, p=.134$) on the confidence of completing the task while using these conditions. There was also no significant interaction between POINTER and EYETRACKING ($F(1,12) = 2.678, p = 0.128$). Table 5.17 show the descriptive statistics.

Condition	Mean	Std. Deviation	N
NONE	5.615	1.1209	13
ET	6.154	.6887	13
P	6.231	.7250	13
BOTH	6.231	.9268	13

Table 5.17: Mean: Task completion confidence (out of 7) for different conditions (Helper)

Q8. My partner and I worked together well on the task

Local User: Worker

A repeated measure two-way ANOVA with ART revealed

that for the local users there was a significant main effect of EYETRACKING ($F(1, 12)=7.303, p=.019$) but no significant main effect of POINTER ($F(1, 12)=0.386, p=.546$) on the perceived level of how well the participants worked together. There was also no significant interaction between POINTER and EYETRACKING ($F(1,12) = 1.581, p = 0.233$). Descriptive statistics (see Figure 5.35 and Table 5.18) shows that the conditions with EYETRACKING (*ET and BOTH combined Mean = 6.2695, Std. Dev = 0.7381*) were rated higher than those conditions without EYETRACKING (*NONE and P combined, Mean = 5.7305, Std. Dev. = 0.91485*).

Condition	Mean	Std. Deviation	N
NONE	5.538	.9674	13
ET	6.308	.7511	13
P	5.923	.8623	13
BOTH	6.231	.7250	13

Table 5.18: Mean: Worked together (out of 7) for different conditions (Worker)

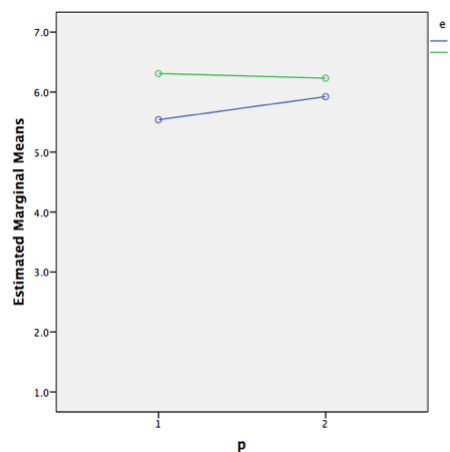


Figure 5.35: Worked together: Interaction between the conditions (Worker)

Remote User: Helper

A repeated measure two-way ANOVA with ART revealed that for the remote users there was a significant main effect of POINTER ($F(1, 12)=5.172, p=.042$) but no significant main effect of EYETRACKING ($F(1, 12)=0.777, p=.395$) on the working together to complete the task while using these conditions. However, there was a significant interaction between POINTER and EYETRACKING ($F(1, 12) = 9.060, p=.011$). Descriptive statistics (see Figure 5.36 on 117 and Table 5.19 on 116) show that the NONE condition ($Mean=5.308, std. dev.=1.2506$) is rated comparatively lower than the other three conditions: P condition ($Mean=6.231, std. dev.=0.7250$), ET ($Mean=5.769, std. dev.=.8321$), and BOTH ($Mean = 6.154, std. dev. = 0.6887$). While the BOTH condition is rated lower than P condition, the difference was not statistically significant based on Wilcoxon Signed Rank test ($Z = -0.447, p = 0.655$), whereas NONE and ET is significantly different based on Wilcoxon Signed Rank test ($Z = -2.121, p = 0.034$). The interaction between the two factors appears to be due to the effects of the two factors being saturated as they are combined together.

Condition	Mean	Std. Deviation	N
NONE	5.308	1.2506	13
ET	5.769	.8321	13
P	6.231	.7250	13
BOTH	6.154	.6887	13

Table 5.19: Mean: Worked together confidence (out of 7) for different conditions (Helper)

Q9. I felt that I was able to express myself clearly to my partner

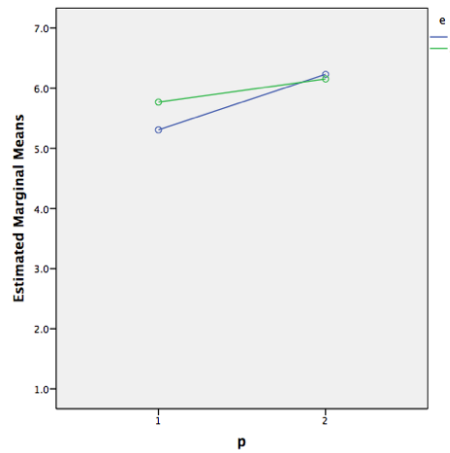


Figure 5.36: Worked together (Helper): A repeated measure two-way ANOVA followed by ART

Local User: Worker

A repeated measure two-way ANOVA with ART revealed that for the local users (Worker) there was a significant main effect of both POINTER ($F(1, 12)=6.381, p=.027$) and EYE-TRACKING ($F(1, 12)=17.388, p=.001$) on expressing clearly to their partners while using these conditions. No significant interaction was found between POINTER and EYETRACKING ($F(1,12) = 3.275, p=0.095$). Descriptive statistics (see Figure 5.37 on page 118 and Table 5.20) show that NONE condition ($Mean=4.385, std. dev.=1.0439$) is rated comparatively lower than P condition ($Mean=5.154, std. dev.=0.8987$) and ET condition ($Mean=5.769, std. dev.=0.8321$). The BOTH condition ($Mean = 6.000, std. dev. = 0.8165$) is rated higher than the other three conditions.

Condition	Mean	Std. Deviation	N
NONE	4.385	1.0439	13
ET	5.769	.8321	13
P	5.154	.8987	13

Condition	Mean	Std. Deviation	N
BOTH	6.000	.8165	13

Table 5.20: Mean: Expressed clearly (out of 7) for different conditions
(Worker)

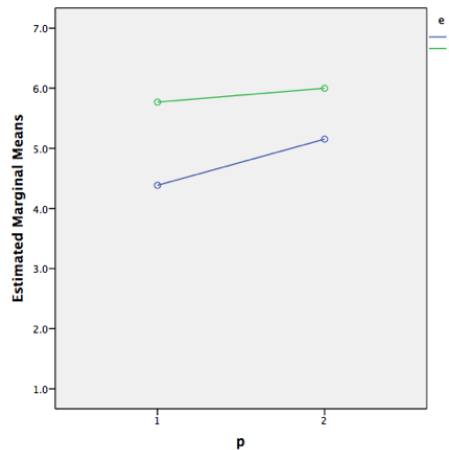


Figure 5.37: Expressed Clearly: Interaction between the conditions
(Worker)

Remote User: Helper

A repeated measure two-way ANOVA with ART revealed that for the local users (Worker) shows that there was a significant main effect of both POINTER ($F(1, 12)=13.119, p=.004$) and EYETRACKING ($F(1, 12)=14.944, p=.002$) on the information provided by Helper while using these conditions. No significant interaction was found between POINTER and EYETRACKING ($F(1,12) = 0.059, p=0.813$). Descriptive statistics (see Figure 5.38 and Table 5.21) show that NONE condition ($Mean=4.385, std. dev.=1.3868$) is rated comparatively lower than P condition ($Mean=5.538, std. dev.=0.9674$) and ET condition ($Mean=5.000, std. dev.=1.000$). And the BOTH condi-

tion ($Mean = 6.308$, $std. dev. = 0.8549$) is rated higher than the other three conditions.

Condition	Mean	Std. Deviation	N
NONE	4.385	1.3868	13
ET	5.000	1.000	13
P	5.538	0.9674	13
BOTH	6.308	.8549	13

Table 5.21: Mean: Expressed clearly (out of 7) for different conditions (Helper)

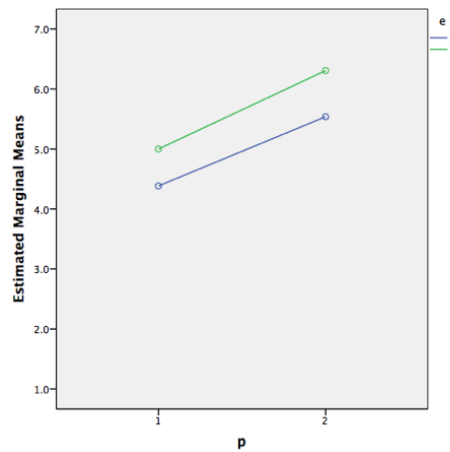


Figure 5.38: Expressed Clearly: Interaction between the conditions (Helper)

Q10. I was able to understand what my partner was communicating to me

Local User: Worker

A repeated measure two-way ANOVA with ART revealed that for the local users (Worker) there was a significant main effect of both POINTER ($F(1, 12)=14.690$, $p=.002$) and EYE-TRACKING ($F(1, 12)=14.739$, $p=.002$) on understanding what

the Helper was communicating while using these conditions. No significant interaction was found between POINTER and EYETRACKING ($F(1,12) = 1.389, p=0.261$). Descriptive statistics (see Figure 5.39 and Table 5.22) show that NONE condition ($Mean=4.923, std. dev.=1.1152$) is rated comparatively lower than P condition ($Mean=6.154, std. dev.=0.6887$) and ET condition ($Mean=5.615, std. dev.=0.7679$). The BOTH condition ($Mean = 6.462, std. dev. = 0.7763$) is rated higher than the other three conditions.

Condition	Mean	Std. Deviation	N
NONE	4.923	1.1152	13
ET	5.615	.7679	13
P	6.154	.6887	13
BOTH	6.462	.7763	13

Table 5.22: Mean: Understood Partner (out of 7) for different conditions (Worker)

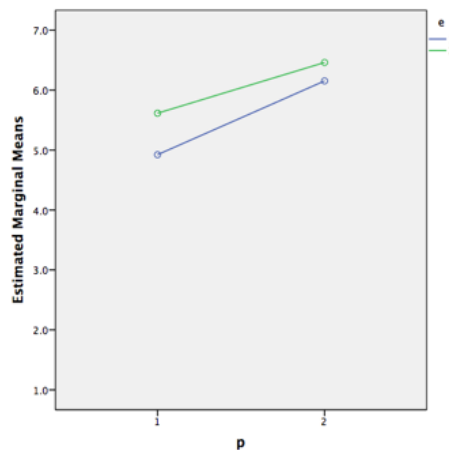


Figure 5.39: Understood the partner: Interaction between the conditions (Worker)

Remote User: Helper

A repeated measure two-way ANOVA with ART revealed that for the local users (Worker) shows that there was a significant main effect of both POINTER($F(1, 12)=9.273, p=.010$) and EYETRACKING ($F(1, 12)=18.381, p=.001$) on the information provided by Helper while using these conditions. No significant interaction was found between POINTER and EYETRACKING ($F(1,12) = 1.523, p=0.241$). Descriptive statistics (see Figure 5.40 and Table 5.23) show that NONE condition ($Mean=4.538, std. dev.=1.4500$) is rated comparatively lower than P condition ($Mean=5.385, std. dev.=.8697$) and ET condition ($Mean=5.692, std. dev.=0.9473$). And the BOTH condition ($Mean = 6.308, std. dev. = 0.8549$) is rated higher than the other three conditions.

Condition	Mean	Std. Deviation	N
NONE	4.538	1.4500	13
ET	5.692	0.9473	13
P	5.385	0.98697	13
BOTH	6.308	.8549	13

Table 5.23: Mean: Expressed clearly (out of 7) for different conditions (Helper)

Q11. The information provided by my partner in this condition helped in easily performing the task.

Local User: Worker

A repeated measure two-way ANOVA with ART revealed that for the local users (Worker) there was a significant main

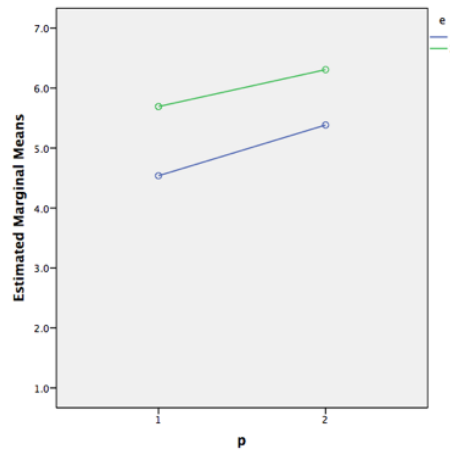


Figure 5.40: Expressed Clearly: Interaction between the conditions (Helper)

effect of both ($F(1, 12)=11.766, p=.005$) and EYETRACKING ($F(1, 12)=9.946, p=.008$) on the information provided by Helper while using these conditions. No significant interaction was found between POINTER and EYETRACKING ($F(1, 12) = 4.206, p=0.063$). Descriptive statistics (see Figure 5.41 and Table 5.24) show that NONE condition ($Mean=4.846, std. dev.=1.0682$) is rated comparatively lower than P condition ($Mean=6.154, std. dev.=0.6887$) and ET condition ($Mean=5.846, std. dev.=0.5547$). The BOTH condition ($Mean = 6.462, std. dev. = 0.7763$) is rated higher than the other three conditions.

Condition	Mean	Std. Deviation	N
NONE	4.846	1.0682	13
ET	5.846	.5547	13
P	6.154	.8987	13
BOTH	6.231	.5991	13

Table 5.24: Mean: information from partner (out of 7) for different conditions (Worker)

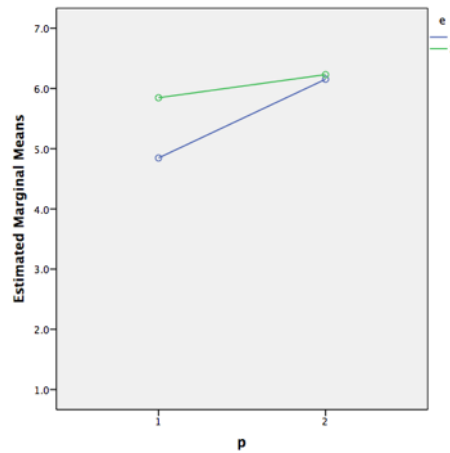


Figure 5.41: Information from partner: Interaction between the conditions (Worker)

Remote User: Helper

A repeated measure two-way ANOVA with ART revealed that for the local users (Worker) shows that there was a significant main effect of both POINTER ($F(1, 12)=11.766, p=.005$) and EYETRACKING ($F(1, 12)=9.946, p=.008$) on the information provided by Helper while using these conditions. No significant interaction was found between POINTER and EYETRACKING ($F(1,12) = 4.206, p=0.063$). Descriptive statistics (see Figure 5.42 and Table 5.25) show that NONE condition ($Mean=4.846, std. dev.=1.0682$) is rated comparatively lower than P condition ($Mean=6.154, std. dev.=0.8987$) and ET condition ($Mean=5.846, std. dev.=0.5547$). And the BOTH condition ($Mean = 6.231, std. dev. = 0.5591$) is rated higher than the other three conditions.

Condition	Mean	Std. Deviation	N
NONE	4.846	1.0682	13
ET	5.846	0.5547	13

Condition	Mean	Std. Deviation	N
P	6.154	0.8987	13
BOTH	6.231	.5591	13

Table 5.25: Mean: information from partner (out of 7) for different conditions (Helper)

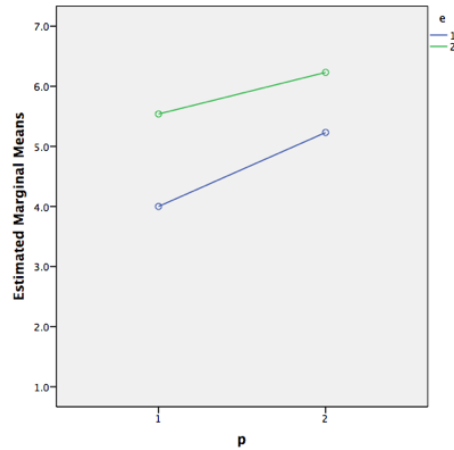


Figure 5.42: Information from partner: Interaction between the conditions (Helper)

Ranking Questionnaire

After participants tried all of the experimental conditions, we asked them to rank the four conditions as per the question statements. To determine the significant differences in the ranks from the responses of participants for different questions, a Friedman Test with post-hoc analysis using Wilcoxon signed-rank with Bonferroni correction was conducted. The significant level of Friedman test was 0.05 whereas for Wilcoxon test with Bonferroni correction ($\alpha = 0.05/6 = 0.0083$). Table 5.26 shows the list of the ranking questions. In the rest of this section, we report the results on ranking for each question.

No.	Question
1	Which condition was best at helping you to enjoy the task?
2	Which condition was best at making you feel connected with your partner?
3	Which condition was best at helping you stay focused on assembling the model?
4	Which condition was best at making you feel that you were present with your partner at same workspace while performing the task?
5	Which condition was best for you to tell that your partner needed assistance/ your partner knew that you needed assistance?
6	Which condition was best at helping you understand what your partner was communicating to you?

Table 5.26: Ranking Questions

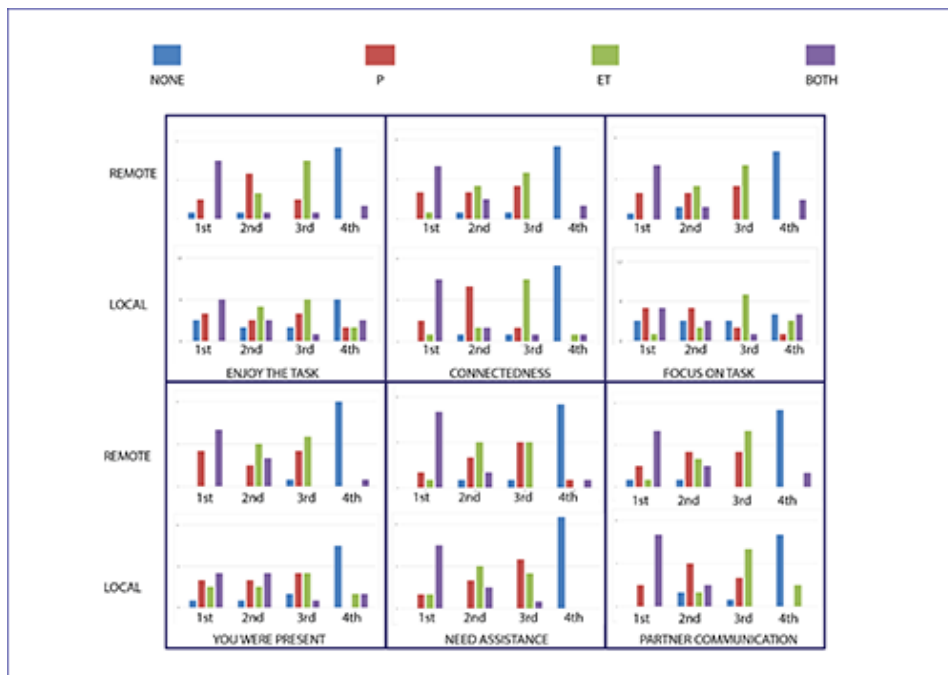


Figure 5.43: Results of Ranking under different aspects

Q1. Which condition was best at helping you to enjoy the task?

Local User: Worker

Condition	Mean	Std. Deviation	N
NONE	2.154	1.2810	13
P	2.692	1.1094	13
ET	2.231	.7250	13
BOTH	2.923	1.2558	13

Table 5.27: Mean: Task Enjoyment for different conditions (Worker)

There was no statistically significant difference in the ranking question of task enjoyment between conditions from the participant's response ($\chi^2(3) = 3.185, p = 0.364$).

Remote User: Helper

Condition	Mean	Std. Deviation	N
NONE	1.385	.9608	13
P	3.000	.7071	13
ET	2.308	.4804	13
BOTH	3.308	1.1821	13

Table 5.28: Mean: Task Enjoyment for different conditions (Helper)

There was a statistically significant difference in ranking in terms of the task enjoyed ($\chi^2(3) = 17.031, p = 0.001$). Median (IQR) perceived effort levels for NONE, P, ET, and BOTH con-

ditions were 1 (1 to 1), 3 (2.5 to 3.5), 2 (2 to 3) and 4 (2.5 to 4), respectively. Post-hoc analysis with Wilcoxon signed-rank tests with Bonferroni correction applied. The results showed that there was a significant difference between P and NONE ($Z = -3.022, p = 0.003$), whereas no significant differences between the rest of the pairs: ET and NONE ($Z = -2.166, p = 0.030$), BOTH and NONE ($Z = -2.543, p = 0.011$), ET and P ($Z = -2.066, p = 0.039$), BOTH and P ($Z = -0.608, p = 0.543$), BOTH and ET ($Z = -2.409, p = 0.016$).

Q2. Which condition was best at making you feel connected with your partner?

Local User: Worker

Condition	Mean	Std. Deviation	N
NONE	1.231	.5991	13
P	3.154	.6887	13
ET	2.231	.7250	13
BOTH	3.462	.9674	13

Table 5.29: Mean: Felt connected with partner for different conditions (Worker)

There was a statistically significant difference between the conditions in terms of ranking based on the feeling of being connected with the partner ($\chi^2(3) = 22.907, p < 0.001$). Median (IQR) perceived effort levels for each condition were *NONE*: 1 (1 to 1), *P*: 3 (3 to 4), *ET*: 2 (2 to 2.5) and *BOTH*: 4 (3 to 4). Post-hoc analysis with Wilcoxon signed-rank tests with Bonferroni correction showed there were significant differences between P and NONE ($Z = -3.270, p = 0.001$), and BOTH and NONE (Z

= -2.217, $p = 0.002$), whereas no significant differences found between the other pairs: ET and NONE ($Z = -2.543$, $p = 0.011$), ET and P ($Z = -2.217$, $p = 0.027$), BOTH and P ($Z = -0.988$, $p = 0.323$), and BOTH and ET ($Z = -2.476$, $p = 0.013$).

Remote User: Helper

Condition	Mean	Std. Deviation	N
NONE	1.231	.5991	13
P	2.923	.8623	13
ET	2.538	.6602	13
BOTH	3.308	1.1094	13

Table 5.30: Mean: Felt connected with partner for different conditions (Helper)

There was a statistically significant difference the rankings of how well each condition helped the helper feel connected with their partner. ($\chi^2(3) = 19.062$, $p < 0.001$). Median (IQR) perceived effort levels for None (No pointer, No Eye-Tracker), only Point, Only Eye-Tracker, Both (Point and Eye-Tracker) were 1 (1 to 1), 3 (2 to 4), 2 (2 to 3) and 4 (3 to 4), respectively. Post-hoc analysis with Wilcoxon signed-rank tests with Bonferroni correction showed that there were significant differences between NONE and the other three conditions (P: $Z = -3.236$, $p = 0.001$; ET: $Z = -2.951$, $p = 0.003$; BOTH: $Z = -2.951$, $p = 0.003$), whereas there were no significant differences between the rest of the pairs: ET and P ($Z = -1.020$, $p = 0.308$), BOTH and P ($Z = -0.892$, $p = 0.372$), or BOTH and ET ($Z = -1.842$, $p = 0.066$).

Q3. Which condition was best at helping you stay fo-

cused on assembling the model?

Local User: Worker

Condition	Mean	Std. Deviation	N
NONE	2.385	1.1929	13
P	3.077	.9541	13
ET	2.077	.8623	13
BOTH	2.692	1.3156	13

Table 5.31: Mean: Stay focused on assembling the model for different conditions (Worker)

There was no statistically significant difference in ranking based on the focus of Worker on assembling the model in a remote collaboration situation ($\chi^2(3) = 4.302, p = 0.231$).

Remote User: Helper

Condition	Mean	Std. Deviation	N
NONE	1.538	1.0500	13
P	2.923	.8623	13
ET	2.385	.5064	13
BOTH	3.154	1.2810	13

Table 5.32: Mean: Stay focused on assembling the model for different conditions (Helper)

There was a statistically significant difference in the average ranking of how well each condition helped the Helper stay focused on the assembly task ($\chi^2(3) = 12.046, p = 0.007$).

Median (IQR) perceived effort levels for None (No pointer, No Eye-Tracker), only Point, Only Eye-Tracker, Both (Point and Eye-Tracker) were 1 (1 to 2), 3 (2 to 4), 2 (2 to 3) and 4 (2 to 4), respectively. Post-hoc analysis with Wilcoxon signed-rank tests with Bonferroni correction showed that there was a significant difference between P and NONE ($Z = -2.946, p = 0.003$), whereas there were no significant differences between the rest of the pairs: ET and NONE ($Z = -1.942, p = 0.052$), BOTH and NONE ($Z = -2.294, p = 0.022$), ET and P ($Z = -1.493, p = 0.135$), BOTH and P ($Z = -0.427, p = 0.670$), BOTH and ET ($Z = -2.140, p = 0.032$).

Q4. Which condition was best at making you feel that you were present with your partner at same workspace while performing the task?

Local User: Worker

Condition	Mean	Std. Deviation	N
NONE	1.769	1.1658	13
P	3.000	.9129	13
ET	2.538	1.0500	13
BOTH	3.000	1.0801	13

Table 5.33: Mean: You were present with your partner for different conditions (Worker)

There was no statistically significant difference in ranking based on making user feel that she was present with her partner at the same workspace while performing the task ($\chi^2(3) = 7.031, p = 0.071$).

Remote User: Helper

Condition	Mean	Std. Deviation	N
NONE	1.077	.2774	13
P	3.000	.9129	13
ET	2.462	.5189	13
BOTH	3.462	.8771	13

Table 5.34: Mean: You were present with your partner for different conditions (Helper)

There was a statistically significant difference in the average rankings of how well each condition made the Helper feel like they were present with their partner in the same workspace. ($\chi^2(3) = 24.969, p < 0.001$). Median (IQR) perceived effort levels for None (No pointer, No Eye-Tracker), only Point, Only Eye-Tracker, Both (Point and Eye-Tracker) were *1 (1 to 1)*, *3 (2 to 4)*, *2 (2 to 3)* and *4 (3 to 4)*, respectively. Post-hoc analysis with Wilcoxon signed-rank tests with Bonferroni correction showed that there was a significant difference between P and NONE ($Z = -3.219, p = 0.001$), ET and NONE ($Z = -3.286, p = 0.001$), and BOTH and NONE ($Z = -3.203, p = 0.001$), whereas there were no significant differences between the rest of the pairs: ET and P ($Z = -1.493, p = 0.135$), BOTH and P ($Z = -1.181, p = 0.238$), BOTH and ET ($Z = -2.476, p = 0.013$).

Q5. Which condition was best for you to tell that your partner knew that you needed assistance?

Local User: Worker

Condition	Mean	Std. Deviation	N
NONE	1.231	.8321	13
P	2.615	.7679	13
ET	2.769	.7250	13
BOTH	3.615	.6504	13

Table 5.35: Mean: Needed assistance while performing tasks for different conditions (Worker)

There was a statistically significant difference in the conditions that could tell that her partner knew that she needed assistance, ($\chi^2(3) = 22.256, p < 0.001$). Median (IQR) perceived effort levels for each condition were *NONE*: 1 (1 to 1), *P*: 2 (2 to 3), *ET*: 3 (2 to 3) and *BOTH*: 4 (3 to 4). Post-hoc analysis with Wilcoxon signed-rank tests with Bonferroni correction showed there were significant differences between P and NONE ($Z = -3.140, p = 0.002$), ET and NONE ($Z = -2.676, p = 0.007$), and BOTH and NONE ($Z = -3.196, p = 0.001$), whereas no significant differences found between the other pairs: ET and P ($Z = -0.369, p = 0.712$), BOTH and P ($Z = -2.409, p = 0.016$), and BOTH and ET ($Z = -2.112, p = 0.035$).

Remote User: Helper

Condition	Mean	Std. Deviation	N
NONE	1.462	.9674	13
P	2.538	.8771	13
ET	2.615	.6504	13
BOTH	3.615	.8697	13

Table 5.36: Mean: Needed assistance while performing tasks for different conditions (Helper)

There was a statistically significant difference in the average ranking of the conditions in terms of how well the Helper could tell that their partner needed assistance, ($\chi^2(3) = 18.256, p < 0.001$). Median (IQR) perceived effort levels for None (No pointer, No Eye-Tracker), only Point, Only Eye-Tracker, Both (Point and Eye-Tracker) were 1 (1 to 1.5), 2 (2 to 3), 3 (2 to 3) and 4 (3.5 to 4), respectively. Post-hoc analysis with Wilcoxon signed-rank tests with Bonferroni correction showed that there was a significant difference between , ET and NONE ($Z = -2.697, p = 0.007$), and BOTH and NONE ($Z = -3.007, p = 0.003$), whereas there were no significant differences between the rest of the pairs: P and NONE ($Z = -2.254, p = 0.024$), ET and P ($Z = -0.037, p = 0.971$), BOTH and P ($Z = -1.181, p = 0.238$), BOTH and ET ($Z = -2.586, p = 0.010$).

Q6. Which condition was best at helping you understand what your partner was communicating to you?

Local User: Worker

Condition	Mean	Std. Deviation	N
NONE	1.538	.8771	13
P	2.923	.7596	13
ET	1.923	.6405	13
BOTH	3.769	.4385	13

Table 5.37: Mean: Understood Partner's communication for different conditions (Worker)

There was a statistically significant difference in the condition that helped her in understanding what her partner was communicating with her, ($\chi^2(3) = 24.535, p < 0.001$). Median

(IQR) perceived effort levels for each condition were *NONE*: 1 (1 to 2.5), *P*: 3 (2 to 3.5), *ET*: 2 (1.5 to 2) and *BOTH*: 4 (3.5 to 4). Post-hoc analysis with Wilcoxon signed-rank tests with Bonferroni correction showed there were significant differences between *P* and *NONE* ($Z = -2.738, p = 0.006$), *BOTH* and *NONE* ($Z = -3.134, p = 0.002$), and *BOTH* and *ET* ($Z = -3.223, p = 0.001$), whereas no significant differences were found between the other pairs: *ET* and *P* ($Z = -2.221, p = 0.026$), *ET* and *NONE* ($Z = -0.910, p = 0.363$), and *BOTH* and *P* ($Z = -2.221, p = 0.026$).

Remote User: Helper

Condition	Mean	Std. Deviation	N
NONE	1.385	.9608	13
P	2.846	.8006	13
ET	2.462	.6602	13
BOTH	3.308	1.1094	13

Table 5.38: Mean: Understood Partner's communication for different conditions (Helper)

There was a statistically significant difference in the average rankings of conditions in terms of which was best at helping the Helper understand what their partner was communicating., ($\chi^2(3) = 15.738, p = 0.001$). Median (IQR) perceived effort levels for None (No pointer, No Eye-Tracker), only Point, Only Eye-Tracker, Both (Point and Eye-Tracker) were 1 (1 to 1), 3 (2 to 3.5), 2 (2 to 3) and 4 (3 to 4), respectively. Post-hoc analysis with Wilcoxon signed-rank tests with Bonferroni correction showed that there was a significant difference between , *P* and *NONE* ($Z = -2.961, p = 0.003$), whereas there were no

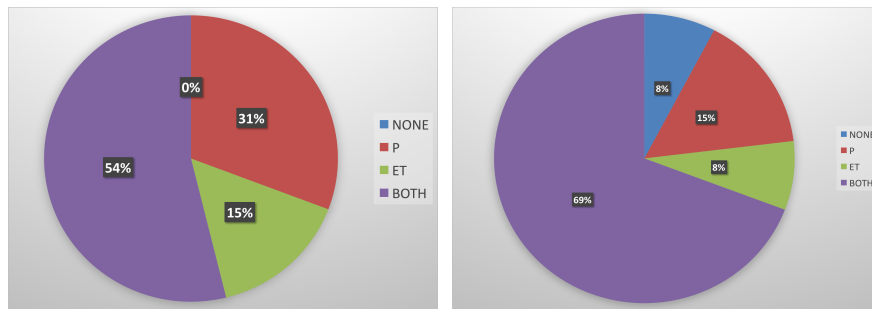
significant differences between the rest of the pairs: ET and NONE ($Z = -2.254, p = 0.024$), BOTH and NONE ($Z = -2.476, p = 0.013$), ET and P ($Z = -1.020, p = 0.308$), BOTH and P ($Z = -1.038, p = 0.299$), and BOTH and ET ($Z = -2.221, p = 0.026$).

Summary

After analysis of ranking questions, For the local user, we found that the BOTH condition (Worker sharing her eye gaze information with the Helper and the Helper using pointer annotation to assist the Worker) was ranked 1st for all the 6 questions, however P (Helper uses pointer annotation to assist Worker) was tied with BOTH for rank 1st for Q3 asking about the focus of participant on the task using these conditions. Whereas, NONE (no cue) was the worst ranked i.e. ranked 4th for all 6 questions except for question 3 where it tied for the worst rank with BOTH. BOTH was also ranked 1st for all the 6 questions for the remote user and NONE was ranked 4th for all the 6 questions.

5.6.3. Questionnaire: Qualitative Feedback

At the end of each condition, we asked participants to write down their views about the interface they had just used. They were asked to write about their understanding of the attention of their partner with each condition including the need of knowing the focus of attention. At the end we also asked about the condition they thought was better in improving the task performance.



(a) Local User (Worker)

(b) Remote User (Helper)

Figure 5.44: Summary of qualitative result: Focus of attention of partner

I could understand where my partner was focusing their attention.

As per the responses from Workers (Local User), almost 85% (11 participants) preferred the condition in which the pointer feature was available (i.e. P and BOTH) as it was helping them to understand their partner’s focus and provide clear instructions. One participant said: *“With Pointer, I can relate what he is talking about, because I could understand him more.”* Whereas, from the Helper’s (Remote User) responses, around 69% (9 participants) suggested to use BOTH condition since they were able to see the place where their partner was looking and use the pointer to help provide instruction, participant said: *“Eye Tracker help me to look in the same view of my partner, and I know what he is doing and will do next”*. However, 15% (2 participants) said that the focus of attention was not necessary, so they didn’t notice. One of them said: *“I don’t feel I need to know where he was looking at, the important was whether he understood where I want him to put the piece”*.

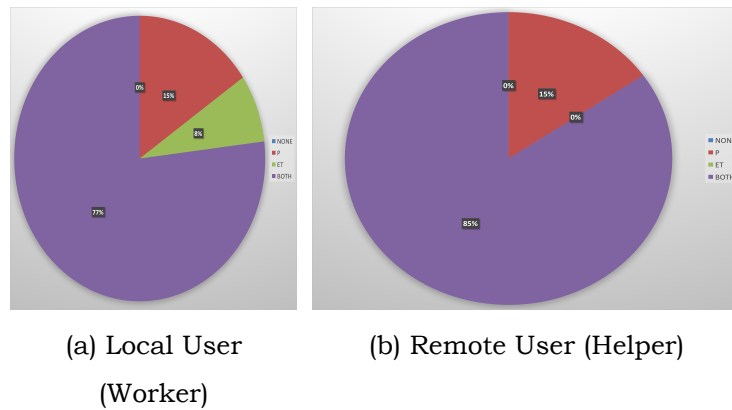


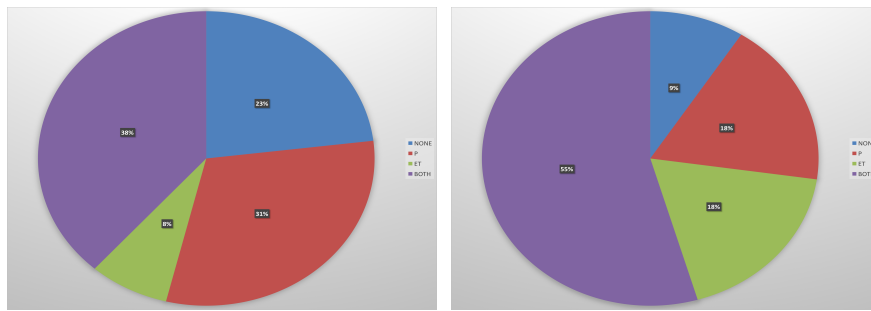
Figure 5.45: Summary of qualitative result: Task performance

Impact of these conditions on the task performance

On asking Workers (Local Users) about the condition that helped them more in performing the task efficiently and quickly, in terms of the information provided, *10 Workers (77%)* said that BOTH condition was better than the others. As one participant said, *“Eye tracker was giving my partner more information about where I looked at pointer was for giving me the instruction from my partner, where I should look at and which piece I should take”*. We got a similar response from the Helper (Remote User) as well, approximately *11 Helpers (85%)* said that BOTH condition was better than others. One participant said: *“To be able to give some form of visual feedback other than audio helps to perform the task together. Also to have some form of visual feedback from the partner, helps me gauge where my partner is looking. It helps when my partner is not talking at points in time.”*

Effect of Timer on performing task using different condition

The timer was meant to provide another point of focus in the task of constructing structures using LEGO pieces. We didn't



(a) Local User (Worker)

(b) Remote User (Helper)

Figure 5.46: Summary of qualitative result: Effect of Timer

use the number of times participants forgot to check the timer in our measurements, they didn't really focus on the timer as it was not the main task. In overall for local users (Worker), 9 (69%) participants said that providing Pointer cue (i.e. P and BOTH) by remote user (Helper) helped him in knowing when he has to press the button as most of the time their focus was on the HMD while performing the task and one of them said that *"if I keep looking at the display, with the help of his mouse click on the block, I can finish the task quickly"*. However, for remote users (Helpers), two of them (15%) didn't focus on the timer at all, as the timer was out of the camera view due to the excessive head movement, but 8 (73%) suggested that eye tracking information was very useful to keep track of the timer since he knows whether the local user is paying attention to the timer or not, so that he can focus only on task and if local user is not paying attention he can ask him to check the timer regularly. When needed about the relation of timer and task performance, one user said that *"I know that she was checking the timer regularly, so I focused only on instructing her about the next move."*

5.7. Discussion

In this section we discuss the results from the user experiment and the lessons learned.

From the experiment analysis, we found that the time taken by the pair of subjects to complete the task was significantly faster in the Both (Eye Tracker & Pointer) condition than using only Pointer or only Eye Tracker or Video only conditions. Whereas, There is a significant main effect of both pointer and eye tracking cues and through mean comparison of time taken to complete the task, Video only (NONE) condition was slowest among all. In the Both condition, the visual cues provided to both users helped them to perform the task more quickly. The Helper was able to use the virtual pointer to give direct guidance to the local Worker; as said by one local Worker, “The information provided by my partner in this condition helped in easily performing the task”. This enabled them to work significantly better than without pointer marker. Similarly, the eye tracker on the local Worker showed their focus of attention to the Helper, making it much easier for them to instruct the Worker quickly in the eye tracker condition as compared to without eye tracker condition.

The use of the pointer and eye-tracker cues enabled them to communicate more effectively. Since the Worker can see the Helper pointer marker directly on his or her HMD, the Helper able to use more deictic oriented language. For example, instead of using “There is a RED 4 BY 2 RECTANGULAR THICK BLOCK near the GREEN COLORED FLOWER. PICK that.”, the Helper could point at the block and just say “PICK THIS BLOCK”. Similarly, with the eye-tracking feed-

back, when Worker was searching for a blue block described by Helper, the Helper could reference the block as “Yes THAT BLUE BLOCK at which YOU JUST LOOKED”. These kind of cues helped in reducing the time taken by the Helper to understand what the Worker was doing.

The questionnaire results support the performance time results, showing a significant difference in the results to the questions “understanding what my partner was communicating”, “I was able to express myself clearly” for both Workers and Helpers with eye tracking or pointer factor and for “worked together” for Worker with eye tracking factor and Helper with pointer factor as compared to without eye tracking or without pointer factors.

The condition without pointer or eye tracker cues (None) performed worst, due to the lack of any visual assistance from Helper. In the Eye Tracking condition the experience for the Worker was the same as in the None condition; the instructions from Helper to Worker were only in verbal form. Interestingly, in this case the performance was significantly better than the None condition, which shows the benefit of making the Helper aware of the Worker’s focus of attention. We provided a separate point to focus that was meant for diverting participant’s attention from only one point, that was an attempt to simulate real world applications. One Helper explained a benefit of knowing the status of focus of attention of Worker as “I know that she was checking the timer regularly, so I focused only on instructing her about the next move.”. This feature was not present in the None condition and that made it very difficult to coordinate with pressing button of the timer on time and working on the task simultaneously.

It can also be seen from the questionnaire result that the remote user enjoyed the task more if the eye tracking information, i.e. focus of attention status of Worker was provided to him as compared to the conditions in which eye tracking information was not provided (NONE and P). Whereas, it can also be seen that if the Worker knows that his partner can see his focus of attention while performing the tasks i.e. while using ET and BOTH conditions, he feel that they worked together better as compared to the conditions without EYETRACKING (NONE and P) as the combined mean rating of EYETRACKING was reaching the highest possible value and higher than without EYETRACKING conditions.

We asked questions in our per-task questionnaire (Likert Scale) and post-task questionnaire (Ranking) regarding the copresence between Worker and Helper, including questions such as “I felt connected with my partner”, “I felt that I was present with my partner with my partner on the same workspace while performing the task”, “My partner was able to sense that I was present with him on the same workspace while performing the task”, “Which condition was best at making you feel connected with your partner” and “Which condition was best at making you feel that you were present with your partner at same workspace while performing the task”. From the responses of the Workers and Helpers we observed that copresence was higher when the Worker’s eye tracking data was sent to Helper, and/or a pointer marker was used by Helper. tFor example, according to one Worker, “I know that Helper sensed what I am going to do and to correct me he pointed on the correct Block”.

This was giving the Worker a sense that she is present with

the Helper on the same workspace as Helper knew about her focus of attention. On asking a Helper about the difference she found between the sharing of task space video in which the perspective of the video is like it is shot directly from the eyes of the Worker, and normal scene video where a camera is in front of the Worker showing the Worker's body and the task space from an angle, she said "Obviously, If I can see from his perspective, how he is looking at the task space, how he is using the blocks , then my instructions are clearer to him as compared to later video case where I have to think about the instructions. I have to say pick the object on your right where I have to make sure I am talking about his right, i.e. my left in case of video from the front. It's confusing. Best would be to have a shared gaze, where I can see that he is looking at the correct direction." She also suggested to use her focus of attention to show the next instruction to the Worker on his HMD, as if we can provide direct looking at the object instead of pointing, it can be much faster to complete the task and might increase the presence between both. We will try to explore this technique of instructing in our future work.

However, a few Helpers complained that "the eye tracker information is misleading as whenever I point on a particular block, he didn't look at it. That is annoying, why he is looking at some where else and not on the block even when I am constantly saying that block, look at the pointer". The reason behind this was the use of a monocular optical see-through HMD. The display is on left eye side and so whenever the Helper points to an object, the Worker looks at the display that switches his focus from real world to the HMD. This affects the eye tracker since it was calibrated for the real world workspace and not for the image viewed on the HMD. The so-

lution to this problem could be the use of a binocular video see-through HMD and by providing only one workspace.

5.7.1. Implications of the Research

Based on the results and discussion of the research, we are proposing implicated design guidelines for people who will develop Head Mounted eye tracker based remote collaborate systems.

- Keep the head mounted system light in weight. From our observation during the experiments and participant's comments, the HMD was a bit heavy and uncomfortable at nose and forehead of the user.
- Calibration of the eye tracker is the most important factor to provide reliable eye gaze information. So, HMD should be robust enough that even in the physical tasks where user has to move his head freely, HMD should not dislocate otherwise, it will affect the calibration and will start giving bogus data.
- Always provide a method for the Helper to visually communicate with the Worker on spatial tasks.

5.7.2. Limitations

Although the experimental results are very interesting, there are a number of limitations that should be address in future research, including:

- The prototype was heavy because of using cameras, HMD and eye tracker all mounted on the same frame. In future

prototypes, we could develop more ergonomic systems with a light weight camera and unobtrusive eye tracker.

- The prototype was not stable on the user's head and sometimes moved around. Although, we used headbands for stabilization, in the future we could develop a more stable system.
- The eye tracker was affected by HMD movement in response to head motion, and could move from the initial position at which it was calibrated. Once the HMD moves too much, it will not show the eye gaze marker on the actual eye pupil position mapped on the real world.
- The task was not completely ideal for this study as the Helper could give clear description of the blocks, not using the features of the prototype (e.g the mouse pointer) and successfully complete the task. In the future, we will try to explore more complex tasks with that cannot be completed easily with voice cues alone
- We tried to use a multitasking activity by providing the timer that needed to be monitors. However, this was not completely successful as participants didn't feel any need to keep track of the timer as it was not affecting the task. On asking participants why they didn't notice the timer, one participant said that since the timer was not useful in constructing the task it was just an annoying diversion. In the future, we will try to use another tool to provide a separate point for focus
- The color of pointer marker and eye gaze marker was red and pink respectively which were similar and was

sometimes difficult for the Worker to distinguish between both. We need to change the color of the markers.

- We used monocular display instead of stereo display, so it was very difficult to overlay virtual cues at the same focal point as the LEGO blocks being assembled. So in future, we will use a different wearable display.
- In our experimental measures we took pointer and eye tracker as independent variables which was not good for design since we were focusing more on exploring the benefit of having focus of attention information in remote collaboration. So pointer was not solving any purpose in that. Also the evaluation methods e.g. questionnaire and surveys were not a standard questionnaire, so in future we will try to use a better experimental design with standard questionnaire. Also to measure user's workload while performing the task was not measured, so in future we will use surveys like NASA TLX survey or other experimental tools for this.

5.8. Conclusions

The main objective of our evaluation was to determine the effect of adding eye tracking information in a remote assistance task, in terms of copresence and performance time.. We developed an eye tracking tool that can be mounted on already existing HMD and prototyped the whole system to use it with our system.

At the start of the chapter, we stated two main hypotheses that we wanted to investigate in the user study.

Overall in terms of connectedness, presence, needing / needed assistance, we found there was a significant main effect in Pointer and Eye Tracking factors most of the time with a significant interaction of Pointer and Eye tracking on each other. This leads to a conclusion of supporting H1.

From assessing the task completion time, we found both Pointer and Eye Tracker factors had a significant impact on the task performance for the given task. However, there was no interaction between the two factors. The participants' task performance was the best in BOTH condition compared to the other conditions. Based on this findings we conclude that the study results support our second Hypothesis H2.

6

Conclusion and Future Work

6.1. Conclusion

The interpersonal and consistent contact of wearable computers with the human body makes them a suitable technology for recognizing and sharing of emotional states with a remote person. There are various modes that can be used to represent the shared emotional states such as visual, audio, haptic feedbacks etc. With the increasing use of HMDs and wearable devices with cameras, visual and audio modes could be used to share emotion. However, for unobtrusive wearable devices such as watches, apparel, and shoes, etc. could be used.

In this thesis, the main aim of our research was to contribute to the field of affective communication and remote collaboration by exploring various modes of sharing emotional states and creating connectedness by using wearable devices. We achieved this by designing two different prototypes to in-

investigate different modes for affective communication with a remote person, and one prototype investigating the role that focus of attention plays in remote collaboration.

For the first prototype system, CSense, after needs analysis, brainstorming, and trying several alternative solutions, we developed a wrist band that uses vibration patterns and band tightening to share emotional states with a remote user. From pilot testing, we found that active and passive emotions could be differentiated when using the system. Since the patterns of vibrations was not properly detected as specific emotions, in the future we would like to think of some different patterns of vibrations that can be explored to make it a robust system for sharing emotions.

In the next prototype, CoSense, we developed a system that combined video conferencing with sharing the emotional states of the user that was recognized using physiological sensors. A formal user study was conducted with four different interface conditions; (1) just video, (2) video with color overlay, (3) video with color overlay and emotional state as text along with heart rate, and (4) video with color overlay and emotional state as text along with heart rate and other physiological sensor data in graphs. We found that interface combining video with simple emotional cues enabled the remote user to empathize better than the other conditions. This was because it used only the most appropriate information, rather than no information, or too much information (e.g showing the raw sensor values).

Finally, we know that remote collaboration using video conferencing is one of the common tools that enable remote people to work together and information. Wearable devices with

HMDs, like Google Glass, etc. that have a camera, microphone and display, can be used to assist workers by overlaying annotations on the object they want to manipulate. So we developed a third prototype in order to see if showing the focus of attention of the worker makes any impact on the co-presence between the worker and helper. We found that the performance, connectedness and presence was increased when eye tracking information from the worker was shared with the helper, and the helper can show pointer annotations on the worker's HMD.

6.2. Future Work

These three prototypes explored some basic but important input modes, i.e. vibration feedback haptic, contraction and expansion feedback, color as visual feedback, gaze tracking and audio feedback. They showed some preliminary positive outcomes regarding the use of wearable devices for affective remote collaboration. However there is future work that could be done to improve the existing interface system, and the experiments conducted.

The current studies used a very small sample size, so in the future we would like to conduct a study with larger sample size.

The LEGO construction task was not ideal for the experiment related to focus of attention. Also the experimental design and evaluation measures, so in the future we would like to conduct a study with better task to perform, better experiment design and better evaluation measures.

We will also investigate the possibility of providing eye tracking information i.e. focus of attention as an annotation tool for remote collaboration.

From the third prototypes, we found that use of eye tracking to share the focus of attention enhances copresence, and colored overlay and simple emotion representation enhances empathy between the local and remote users. In the future, we would like to explore the level of copresence and empathy created when presenting both focus of attention and emotional states are merged in one interface. For example, the color of the eye gaze marker could change according to the user's emotional state (see figure 6.1).



Figure 6.1: Idea concept image

We would also like to research where a person can tag specific areas around her with emotional tags by looking at those areas and share it with other people has a great potential in well-being of the society. This could be used by paralyzed people who want to express their emotions but cannot, or by autistic non-verbal kids as with this system they can tag their surroundings according to the things they like and the things

that annoy them and it can be shared with their parents or concerned people.

We would also like to investigate the use of different cues other than color to represent the focus of attention or emotional state. For example, focus of attention with different spatialized audio nodes assigned for specific emotional states. We are also interested in developing a language of vibration patterns for different emotional states, and using haptic feedback to represent the focus of attention along with emotional states.

Lastly, the increase in the range and variety of wearable devices, provides a scope for using them for sharing of emotional states with verbal or non-verbal cues. We will try to redesign our system such that we can use these new devices as well. In the future we could use a mesh of wearable devices that on combining the available sensory data from those devices, can provide a platform to more accurately detect and share the emotional states.

Bibliography

- [1] Simon Guigui aka fyhertz. spydroid-ipcamera: A powerful and funny android app. that streams the camera and microphone of your phone to your browser or to vlc.
- [2] Arduino. Arduino development board.
- [3] David G Armstrong, Timothy M Rankin, Nicholas A Giovinco, Joseph L Mills, and Yoky Matsuoka. A heads-up display for diabetic limb salvage surgery a view through the google looking glass. *Journal of diabetes science and technology*, page 1932296814535561, 2014.
- [4] Christopher Heine at Adweek. Google glass: What it's like to have 'an iphone attached to your face'.
- [5] MacNN Staff at MacNN. Brother upgrades airscouter head mounted display.
- [6] Charlie Sorrel at Wired. Nike+ heart-rate monitor on sale june 1st.
- [7] Matt Morgan at Wired. Review: Does the vuzix wrap 1200 take tv viewing into the future?
- [8] Phandroid at Youtube. Looking through google glass: Real life example.
- [9] Sudhanshu SDP Ayyagari, Kunal Gupta, Matt Tait, and Mark Billingham. Cosense: Creating shared emotional

- experiences. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pages 2007–2012. ACM, 2015.
- [10] Jeremy N Bailenson, Nick Yee, Scott Brave, Dan Merget, and David Koslow. Virtual interpersonal touch: expressing and recognizing emotions through haptic devices. *Human-Computer Interaction*, 22(3):325–353, 2007.
- [11] Maha Bali. What is empathy?
- [12] Martin Bauer, Gerd Kortuem, and Zary Segall. "where are you pointing at?" a study of remote collaboration in a wearable videoconference system. In *ISWC*, volume 99, page 151, 1999.
- [13] Jimmy Bike. Glass rear view.
- [14] BITalino. Bitalino | diy biosignals.
- [15] Leonardo Bonanni, Cati Vaucelle, Jeff Lieberman, and Orit Zuckerman. Taptap: a haptic wearable for asynchronous distributed touch therapy. In *CHI'06 extended abstracts on Human factors in computing systems*, pages 580–585. ACM, 2006.
- [16] Scott Brave and Andrew Dahley. intouch: a medium for haptic interpersonal communication. In *CHI'97 Extended Abstracts on Human Factors in Computing Systems*, pages 363–364. ACM, 1997.
- [17] Susan E Brennan, Xin Chen, Christopher A Dickinson, Mark B Neider, and Gregory J Zelinsky. Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*, 106(3):1465–1477, 2008.

- [18] Brother. Brother airscouter.
- [19] Carlos Busso, Zhigang Deng, Serdar Yildirim, Murtaza Bulut, Chul Min Lee, Abe Kazemzadeh, Sungbok Lee, Ulrich Neumann, and Shrikanth Narayanan. Analysis of emotion recognition using facial expressions, speech and multimodal information. In *Proceedings of the 6th international conference on Multimodal interfaces*, pages 205–211. ACM, 2004.
- [20] WearScript by Drapper Vision Inc. Wearsript: Eye tracking.
- [21] Walter B Cannon. The james-lange theory of emotions: A critical examination and an alternative theory. *The American Journal of Psychology*, pages 106–124, 1927.
- [22] Angela Chang, Sile O’Modhrain, Rob Jacob, Eric Gunther, and Hiroshi Ishii. Comtouch: design of a vibrotactile communication device. In *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques*, pages 312–320. ACM, 2002.
- [23] Angela Chang, Ben Resner, Brad Koerner, XingChen Wang, and Hiroshi Ishii. Lumitouch: an emotional communication device. In *CHI’01 extended abstracts on Human factors in computing systems*, pages 313–314. ACM, 2001.
- [24] Manfred Clynes and Yehudi Menuhin. *Sentics: The touch of emotions*. Anchor Press Garden City, NY, 1977.
- [25] cooking hacks. e-health sensor platform complete kit v2.0 for arduino, raspberry pi.

- [26] Claudia Cortes. Claudia meaning of color.
- [27] Beatrice de Gelder, Koen BE Böcker, Jyrki Tuomainen, Menno Hensen, and Jean Vroomen. The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. *Neuroscience letters*, 260(2):133–136, 1999.
- [28] Dictionary. Empathy.
- [29] Carl DiSalvo, Francine Gemperle, Jodi Forlizzi, and Elliott Montgomery. The hug: an exploration of robotic form for intimate communication. In *Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003. The 12th IEEE International Workshop on*, pages 403–408. IEEE, 2003.
- [30] Vanessa Urch Druskat and Steven B Wolff. Building the emotional intelligence of groups. *Harvard Business Review*, 79(3):80–91, 2001.
- [31] Paul Ekman. Facial expression and emotion. *American psychologist*, 48(4):384, 1993.
- [32] Emotient. Emotient.
- [33] Empatica. Empatica: Human data in real life.
- [34] Florian Eyben, Martin Wöllmer, and Björn Schuller. Openear—introducing the munich open-source emotion and affect recognition toolkit. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pages 1–6. IEEE, 2009.
- [35] Tiffany Field. Touch for socioemotional and physical well-being: A review. *Developmental Review*, 30(4):367–383, 2010.

- [36] Britannica. Britannica Online for Kids. Communication - affective communication.
- [37] Fritzing. Fritzing.
- [38] Susan R Fussell, Robert E Kraut, and Jane Siegel. Coordination of communication: Effects of shared visual context on collaborative work. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, pages 21–30. ACM, 2000.
- [39] SUSAN R Fussell and LESLIE D Setlock. Using eye-tracking techniques to study collaboration on physical tasks: Implications for medical research. *Unpublished manuscript, Carnegie Mellon University*, 2003.
- [40] Susan R Fussell, Leslie D Setlock, and Robert E Kraut. Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 513–520. ACM, 2003.
- [41] Susan R Fussell, Leslie D Setlock, and Elizabeth M Parker. Where do helpers look?: Gaze targets during collaborative physical tasks. In *CHI'03 Extended Abstracts on Human Factors in Computing Systems*, pages 768–769. ACM, 2003.
- [42] Alberto Gallace and Charles Spence. The science of interpersonal touch: an overview. *Neuroscience & Biobehavioral Reviews*, 34(2):246–259, 2010.
- [43] Google. Google glass.
- [44] Antal Haans and WA Usselsteijn. The virtual midas

- touch: Helping behavior after a mediated social touch. *Haptics, IEEE Transactions on*, 2(3):136–140, 2009.
- [45] Jennifer Anne Healey. *Wearable and automotive systems for affect recognition from physiology*. PhD thesis, Massachusetts Institute of Technology, 2000.
- [46] Philips Healthcare. Experience the future of wearable technology - philips healthcare.
- [47] Javier Hernandez and Rosalind W Picard. Senseglass: Using google glass to sense daily emotions. In *Proceedings of the adjunct publication of the 27th annual ACM symposium on User interface software and technology*, pages 77–78. ACM, 2014.
- [48] Fraunhofer IIS. Shore: Object and face recognition using google glass.
- [49] Invision. Invision app.
- [50] Ken Iwasaki, Takashi Miyaki, and Jun Rekimoto. Affectphone: A handset device to present user’s emotional state with warmth/coolness. In *B-Interface*, pages 83–88, 2010.
- [51] Moritz Kassner, William Patera, and Andreas Bulling. Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction. April 2014.
- [52] Jiro Kawakita. The original kj method. *Tokyo: Kawakita Research Institute*, 1991.
- [53] John F Kelley. *Cal-a natural language program developed with the oz paradigm: Implications for supercomputing systems*. IBM Thomas J. Watson Research Center, 1985.

- [54] Janice R Kelly and Sigal G Barsade. Mood and emotions in small groups and work teams. *Organizational behavior and human decision processes*, 86(1):99–130, 2001.
- [55] and Jun Rekimoto Ken Iwasaki, Takashi Miyaki. Affect-phone: A handset device to present user’s emotional state with warmth/coolness.
- [56] Seungwon Kim, Gun Lee, Nobuyasu Sakata, and Mark Billinghurst. Improving co-presence with augmented visual communication cues for sharing experience through video conference. In *Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on*, pages 83–92. IEEE, 2014.
- [57] Robert E Kraut, Darren Gergle, and Susan R Fussell. The use of visual information in shared visual spaces: Informing the development of virtual co-presence. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*, pages 31–40. ACM, 2002.
- [58] MIT Media Lab. Mit media lab.
- [59] Stephen RH Langton. The mutual influence of gaze and head orientation in the analysis of social attention direction. *The Quarterly Journal of Experimental Psychology: Section A*, 53(3):825–845, 2000.
- [60] Joseph LeDoux. *The emotional brain: The mysterious underpinnings of emotional life*. Simon and Schuster, 1998.
- [61] Clayton Lewis. *Using the” thinking-aloud” method in cognitive interface design*. IBM TJ Watson Research Center, 1982.

- [62] Steve Mann. *Wearable Computing*. The Interaction Design Foundation, Aarhus, Denmark, 2014.
- [63] MIT. Mit app inventor.
- [64] Albert H. Munsell. Munsell color space.
- [65] KAYA NAz and Helena Epps. Relationship between color and emotion: A study of college students. *College Student J*, 38(3):396, 2004.
- [66] Mark B Neider, Xin Chen, Christopher A Dickinson, Susan E Brennan, and Gregory J Zelinsky. Coordinating spatial referencing using shared gaze. *Psychonomic bulletin & review*, 17(5):718–724, 2010.
- [67] T Nettelbeck, C Henderson, and R Willson. Communicating emotion through sound: An evaluation of clynes’ theory of sentics. *Australian journal of psychology*, 41(1):25–36, 1989.
- [68] NeuroSky. Neurosky: Mindwave.
- [69] Joy Nicholson, Kazuhiko Takahashi, and Ryohei Nakatsu. Emotion recognition in speech using neural networks. *Neural computing & applications*, 9(4):290–296, 2000.
- [70] Niels A Nijdam. Mapping emotion to color. *(Eds.):‘Book Mapping emotion to color’(2009, edn.)*, pages 2–9, 2009.
- [71] Jiazhi Ou, Lui Min Oh, Susan R Fussell, Tal Blum, and Jie Yang. Predicting visual focus of attention from intention in remote collaborative tasks. *Multimedia, IEEE Transactions on*, 10(6):1034–1045, 2008.

- [72] Jiazhi Ou, Lui Min Oh, Jie Yang, and Susan R Fussell. Effects of task properties, partner actions, and message content on eye gaze patterns in a collaborative task. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 231–240. ACM, 2005.
- [73] Rosalind W Picard and Roalind Picard. *Affective computing*, volume 252. MIT press Cambridge, 1997.
- [74] Rosalind W Picard and Jocelyn Scheirer. The galvactivator: A glove that senses and communicates skin conductivity. In *Proceedings 9th Int. Conf. on HCI*, 2001.
- [75] Processing. Processing.
- [76] Bernard Rimé. Emotion elicits the social sharing of emotion: Theory and empirical review. *Emotion Review*, 1(1):60–85, 2009.
- [77] Bernard Rime, Batja Mesquita, Stefano Boca, and Pierre Philippot. Beyond the emotional event: Six studies on the social sharing of emotion. *Cognition & Emotion*, 5(5-6):435–465, 1991.
- [78] Fabien Ringeval, Andreas Sonderegger, Basilio Noris, Aude Billard, Jens Sauer, and Denis Lalanne. On the influence of emotional feedback on emotion awareness and gaze behavior. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*, pages 448–453. IEEE, 2013.
- [79] Jocelyn Scheirer, Raul Fernandez, and Rosalind W Picard. Expression glasses: a wearable device for facial expression recognition. In *CHI'99 Extended Abstracts on*

Human Factors in Computing Systems, pages 262–263. ACM, 1999.

- [80] SolidWorks. Solidworks.
- [81] Randy Stein and Susan E Brennan. Another person’s eye gaze as a cue in solving programming problems. In *Proceedings of the 6th international conference on Multimodal interfaces*, pages 9–15. ACM, 2004.
- [82] Cara A Stitzlein, Jane Li, and Alex Krumm-Heller. Gaze analysis in a remote collaborative setting. In *Proceedings of the 18th Australia conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments*, pages 417–420. ACM, 2006.
- [83] Google Play store. Skype.
- [84] Sseed Studio. Sseed studio.
- [85] Think-Dezine. User experience design.
- [86] Dzmitry Tsetserukou and Alena Neviarouskaya. ifeel_im!: augmenting emotions during online communication. *IEEE computer graphics and applications*, (5):72–80, 2010.
- [87] Dzmitry Tsetserukou and Alena Neviarouskaya. World’s first wearable humanoid robot that augments our emotions. In *Proceedings of the 1st Augmented Human International Conference*, page 8. ACM, 2010.
- [88] Twitch. Twitch.
- [89] Thurid Vogt, Elisabeth André, and Nikolaus Bee. Emovoice—a framework for online recognition of emo-

- tions from voice. In *Perception in multimodal dialogue systems*, pages 188–199. Springer, 2008.
- [90] Johann Wolfgang Von Goethe and Charles Lock Eastlake. *Theory of colours*. Number 3. Mit Press, 1970.
- [91] Vuzix. Vuzix: View the future.
- [92] Johannes Wagner, Florian Lingenfeller, and Elisabeth André. The social signal interpretation framework (ssi) for real time signal processing and recognition. In *INTERSPEECH*, pages 3245–3248, 2011.
- [93] Che-Wei Wang. Soft pneumatic exoskeleton.
- [94] Alden E Wessman and David F Ricks. Mood and personality. 1966.
- [95] Jacob O Wobbrock, Leah Findlater, Darren Gergle, and James J Higgins. The aligned rank transform for non-parametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 143–146. ACM, 2011.
- [96] Jeffrey Wong, Lui Min Oh, Jiazhi Ou, Carolyn P Rosé, Jie Yang, and Susan R Fussell. Sharing a single expert among multiple partners. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 261–270. ACM, 2007.

A

Appendix A: Interview Questions

A.1. Emotion sharing in context of sports

- Do you play any sport? What kind of Sport? Why?
- What makes you to feel to play this particular sport and not others?
- With whom do you like to play this sport? Why?
- What are the emotions that comes in your mind when you have to go and play your favorite sport? Why? Any particular story?
- How do you make yourself emotionally ready for the game?
- How do you share these emotions? Why?

- Imagine if your best friend or the person you likes to play with the most, didn't come. So How would you feel at that time? Would you like to share those emotions to him/her? How? Why?
- What would be your emotions after seeing strong opponents? Would you like to share these emotions? How?
- What sort of emotions you go through while playing the sport? Why?
- How do you share your these feelings while playing? Why?
- How do you share your emotions after the game? Why?
- When you are not able to cross any level or hurdle, how do you feel? Why? Do you share these feelings with your friends too? Why? How would you like to share?
- What is your views on Emotions Sharing process while playing sports? Why do you think so?

B

Appendix B: Information sheet and consent form



Human Interface Lab (HIT Lab) NZ
Telephone: +64 3 364 2349
Email: kunal.gupta@pg.canterbury.ac.nz
Date:

Comparative study for social presence using eye tracking Head Mounted Device with AR based feedback in remote collaboration

Information Sheet for Participants

You are invited to take part in a remote assistance collaboration research study using an eye tracking HMD. The purpose of this study is to identify the significance of using eye tracking HMD on worker (local user) to send his attention details to helper (remote user) in completing the assigned task in a remote assistance collaboration condition.

Your involvement in this project will be as a remote user or a local user to help in completing the task.

There will be four tasks for each condition and in the task, local user will be asked to construct a simple LEGO Duplo structure with remote user's assistance who has an experience of making that structure and with all the instructions set. There will be a timer with RESET button to create a multi-tasking environment. Data will be recorded in the form of questionnaires and time to complete the task. Also the you will be audio recorded. However your task workspace might be visually recorded, but face or any part of body which might reveal user's identity will not be recorded.

The study will follow the procedure outlined as below:

1. The participants (helper and worker) read and sign the informed consent form.
2. The participants answer to a questionnaire individually on demographic information, their previous experiences of remote assistance using any available technology and previous experiences using HMD.
3. The researcher explains the study setup and experimental task for the participants to perform during the study.
4. The participants performs the experimental tasks which may include:
 - a. Wearing the HMD and calibrating the eye tracker
 - b. Performing the given task, which may include constructing simple or complex structures using LEGO Duplo pieces with the assistance of other participant who will have the complete instruction set to perform the task with him.
 - c. Rate the social presence based on each condition by answering to a questionnaire.
5. The participants answer a questionnaire individually asking for feedback on the overall study.
6. As a follow-up to this investigation, you will be asked to answer few questions for a debriefing interview.

The whole procedure will take approximately 80-100 minutes.

Figure B.1: Information sheet



In the performance of the tasks and application of procedures the risks are minimal. The participants will be sitting and performing the given LEGO construction task, so there will be a very less physical movement. Hence, we do not expect any injury to come upon any participant.

You may receive a copy of the project results by contacting the researcher at the conclusion of the project.

Participation is voluntary and you have a right to withdraw at any stage without penalty. If you withdraw, I will remove information relating to you.

The results of the project may be published, but you may be assured of the complete confidentiality of data gathered in this investigation: your identity will not be made public without your prior consent. To ensure anonymity and confidentiality, we are not going to record your name anywhere in the data except your consent form and you will be given a participant number which will be your identity in the data then onwards. In publications, we will mainly report the results in an aggregate format (by reporting only combined results and never reporting individual ones). All questionnaire will be concealed, and no one other than the investigators will have access to them. The data will be kept securely for up to 5 years and will be destroyed after completion of the project. A thesis is a public document which will be available through the UC Library.

The project is being carried out as a requirement of Masters of Human Interface Technology (MHIT) degree by Kunal Gupta under the supervision of Dr. Christoph Bartneck (christoph.bartneck@canterbury.ac.nz), who can be contacted at kunal.gupta@pg.canterbury.ac.nz. He will be pleased to discuss any concerns you may have about participation in the project.

This project has been reviewed and approved by the University of Canterbury Human Ethics Committee, and participants should address any complaints to The Chair, Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz).

If you agree to participate in the study, you are asked to complete the consent form and you can take this information sheet with you when you leave.

Figure B.2: Information sheet

Human Interface Lab (HIT Lab) NZ
 Telephone: +64 3 364 2349
 Email: kunal.gupta@pg.canterbury.ac.nz
 Date:

Comparative study for social presence using eye tracking Head Mounted Device with AR based feedback in remote collaboration

Consent Form for Participants

I have been given a full explanation of this project and have has the opportunity to ask questions.

I understand what is required of me if I agree to take part in the research.

I understand that participation is voluntary and I may withdraw at any time without penalty. Withdrawal of participation will also include the withdrawal of any information I have provided should this remain practically achievable.

I understand that any information or opinions I provide will be kept confidential to the researcher and the administrators of the research project and that any published or reported results will not identify the participants. I understand that a thesis is a public document and will be available through the UC Library.

I understand that all the recordings and data collected for the study will be kept in locked and secure facilities and/or in password protected electronic form and will be destroyed after five years.

I understand the risks associated with taking part and how they will be managed.

I understand that I am able to receive a report on the findings of the study by contacting the researcher at the conclusion of the project.

I understand that I can contact the researcher Kunal Gupta (kunal.gupta@pg.canterbury.ac.nz) or supervisor Dr. Christoph Bartneck (christoph.bartneck@canterbury.ac.nz) for further information. If I have any complaints, I can contact the Chair of the University of Canterbury Human Ethics Committee, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz)

By signing below, I agree to participate in this researach project.

Participant (Print Name)	Signature	Date
--------------------------	-----------	------



Figure B.3: Consent Form

C

Appendix C: Poster for advertisement

PARTICIPANTS NEEDED FOR RESEARCH IN HUMAN COMPUTER INTERACTION at HIT LAB

We are looking for volunteers to take part in a study of
*remote assistance collaboration using
eye tracking Head Mount Displays (HMD)*.

You would be asked to:

- Construct something cool using LEGO or Help others constructing.
- Answer an anonymous questionnaire.

Study will usually take 80-100 minutes. In appreciation for your time, you will receive two *cafe voucher!*

For more information, or to volunteer for this study, please contact:

Kunal Gupta

0210 2363177

kunal.gupta@pg.canterbury.ac.nz

Study Reviewed by HEC, UC

For details, call/text/Email

Kunal:
021 02363177
kunal.gupta@pg.canterbury.ac.nz

For details, call/text/Email

Kunal:
021 02363177
kunal.gupta@pg.canterbury.ac.nz

For details, call/text/Email

Kunal:
021 02363177
kunal.gupta@pg.canterbury.ac.nz

For details, call/text/Email

Kunal:
021 02363177
kunal.gupta@pg.canterbury.ac.nz

For details, call/text/Email

Kunal:
021 02363177
kunal.gupta@pg.canterbury.ac.nz

For details, call/text/Email

Kunal:
021 02363177
kunal.gupta@pg.canterbury.ac.nz

D

Appendix D: Ethics Approval

HUMAN ETHICS COMMITTEE

Secretary, Lynda Griffioen
Email: human-ethics@canterbury.ac.nz

Ref: HEC 2015/37/LR

9 July 2015

Kunal Gupta
HITLab NZ
UNIVERSITY OF CANTERBURY

Dear Kunal

Thank you for forwarding your Human Ethics Committee Low Risk application for your research proposal “Comparative study for social presence using eye tracking HMD with augmented reality based feedback in remote collaboration”.

I am pleased to advise that this application has been reviewed and I confirm support of the Department’s approval for this project.

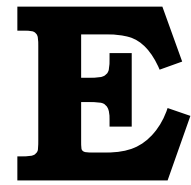
Please note that this approval is subject to the incorporation of the amendments you have provided in your email of 6 July 2015.

With best wishes for your project.

Yours sincerely



Lindsey MacDonald
Chair, Human Ethics Committee



Appendix E: Questionnaires

E.1. Pre-Task Questionnaire

6. How often **you ask someone to assist you** with performing a real world task over video conferencing?

- Never
- Less than once a Month
- Once a Month
- 2-3 Times a Month
- Once a Week
- 2-3 Times a Week
- Daily

7. What **kind of help do you ask for or give to your partner in video conferencing?**

8. Please describe any **special activity that you helped with over a video conference, e.g. working on a project together or any task which might need special attention to it among the other things, etc.**

E.2. Per-Task Questionnaire for Worker (Local User)

Questionnaire (Local) Condition: - Task:

Below are statements about your experience. **Please circle on a number** to indicate your level of agreement or disagreement with each statement

1. I **enjoyed** the experience.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

2. I **felt connected** with my partner.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

3. I was **able to focus** on the task activity (constructing structure)

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

4. I felt that **I was present** with my partner on the same workspace while performing the task

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

5. I think that **my partner was able to sense that I was present** with him on the same workspace while performing the task

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

P.T.O

Questionnaire Continues...

6. I felt that I was able to **express myself clearly** to my partner

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

7. I was able to **understand what my partner was communicating to me**

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

8. I am **confident that we completed the task** correctly

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

9. The **information provided by my partner in this condition helped in easily performing the task**

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

10. My partner and I **worked together well** on the task

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

11. My partner could tell **when I needed assistance**

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

Questionnaire for this condition ends.

E.3. Post-Task Questionnaire for Worker (Local User)

Questionnaire

(L)

After four conditions

Please rank the condition according to your experience.

- A. No Pointer & No Eye Track
- B. Only Pointer
- C. Only Eye Tracker
- D. Eye Tracker & Pointer both

1. Rank the conditions according to *which condition was best at helping you to enjoy the task.*

BEST	SECOND BEST	THIRD BEST	WORST

2. Rank the conditions according to *which condition was best at making you feel connected with your partner.*

BEST	SECOND BEST	THIRD BEST	WORST

3. Rank the conditions according to *which condition was best at help you stay focused on assembling the model.*

BEST	SECOND BEST	THIRD BEST	WORST

4. Rank the conditions according to *which condition was best at making you feel that you were present with your partner at the same workspace while performing the task.*

BEST	SECOND BEST	THIRD BEST	WORST

5. Rank the conditions according to *which condition was best for your partner to tell when you needed assistance.*

BEST	SECOND BEST	THIRD BEST	WORST

6. Rank the conditions according to *which condition was best at helping you understand what your partner was communicating to you.*

BEST	SECOND BEST	THIRD BEST	WORST

7. (Short Interview)

You mostly choose () condition as the best. Could you explain why?

You mostly choose () condition as the worst. Could you explain why?

How do you think you could best improve the user interface to better support remote collaboration?

What kind of applications you would like to use this system in?

How do you think the attention of a person affects the remote collaboration?

E.4. Per-Task Questionnaire for Helper (Remote User)

6. How often **you ask someone to assist you** with performing a real world task over video conferencing?

- Never
- Less than once a Month
- Once a Month
- 2-3 Times a Month
- Once a Week
- 2-3 Times a Week
- Daily

7. What **kind of help do you ask for or give to your partner in video conferencing?**

8. Please describe any **special activity that you helped with over a video conference, e.g. working on a project together or any task which might need special attention to it among the other things, etc.**

Questionnaire (Remote) Condition: - Task:

Below are statements about your experience. **Please circle on a number** to indicate your level of agreement or disagreement with each statement

1. I **enjoyed** the experience.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

2. I **felt connected** with my partner.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

3. I was **able to focus** on the task activity (constructing structure)

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

4. I felt that **I was present** with my partner on the same workspace while performing the task

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

5. I think that **my partner was able to sense that I was present** with him on the same workspace while performing the task

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

E.5. Post-Task Questionnaire for Helper (Remote User)

Questionnaire

(R)

After four conditions

Please rank the condition according to your experience.

- A. No Pointer & No Eye Track
- B. Only Pointer
- C. Only Eye Tracker
- D. Eye Tracker & Pointer both

1. Rank the conditions according to *which condition was best at helping you to enjoy the task.*

BEST	SECOND BEST	THIRD BEST	WORST

2. Rank the conditions according to *which condition was best at making you feel connected with your partner.*

BEST	SECOND BEST	THIRD BEST	WORST

3. Rank the conditions according to *which condition was best at help you stay focused on assembling the model.*

BEST	SECOND BEST	THIRD BEST	WORST

4. Rank the conditions according to *which condition was best at making you feel that you were present with your partner at the same workspace while performing the task.*

BEST	SECOND BEST	THIRD BEST	WORST

5. Rank the conditions according to *which condition was best for you to tell that your partner needed assistance.*

BEST	SECOND BEST	THIRD BEST	WORST

6. Rank the conditions according to *which condition was best at helping you understand what your partner was communicating to you.*

BEST	SECOND BEST	THIRD BEST	WORST

7. (Short Interview)

You mostly choose () condition as the best. Could you explain why?

You mostly choose () condition as the worst. Could you explain why?

How do you think you could best improve the user interface to better support remote collaboration?

What kind of applications you would like to use this system in?

How do you think the attention of a person affects the remote collaboration?
