ACIT5930

Master's Thesis

in

Applied Computer and Information Technology (ACIT)

May 14th, 2021

Universal Design of ICT

An experimental intervention with an eye-tracking device to promote social inclusion of user of augmented communication technologies.

Samip Bhurtel

Department of Computer Science
Faculty of Technology, Art and Design

OSLOMET

Preface

This work represents the culmination of our master's degree program in applied computer information technology with a focus on universal design.

This thesis would not have been possible without the help and guidance of those who encouraged, supported and influenced me during the writing process. Firstly, I would like to thank my supervisor professor Gustavo Mello and Pedro Lind for the motivation, guidance and continuous support from start of the thesis until the last hour of thesis submission. Your constructive suggestions pushed me to improve my thought and raise the quality of my work. I would also like to thank Prof. Dr. Renate Reniers, from the Institute of Clinical Sciences at the University of Birmingham. She provided me with the questionnaire for the experiment presented in this thesis. I would also like to thank my family abroad for their patience, encouragement and continuous support. I am thankful to the people who participated in our study for data collection, without them our study would not have been executed successfully. Lastly, I would like to express my gratitude to all who helped to make this project a reality in every way.

Samip Bhurtel
Oslo, May 14, 2021

Abstract

In this thesis, we design and perform an experiment using an eye-tracker, communication interface and questionnaires to promote empathy towards people with communication difficultes.

People with communication difficulties face lots of problems to be understood. Assistive technology helps individuals who cannot communicate verbally, people suffering from tetraplegia, locked-in syndrome, and brain paralysis.

Assistive technology alone does not entirely fulfill the communication needs because the attitude towards non-verbal people has not always been good. For controlling bad attitudes towards non-verbal people, society should know the difficulties faced by non-verbal people by getting into their shoes. In this context, we design a new framework to promote empathy and to help better social inclusion of the particular group of non-verbal individuals. We use a validated questionnaire, called QCAE (Questionnaire of Cognitive and Affective Empathy), to assess the level of empathy, composed of 31 questions, and add 9 more questions for the specific context of communication with eye-tracker to fit the design of our experiment. The result shows that, while the empathy level of participants does not increase significantly, the statistical power obtained for our extended questionnaire (75%) is not far from the standardly accepted power value (80%) to reject the null hypothesis. Moreover the power of our extended questionnaire is considerably larger than the power obtained with QCAE (63%).

Finally, a discussion for future directions of the study is also presented.

Keywords: Empathy, assitive technology, eye-tracker, non-verbal, tetraplegia, locked-in syndrome.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

Communication enables us to express opinions, feelings, and needs of people. It is easy when people can speak a common verbal language. However, some people cannot speak in our society due to physical disability, either an inborn disability or an accident. Some examples are people suffering from tetraplegia, brain paralysis and locked-in syndrome (Smith & Delargy, 2005). All of them find it difficult to express their emotions and needs simply because they cannot properly communicate verbally.

For instance, being difficult to communicate in public places leads to situations where one is unable to practice fundamental rights. For instance, sometimes local municipalities provide information about services provided to their local people. The municipality does not publish a version of the leaflet that is easy to read to save money. This situation makes it difficult for someone having learning disabilities to get information about the services provided by local administrative bodies (Shaw, Chan, & McMahon, 2012).

People with disabilities often face problems traveling in public transportation (Field, Jette, of Medicine (US) Committee on Disability in America, et al., 2007). To overcome the problems faced by people with disabilities while traveling, accessibility of information is a key to travel fearlessly. Accessibility refers to the availability of services that can be used by possibly the maximum number of people (Gutiérrez & Miravet, 2016). For instance, If there is no voice information about the bus stops, while a person with visual impairment is traveling, they find it hard to identify their destination. In such a situation, the public transport should be aware of implementing the voice-enabled information system about the destination. A person who cannot speak or communicate verbally finds it difficult to be understood (Ekstein & Caruth, 1969). Moreover, because of these difficulties, such individuals often face injustice and discrimination in everyday situations of our daily routines (N. J. Davis, 1992). Therefore, it is important to find new ways for society to learn the needs, opinions, and feelings of these individuals, guaranteeing a healthier level of social inclusion.

To make the communities inclusive, several governments have taken notable actions to cope with some of the challenges to promote social inclusion. For example voice information in public transportation for people with visual impairment, visual displays of destinations for people with hearing impairment. Still, much more needs to be done since people report that individuals with communication difficulties tend to think that society does not prioritize their

needs (Chinnery, 1990).

To further promote the social inclusion of individuals with communication disabilities, two essential aspects need to be considered.

The first one is to implement strategies for developing empathy in all society in general. Empathy is our ability, as individuals, to put ourselves in other's positions, trying to foresee other's reactions, behaviors, and feelings (Kerr & Speroff, 1954). In other words, it is a communicative process through which people understand and respond to the feelings of other (Janssen, 2012). For instance, suppose someone works at a clinic, and an older man comes in, complaining about back pain due to spine-related problem. For the doctor to understand the causes behind that pain, he or she needs to ask the right questions, in order that the information provided by the older man will lead to an accurate diagnosis. To that end, the doctor needs to have developed a certain level of empathy with people coming in, into the clinic, with complaints.

Some recent achievements within the scope of developing empathy in social communities have been made. For example, the regular interaction with someone suffering from disabilities can make people empathetic towards them (M. Kupetz, 2014). Doctors who try to understand how their patients feel can figure out the best way of treatment for their patients. It is believed that patient satisfaction can be improved from cognitive empathy, as the patient will feel comfortable sharing their views to doctors or medical practitioner without feeling being judged (Decety & Fotopoulou, 2015).

The second fundamental aspect for promoting social inclusion is technological development. Various technologies and devices have been invented to make a life of people easier. In the case they address communities with a particular kind of disability, which needs a special solution, they are usually named assistive technologies (Miskelly, 2001).

Assistive technologies have helped to accomplish tasks, which usually people with disabilities would not be able to perform. Some examples are assistive technologies for mobility such as walkers and wheelchairs (Buvanswari & Reddy, 2015), assistive technologies which help people with hearing disabilities (Scherer, 2004) and assistive technologies such as screen-readers for visually impaired people (Alves, Monteiro, Rabello, Gasparetto, & Carvalho, 2009). The latter ones deal with so-called human-computer interfaces (HCI), which can also be used for communication purposes. In case one can track what an individual, who can not speak, is reading or seeing in a computer screen, one could think of possible technological solutions as interfaces for such individual to communicate. There are some communication technologies which helps peo-

Figure 1.1: Figure A, from the top, left is an external processing unit, a controlled environment used for eye-tracking related calculation (Tobii, 2020), Figure B, on down left is tobii^{TM}eye-tracker that we have used for tracking eye-movement. Figure C shows how the data was collected through the participants.

ple communicate through their eyes only. Those technologies are used by those who are unable to move their forlimbs to communicate and rely on their eyes to communicate. To track the movement of our eyes when reading or seeing a computer screen is exactly what an eye-tracker does.

Eye-trackers are widely used in various fields, such as marketing and teaching, and there are different types of eye-trackers available. Some of them are screen-based, while others are wearable, based on glasses or other virtual reality support, or use a webcam eye-tracking system (Bulling, Roggen, & Tröster, 2009). There are already several eye-tracking technologies that assist people who can not speak. For example, screen-readers and gesture-based applications have helped non-verbal people to communicate (Isokoski, 2000).

In this thesis, we have used eye-tracking technology for conducting an experiment to test the ability of an eye-tracker to be used as an instrument for

promoting empathy. In our experiment, empathy has assessed in the particular context of communicating through an eye-tracker. With that experiment, we aimed at establishing a framework where a group of individuals can experience the difficulties associated with communication through an eye-tracking system and eventually, develop their empathy towards those forced to use such systems.

To reach these goals, we, therefore, formulated the following two research questions that is addressed in this thesis:

RQ1 How to develop empathy in verbal individuals towards individuals with special needs for communicating?

RQ2 Using an eye-tracker, how can such a framework for developing empathy be implemented in a way that can easily be reproduced?

There are some researches which have conducted experimental interventions in this domain, the closest one that we came across is the effect of empathy trait on attention to positive emotional stimuli: evidence from eye movements (Liu, Sun, Zhang, & Li, 2020). In this research, most of the equipment used are similar with the equipment that we have used in our research.

While our results show that the intervention performed does not develop empathy significantly, the experimental design introduced in this thesis may lead to more efficient interventions in the near future, as we discuss in the end of the thesis.

We start in Chapter 2 with the background and state of the art that will explain the importance of empathy, assistive technologies, social inclusion, empathy scale, and its validation. In Chapter 3, we present the methodology and experimental design in detail along with the methods used for data analysis. In Chapter 4, we present and describe the results, and in Chapter 5, we discuss possible interpretations, putting the main results in perspective. Finally in Chapter 6 we draw the main conclusions.

Chapter 2
Background and state of the art

In this chapter, we present the recent research achievements in augmented communication technologies and their types. We present the importance of eye-tracking technology as an interface for mitigating difficulties faced by non-verbal people, who cannot communicate verbally. We also present on how people communicate with eye-tracker.

We have also discussed assistive technologies, empathy, the importance of communication interface, and different scales used to measure empathy.

2.1 Assistive technology for communication

Assistive technology, the technology used to communicate for people with communication difficulties, alone does not provide all the solutions to overcome the difficulties faced by non-verbal people. Society can play a vital role in making them believe that they are also equally competent in doing everything that people who speak can do. To make verbal individuals, people who can communicate verbally, aware of the problems faced by nonverbal people, empathy towards non-verbal people is essential.

Several assistive technologies assist people to communicate well in society. Assistive technologies are often employed to help people with different special needs to communicate. We are going to focus on the particular needs derived from paralysis and how they affect communication.

In a head-controlled system, controlling the cursor can be done only with the movement of the head and neck (Evans, Drew, & Blenkhorn, 2000). Users do not need to use their hands to navigate. An example of head-controlled assistive technology is wheelchair (D. Kupetz, Wentzell, & BuSha, 2010).

The use of the tongue can also help control the mouse cursor or navigate through the contents of the computer through tongue motion. We can do that by installing magnetic sensors in tongue (Prabhu & Prasad, 2011). Thus, we can operate multiple devices by using such systems. People use tongue-controlled systems to dial phone numbers from mobile phones and control wheelchairs (Kim, Park, & Ghovanloo, 2012).

In eye-controlled systems, we can do the navigation through eye-tracking. In conditions where people lose the functioning of their limbs, people use eye-tracking technologies. Therefore, people use eye-trackers to operate computers and other electronic devices in such cases (Arai & Mardiyanto, 2011).

2.2 Eye-tracking devices and their applications

2.2.1 Importance of eye-tracker

Eye-tracking technology is a tool that is used to observe the eye-movement (Majaranta & Bulling, 2014). It is used in fields like infant and child research, user experience, medical research, education, and sports (Aslin, 2012).

For instance, some children cannot focus or have a short attention span disorder, which leads the child to under-perform in the classrooms, unable to do daily tasks. Eye-tracker is used on children to monitor the child's activities through the eye movement, the data achieved from eye-tracker can be used to change the cognitive development through proper counseling (Sasson & Elison, 2012).

To improve the user experience and interaction, researchers or software companies use eye-tracker (Nielsen & Pernice, 2010). It is implemented on software experience usability, web usability, mobile device usability to make the application as accessible for many people as possible. For example, data obtained from the eye-tracker can be used to design a better application by knowing user behavior through eye movement on how people spend time or find difficulties while using the applications (Nielsen & Pernice, 2010).

In sports, an eye-tracker is used to measure the focus of athletes while they are performing. For instance, the focus of the athlete who comes first in a race can be different from the other athlete who achieves second or third position. While performing on games, the strategies of the athlete can be achieved through their eye movement; later on, the strategies can be shared with other upcoming athletes (Barfoot, Matthew, & Callaway, 2012).

Likewise, eye-tracker is used in medical research. It is used to get the eye-movement data of patients who cannot speak verbally, which can help identify the messages they want to deliver. For example, it is used in neurological science to treat patients who have Amyotrophic lateral sclerosis (ALS) (Calvo et al., 2008). It is used to study the diseases like Schizophrenia (Carpenter, Kirkpatrick, & Buchanan, 1990).

Similarly, the use of eye-tracking technologies are in many fields where, such as in education, for monitoring students learning process of students (Slykhuis, Wiebe, & Annetta, 2005).

Eye-tracker plays a vital role in communication for people who cannot communicate verbally, conditions in which a person can only move their eyes up down and blink such as tetraplegia, brain paralysis, locked-in syndrome, and

other spine injuries (Bauer, Gerstenbrand, & Rumpl, 1979). There are different types of eye-tracker available according to the purpose of the study and its uses.

2.2.2 Types of eye-tracker

Based on the working principle, the types of eye-trackers are discussed below.

An eye-attached tracking device uses an attachment to the eye, such as a contact lens with sensors. The movement of the sensor attached to the eye is measured assuming that the sensor does not slip while the user rotates their eyes. Researchers use this eye attached tracking to study the dynamics and underlying physiology of eye movement (Robinson, 1963).

Electric potential measurement technology uses electric potential, measured with electrodes placed around the eyes; the process is called electrooculogram (EOG). The change in electric potential in electrodes due to the eye movement results in a change in the measured signal. Electric potential measurement is considered a robust technique for measuring eye movement associated with gaze shifts and detecting eye blinks. It also allows recording of eye movement even with the closed eye and is used in sleep research. However, it is challenging to use EOG to measure slow eye movement and to detect gaze direction (Elbert, Lutzenberger, Rockstroh, & Birbaumer, 1985).

Optical tracking uses a non-contact optical method for measuring eye movement. A specially designed camera senses the infrared light reflected from the eye. The data is then analyzed to extract the rotation of the eyes. This type of technology is widely used for eye-tracking and is usually preferred because they are inexpensive and non-invasive (Crane & Steele, 1985). Following is the working principle of the eye-tracker: the light of the camera is coordinated towards the participant's eye, making an appearance in the cornea and pupil of the eye (Cheng & Vertegaal, 2004).

2.2.3 Communicating with eye-tracker

There are different ways of communicating with eye-tracker. The most common communication approach is text entry by using gaze points directly, by setting up an on-screen keyboard. The user communicates by typing with the help of eye-gaze. The user focuses on the letters that they want to select on the digital keyboard on the screen. The computer then analyses the gaze behavior of the user to detect the key that the user intends to type during the same time (Majaranta & Räihä, 2007).

Eyelink™is a transparent board containing alphabets, numbers, and some crucial keys such as space, backspace, and start over (Swift, 2012). The board is held vertically to the user with letters facing them. It is, therefore, oriented correctly for the user of the device. The user points the eye-gaze to desired letters while the eye-tracker scans through the board at the user's eyes (Swift, 2012).

MINSPEAK™is a method to define languages in symbols (van der Merwe & Alant, 2004). MINSPEAK™is used to give several meanings with one picture. For instance, if there is an icon or symbol of an orange, then it can be associated with the ideas like fruit, color, and hungry, a way to communicate many vocabularies with a single icon (van der Merwe & Alant, 2004).

To overcome the slower communication, there is another technology called GazeSpeak™. It is an eye gaze communication system that is compatible with smartphones. GazeSpeak interprets the real-time eye gesture and helps nonverbal people to communicate (Zhang, Kulkarni, & Morris, 2017).

Most eye-tracking devices are expensive and cannot perform well under sunlight. There is some low-tech alternative such as E-tran boards which helps in communicating for the people with disabilities. An E-tran board is an interface used to communicate through the help of eyes (Grisold, 2004). E-tran frame is generally made up of transparent plexiglass. Letters in the E-tran frame are grouped together, and each group is identified with the color next to it (Grisold, 2004). E-tran frames can be customized in messages instead of letters and numbers. E-tran frame is considered to be an effective way for the people who rely on their eyes movement for communication (Grisold, 2004).

2.3 The concept of empathy and empathy quotient scales

Empathy is the process of understanding and responding to people's feelings by putting yourself in their position (Janssen, 2012). The best analogy for empathy means putting yourself in someone else's shoes. The term empathy is often confused with the word sympathy, but they are slightly different. Sympathy generally means to feel sorry for someone, whereas empathy means putting yourself in someone else's position and understanding the feelings of that person (Wispé, 1986). Empathy is considered to be the basis for understanding, trust in society. It is essential for understanding social functioning. It helps us

understand other people's objectives through their emotions, which are triggered by the various factors of their actions and expressions.

The degree of empathy may vary from person to person. For example, a person feels empathetic by looking at someone getting bullied; the same scenario may not enable empathy in another person looking at it. To test individuals empathy score, empathy scales are designed to measure or achieve the empathy (Paolo Senese, De Nicola, Passaro, & Ruggiero, 2018).

The empathy scales are used widely across the world and are validated in different languages. Empathy scales are designed to measure various attributes of people. Some are designed to measure mental capabilities or behavioral styles such as psychometric properties. At the same time, some are designed to measure individual differences and social psychology (Johnson, Cheek, & Smither, 1983).

Furthermore, as the empathy scales can be found in different languages (Paolo Senese et al., 2018), the accuracy in measuring empathy may vary due to linguistic differences, culture, background. For example, the tests such as the Social skills improvement system (SSIS) were designed as a comprehensive tool for finding the risk of children having difficulties in social behavior (Klaussen & Rasmussen, 2013).

2.4 Empathy test validation

Empathy test validation is necessary to achieve a precise result. For instance, if the designer is designing a new interface for people who have the nonverbal condition. The designer should be empathetic towards the user group. In this way, he will design a better interface for nonverbal people. Likewise, if a person knows the difficulties of the people suffering from diseases such as paralysis, they will be more empathic.

There are differences in empathy across a person. Theory suggests that aggressive people are lower in empathy compared to non-aggressive people. Aggressive people show lower affective empathy than those who are less non-aggressive in nature (Shechtman, 2002).

It has been explained that the level of empathy can be improved through intensive training methods (Furman, 2005). The methods to achieve empathy can be derived through various physiological, facial, and behavioral measurements (Kim & Lee, 2010). In other words, empathy is defined to understand the problems that other individuals face. To make people understand the problem

faced by non-verbal people, researchers use empathy scales with picture story methods and questionnaires (Kim & Lee, 2010).

There are several empathy tests available. For instance, the balanced emotional empathy scale; is a one-dimensional measure that considers empathy as increased reactiveness to the emotion experienced by other people. A balanced emotional scale measures the level of happiness or suffering experienced by the participants (Mehrabian, 1996). The scale suggests, a person having a high tendency of emotional empathy scale is more likely to be more emotional as they felt like weep after participating in test (Mehrabian, 1996). People were more tolerant to the infant crying and were less abusive to their child after performing the test. People were more unselfish in their behavior towards others and helped others as a volunteer. People were likely to be less aggressive (Mehrabian, 1996).

The basic empathy scale is understanding and sharing other people's state according to emotions (Cohen & Strayer, 1996). The measurement of empathy is done based on sadness, fear, anger, happiness, and these measures are related to cognitive and affective empathy (Jolliffe & Farrington, 2006).

The multidimensional emotional empathy scale is more focused on the affective component and is designed for adults and adolescence (Caruso & Mayer, 1998). Affective empathy can be perceived as understanding the emotions of another person and responding to them appropriately. Multidimensional emotional empathy test comprises six subscales, i.e., positive sharing, empathic suffering, responsive crying, emotional attention, emotional contagion, and feeling for others. The total score of empathy is achieved by adding all the six measures (Caruso & Mayer, 1998).

The empathy quotient scale is designed to measure affective and cognitive components of empathy in a concise and accessible way (Baron-Cohen & Wheelwright, 2004). It comprises 40 items of empathy with 20 control items. The scale consists of four scales, i.e., disagree strongly, disagree slightly, agree slightly, strongly agree (Baron-Cohen & Wheelwright, 2004).

The feeling and thinking scale is designed for children (M. H. Davis et al., 1980). It contains four independent subscales, i.e., perspective-taking, empathic concern, personal distress, fantasy (M. H. Davis et al., 1980).

Griffith's empathy measure was made to support the assessment of multi-informant. It was designed to measure the empathy of adolescence and children (Dadds et al., 2008). It is designed to measure cognitive and affective components of empathy (Dadds et al., 2008). Griffith's empathy scale was validated in China in a Chinese context.

The development of the Toronto empathy scale was done to consider the concept of empathy on a broad level (Spreng, McKinnon, Mar, & Levine, 2009). It consists of the items from 11 different empathy scales questionnaire. Toronto empathy questionnaire has an emphasis on emotional components of empathy (Spreng et al., 2009). Items on Toronto, the empathy questionnaire are rated based on 5 point scale, i.e. (never, rarely, sometimes, often, always) based on the experience the participants (Spreng et al., 2009).

The interpersonal reactivity index separates empathy in four different aspects and is related to the aspects of emotionality, others sensitivity, self-esteem, and social functioning (De Corte et al., 2007). Every scale shows the patterns of relations with the other measure. The scale measurement is designed in 5 categories (M. H. Davis, 1983).

The questionnaire for cognitive and affective empathy scale was validated in England. While validating the QCAE, it was found to be positively tallied with a balanced empathy scale(BES) (Reniers, Corcoran, Drake, Shryane, & Völlm, 2011). The item pool was created to develop the QCAE consisted of 65 items derived from existing questionnaires. Participants completed the 65 item version, and using exploratory factor analysis (principal components analysis), researchers reduced the questionnaire to the final 31 item version that was then validated using confirmatory factor analysis (Reniers et al., 2011).

Researchers collect data by different means, such as, interviews, questionnaire, experiments. The next phase is always getting the insights of the data or to analyse data to get the insights of the study, that is where statistics comes into account. Study of data is called statistics, researchers use statistics to define data, analyse the data and draw a conclusion for their study (Ostle et al., 1963).

2.5 Concepts from statistics: hypothesis testing and power analysis

To validate the findings in our research, we need to carry out hypothesis testing defining an experimental hypothesis and a null hypothesis. The result always supports one hypothesis and rejects the another.

Null hypothesis is the hypothesis where there is no difference in the result between two groups, one where no intervention is done (control group) and another which experiences the intervention (experimental group). In any experiment, to accept the null hypothesis, no effect can be observed. For instance,

there is no change in effect to promote empathy, even by giving the different treatment in the experiment for control and test group (Rouder, Speckman, Sun, Morey, & Iverson, 2009).

The hypothesis where we assume that an effect is likely to be seen after the experiment. For example, the experimental hypothesis can be that the use of eye tracker and communication interface to communicate, increases empathy. If it is accepted, then we reject the null hypothesis. If there is no difference in the result, we accept the null hypothesis (Rouder et al., 2009).

To better understand data collection and data analysis, we use various statistical tools which provide knowledge about the data. Some of these tools are the mean of observed values, standard deviation of the data values. Some of these basic statistics concepts will be needed for hypothesis testing (accept or reject null hypothesis)(Matthews & Farewell, 2015), and there we list next the main concepts.

- Mean: Mean is the sum of all the values divided by the number of samples:

$$\frac{1}{n}\sum_{i=i}^{n} x_i \tag{2.1}$$

  Generally, it represents the typical value of the dataset. However, it may not fairly represent one typical value, for instance, if there are outliers. To find the spread of data in the dataset, we use standard deviation (Matthews & Farewell, 2015).

- Standard deviation: Standard deviation tell us how broad the range of observed values is, for a particular dataset. It also tells us typically how far or how close is each individual data from the mean (Matthews & Farewell, 2015).

- Normal Distribution: The normal distribution is a continuous probability distribution which plays a central role in statistics. Standard normal distribution is represented by a bell curve where the mean is zero and the standard deviation is one. The distribution tells us that the data distributes typically around the mean and not so frequently, as we go far from the mean value (Matthews & Farewell, 2015).

- Z-score: Z-score is used to identify the distance in the normal distribution between the mean and a data point in a normally distributed dataset. It is

measured in number of standard deviations below or above of mean value. For example, the z-score of 1 means the value z which is one standard deviation above the mean. Likewise, the Z-score of -1 implies the value which is one standard deviation below the mean value. The Z-score play a fundamental role in determining the likelihood (probability) of type-I and type-II errors (Matthews & Farewell, 2015).

- Type-I error: It is a type of error, which rejects the null hypothesis, when the null hypothesis is true.

- Type-II error: It is a type of error, which fails to reject the null hypothesis, when the null hypothesis false.

- Effect size: The effect size measures the effect of relationship between two given dataset. Correlation between two variables can be a good representation of effect size. It plays a vital role in statistical analysis to test a hypothesis. It is used to evaluate the strength of statistical claim. It is widely used in statistics namely power analysis for sample size planning. The Effect size is calculated using ($\mu_c = \mu_e$), is given by

$$D = \frac{|\mu_c - \mu_e|}{\tilde{\sigma}}\,, \tag{2.2}$$

  where $\tilde{\sigma}$ is the so-called pooled standard deviation, given by the weighted average of the standard deviations of each group (σ_e and σ_e):

$$\tilde{\sigma} = \sqrt{\frac{\sigma_c^2 + \sigma_e^2}{2}}\,, \tag{2.3}$$

- Power($1 - \beta$): We use power to validate the fact that the findings bring some significance. It is affected by the overlap between the two distributions and the sample size. It is represented by $1 - \beta$, where β is the probability of having type-II errors. It takes values between 0 and 1. A power close to 1 is desirable. However, a power of 0.8 is commonly used, meaning that there is an 80% probability that it rejects the null hypothesis (Matthews & Farewell, 2015).

- Level of significance (α): It is used to find the significance of the study while rejecting the null hypothesis. Usually it gives the probability of having type-I errors. The alpha value is set to 0.05, which is usually considered as standard (Matthews & Farewell, 2015).

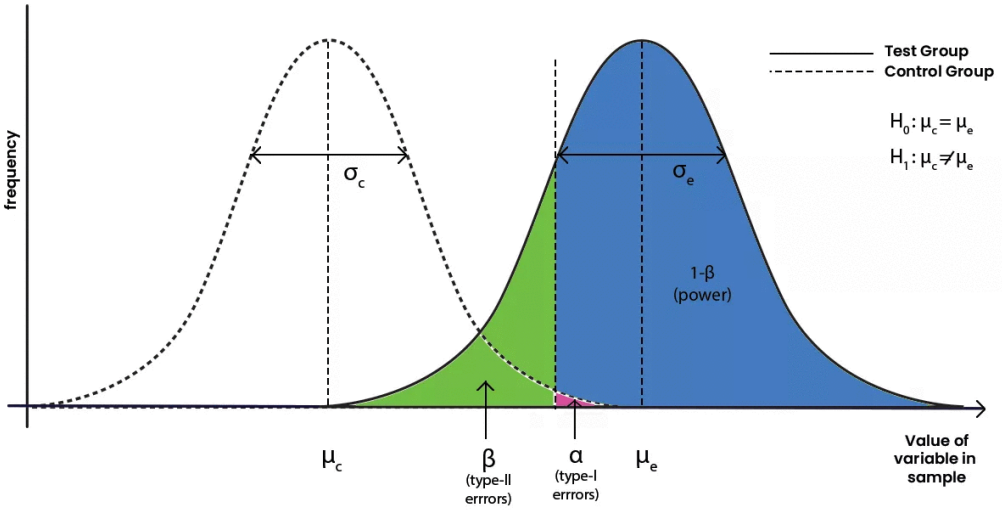

Figure 2.1: Power analysis.

For two datasets with the same sample size, the effect size D can be used together with chosen Z-scores with a given significance level α and for a given power $1-\beta$, to provide an estimate of the minimum sample size in each group:

$$N_c \sim N_e \sim \frac{(Z_\alpha + Z_{1-\beta})^2}{D^2}\,. \tag{2.4}$$

where N_c and N_e are the sample sizes of each group.

The analysis that determines the minimum sample size to accept or reject the null hypothesis is called a power analysis. We use factors like effect size, Z-score, mean, standard deviation, level of significance in normal distribution to calculate the power of the study. Power analysis deals with both type I and type II errors. The graphical representation of power analysis can be seen in figure 2.1.

In the next chapter, we discuss experimental design.

Chapter 3
Methodology and experimental design

3.1 Overview

We devise an experiment for verbal individuals to develop empathy towards individuals with problems in communicating. The experiment is based on a framework that we implement for communication through an eye-tracker. We hypothesize that, after training communication through an eye-tracker, using our framework, participants will be more aware of the inherent difficulties and consequently will gain more empathy towards those who need such types of equipment. To measure the level of empathy, we introduce a questionnaire which extends a established questionnaire Appendix A for measuring empathy levels. In the rest of this section, we present a description about the experiment, recruitment of participants, use of equipment and software, procedure illustrating the stages of experiment, schedule, and approach for data analysis. An overview sketch of all the experimental design is given in figure 3.1.

3.2 Participants, equipment and software

We recruited a total of 44 participants for the experiment, 22 used for the control group, and 22 for experimental group. The participants were bachelor and master students at Oslo Metropolitan University, NTNU, University of Oslo, with the approximate ages between 20 and 40, from both genders. Only individuals who have not suffered any paralysis-related conditions in the past two years were elected to participate. We invited participants through email, which contained information about the purpose of the study, the instructions to register, and additional resources necessary to participate in the experiment successfully. We carried out the registration process online, which included a consent to be signed by each participant before joining the study and a questionnaire to collect demographic data, namely gender, name, email, country of origin. See section 3.5 about ethical requirements.

The equipment and software for our experiment included:

- An eye-tracking interface
- A communication interface

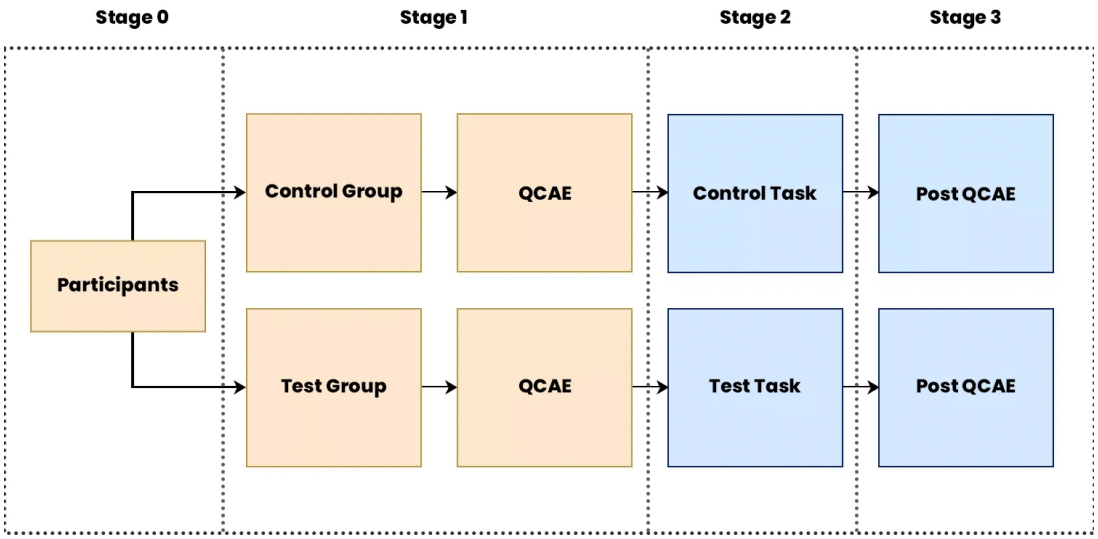

Figure 3.1: Experimental design.

- A questionnaire.

We used the Tobii [1] package, which consists of the Tobii Pro X3-120 Hardware and the software Tobii Pro Lab Presenter Edition, as an eye-tracking interface. The Tobii Pro X3-120 combines an infrared light projector and a camera. The projector emits a pattern of infrared light that is reflected by the back of the retina. This reflection increases the contrast between the image of the eye and the pupil. The camera component can capture visible and infrared light at a sampling rate of 120 Hz. To operate the Tobii Pro X3-120, we used the Tobii Pro Lab. This software allows transforming video data captured by the Tobii Pro X3-120 into coordinates on the screen that map where the participant is looking at, its gaze. In addition, the Tobii Pro Lab allows designing experiments that use the gaze coordinates on the screen as inputs to the computer. Tobii Pro Lab supports the windows operating system; it runs in windows seven, windows 8.1 (64-bit), and windows 10 (64-bit). Minimum system requirements for CPU is Intel Core i7 (4 cores), 2.0 GHz, and requirement for RAM is 8 GB. Besides the windows operating system, we can run Tobii Pro X3-120 on Mac or Linux using an external processing unit (EPU) in a controlled environment.

[1] https://www.tobii.com/

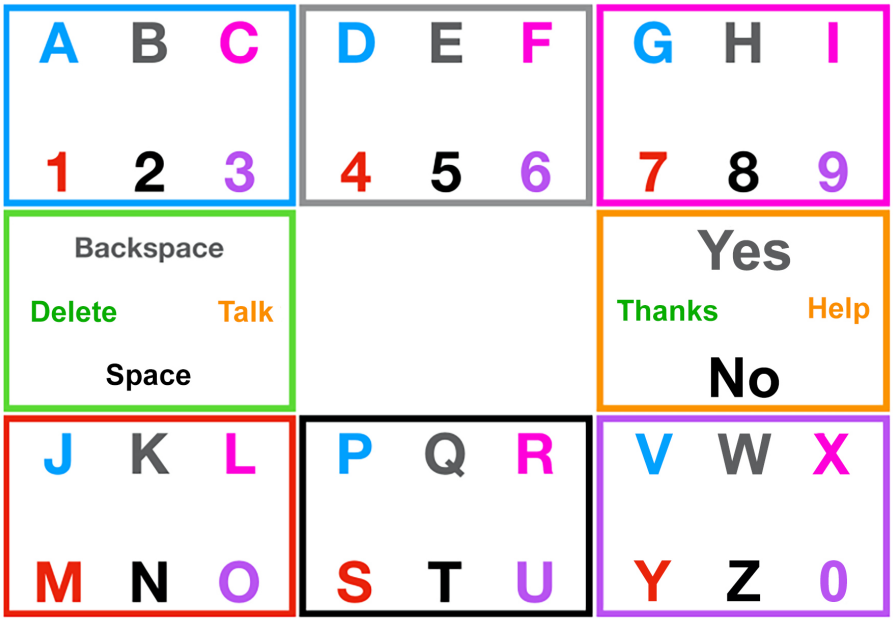

Figure 3.2: E-tran letter board with letters and words from (Krohn et al., 2020) with permission. Two eye movement selection is done. Character selection is done on by selecting one frame, after that color of character is identified by selecting color of the frame.

EPU is a feature of Tobii for providing compatibility to operating systems other than windows (Tobii, 2020).

As a communication interface, we have used a digital version of the E-tran frame. E-tran frames enable the choice of symbols on a letter-board through the eye movement (Grisold, 2004). Boards are split into nine parts. One central part does not contain any letters and numbers, and the remaining eight outlying sections contain letters, numbers, and words. The outlying sections is recognized by direction and color (Swift, 2012).

To measure empathy, we have used the Questionnaire of Cognitive and Affective Empathy (QCAE) developed by Renate Reniers (Lecturer in Psychiatry, Institute of Clinical Sciences on the University of Birmingham). The QCAE measures both cognitive and affective components of empathy (Reniers et al.,

2011). This questionnaire consists of 31 items with four choice scale. The QCAE consists of five sub-scales, namely: online simulation, perspective taking, proximal responsivity, emotional contagion and peripheral responsivity (Reniers et al., 2011). For our purposes we had to extend the questionnaire to adapt it to our specific experiment. See section 3.3, and Appendix A.

To apply the online questionnaire, we made the questionnaire in *Nettskjema* [2], which is a platform from University of Oslo, in which we can create a form to have a data collection for research and which follows all the ethical requirements in Norway for handling data. All the questions were published in it, and each question has four possible answers (i.e., strongly disagree, slightly disagree, slightly agree, strongly agree). The participants were able to choose only one out of the four options, which may be changed until the submission of all answers. In *Nettskjema*, we have created two forms with the same questionnaire. One form was used for the test of empathy before the experiment, and the second form was used collect the data after the experiment (nettskjema, 2020). The consent form is given in Appendix B.

3.3 Towards a new questionnaire for accessing empathy of people

To investigate a posteriori, if our questionnaire extension was properly developed, we separate the analysis so called Baseline and our Complete. The baseline test refers to the cumulative score of thirty-one questions taken from QCAE (Reniers et al., 2011). Apart from that thirty-one questions, we formulated additional nine more questions to address the specifities in our experiment, namely:

- I feel frustrated when I cannot communicate my ideas.
- I complete the sentences of other people when they find it difficult to communicate their ideas to help them.
- If someone is suffering from stutter, I prefer speaking rather than texting.
- I try to be patient when people speak with me very slowly.
- In a situation, when there are handicapped students in a classroom, the pace of the class should be adapted to the handicapped student.

[2] https://www.uio.no/tjenester/it/adm-app/nettskjema/

- In a situation, when there are only one handicapped student in a classroom, the pace of the class should be adapted to him/her independently of how many students are in the class.

- I feel impatient when people communicate slowly, so I complete other people's sentences to make the communication more efficient.

- I believe that the handicapped students should catch the pace of the group in a classroom to not delay the progress of other students.

- I feel left out (or excluded) when I cannot participate in a conversation.

. To select the additional questions related to our research, we prepared 20 questions altogether. After that, we performed a focus group meeting with 6 participants in two sessions. In the first session, there were 3 participants and was followed by the same number in the second session. In order to get a good feedback, we balanced the group by recruiting three males and three females for focus group discussion. One focus group meeting was held in the AI lab at Oslo Metropolitan University with the supervision of an expert, and another was done online.

As per the discussion, we modified the questions and selected nine questions which now composes the complete questionnaire together with all the questions in the baseline questionnaire. The focus group meeting minute can be seen in Appendix C. Similarly, while performing the calculations, we have separated the results of QCAE as a baseline score (Reniers et al., 2011). As we will see the extendended questionnaire enables more accurate results reflected in a power of the sample statistics.

3.4 Procedure of experiment

Before the experiment, we randomly divided the participants into two groups and we make sure that both control and test group are well mixed. Both groups had the same number of participants, and they were balanced for age, gender, education achievement, marital status, country of origin. Once the groups were established and adequately balanced, each group passed through the different stages of the experiment as sketched in figure 3.1. The experiment consists of three stages, namely pre-testing, intervention, post-testing.

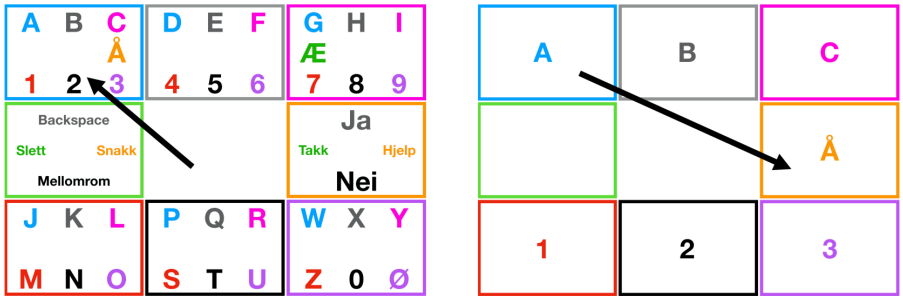

Figure 3.3: Selection of characters by two movement processes from (Krohn et al., 2020) with permission. (Left), a group of characters is selected through the movement of the eye. (Right), the color of the frame matching with the color of a letter or number is selected through the movement of eye.

During the pre-testing, the participants in both groups answered to the complete questionnaire to establish a baseline measure of empathy. Their scores were recorded separately and stored for later analysis.

In the second stage of the experimental intervention, both groups were instructed about the usage of augmented communication technologies and their importance for non-verbal people. After the explanation, each participant carried out a task according to the group they were assigned to, control group or test group.

We performed the sessions alternatively. For instance, every odd number assigned the participant to performed the test sessions and even number assigned performed control task. The first participant performed test task, and the second performed control task, and so on.

Before the experimental session, we fixed everything that was needed for the experiment. In order to do so, we created separate folders in the computer, one for keeping data of the test group and one for the control group. All the resources such as images used in the experiment were stored in respective folders, i.e.test and control groups. After that, we apply the timeline to perform the experiment through the eye-tracker application. See figure 3.1

3.4.1 Test group

In this thesis, we have used test group and experimental group as the same group interchangeably to describe the results.

After welcoming the participant, we explained him/her the uses of eye-tracker and the people who rely on this device to communicate and run their daily lives. The experimental group was trained to use the digital E-tran board and the eye-tracker to build messages, see figure 3.2 and 3.3. After this training, they were requested to make some short dialogues using the eye-tracking interface and the E-tran board exclusively. The tasks to be done in the test group were:

- Please type hello with the letters that are available on the screen.
- Can you type your full name.
- Please type: Can you help me.
- Please type: Can you take me to the restroom.
- Please type: I am hungry, can you do me some eggs.
- Please type: I am uncomfortable. Can you change my position.
- Please type: I am in pain. Can you give me some medication.
- Please type: Dialog. Start with "Hi, how are you".

The time taken for each step in keeping the dialog was recorded during the experiment. We also noted any kinds of difficulties or issues that the participants faced during the experiment. The participants were encouraged to take short breaks between the experiment if they experienced some kind of fatigue or difficulties. After completing the experiment, participants were asked to tell the impression about using the device for communicating. Furthermore, we asked participants to tell the impression about depending on the device that was used to communicate to people.

Figure 3.4: We show this figure to the participants of the control group. In this figure, participants look for a minute and share what they found looking at this image.

3.4.2 Control group

At the same stage after the explanation session, the control group was asked to use the eye-tracker to look at abstract images, not related with frames for communication purposes and was also be shown the E-tran board. Nevertheless, they did not build any message from it. The image used for the experiment for the control group is shown in fig 3.4.

The image in figure 3.4 is shown to the participants. They were given time to see the images for a minute and were asked what they saw on the image. The same image was shown four times separately, and participants were asked about the experience after seeing the image in repetition.

To notice that the output obtained from the control group in this experiment, will be used for another research project, beyond the scope of this thesis. We only kept the data achieved from pre and post questionnaire. The data of eye-gaze achieved from eye-tracker will be used in another research, for which

this experimental procedure contributed with data collection.

3.4.3 Post experiment

Finally, on stage three of the experiment, the QCAE was independently applied one more time to both groups. Compared with the initial empathy levels, the score resulting from the second testing was stored to calculate the intervention stage's effect.

3.4.4 Post experiment interviews

After completing the experiment and doing some analysis, we found some participants who scored more than other average results. To know the reason behind the high score for these specific participants, we organized an online interview with the group of participants who scored high in the empathy test. We perform this interview to see the reason behind the high empathy in an individual, see if the participants were empathic before or he/she became empathic after going through our protocol.

3.5 Data management and ethical requirements

We record the data of each experimental session separately. We have created the folder for the experimental group and control group. All the resources required to perform each session are saved inside the folder in their respective locations. The data collected from each participant in the experiment is kept confidential to maintain the integrity of the data and carry out the survey ethically. Apart from that, we have also taken permission from Norsk Senter for Forskningdata (NSD). NSD is an organization that assesses ethical requirement in projects publishing data collection. The application has reference number 119986, that we prepared for NSD and it is attached to this thesis. See Appendix E.

We made the documentation for data collection approval and submitted it to the Norwegian center for research data (NSD). After a month of assessment, ethical consideration was given by NSD, see Appendix F and it was written:

> Our assessment is that the processing of personal data in this project will comply with data protection legislation, so long as it is

Figure 3.5: Boxplot plot figure used for the exploratory data analysis.

> carried out in accordance with the Notification Form and attachments, dated 28.04.21, as well as in correspondence with NSD and our assessment. Everything is in place for the processing to begin. Good luck with the project!.

3.6 Methods used for data analysis and visualization

After the data collection, we imported all the CSV file formats in the local computer from nettskjema [3]. Secondly, we uploaded the file in *Jupyter notebook* [4], an open-source web application that is used to perform exploratory data analysis and data visualization. Jupyter notebook supports *Python* [5] programming language. Python is a high-level, general-purpose programming language used in software development, machine learning, scientific computing, and data science. It is one of the popular programming languages right now because of its uses in multiple domains (Vallat, 2018).

Therefore, we performed boxplot [6] analysis from the data that we have:

- The box plot sketched in figure 3.5, gives five important information about the data dividing the values into four groups (quartiles) Q1-Q4. The vertical line in the middle of the box is the median, value which lies in the midpoint, also denoted as Q2.

- The two ends of the box mark the upper boundaries of Q1 and Q3.

- The whiskers outside the box show the highest and lowest observations in the data (Williamson, Parker, & Kendrick, 1989).

[3] https://www.uio.no/tjenester/it/adm-app/nettskjema/
[4] https://jupyter.org/
[5] https://www.python.org/
[6] https://seaborn.pydata.org/generated/seaborn.boxplot.html

Furthermore, we have also done swarmplot [7] analysis, an analysis where we put individual empathy score data points in visualization (Reynolds & Peng, 2005). We have integrated the swarm plot with the boxplot analysis, which can be seen in the next chapter.

After that, we have done further analysis using scatter [8] plot, the analysis to find the patterns in our data (Touchette, MacDonald, & Langer, 1985). We have used a scatter plot to show the relationship of empathy score with time-variable, i.e., time spent on the experiment, time spent to fill the questionnaire, the time gap between pre-intervention and post-intervention of the participants.

In the next chapter, we have discussed the results.

[7] https://seaborn.pydata.org/generated/seaborn.swarmplot.html
[8] https://seaborn.pydata.org/generated/seaborn.scatterplot.html

Chapter 4
Results

In this chapter, we present and describe the results that are drawn from the data collection. For presenting the results we have used three different visualization tools, box plot, swarm plot, and scatter plot. The box plot and swarm plot analysis are done in the same figure. Also we have divided the results in two parts: baseline questionnaire (BQ) and complete questionnaire (CQ). We have represented QCAE with 31 questions as baseline questionnaire, and the sum of baseline and the extended questionnaire is represented by complete questionnaire.

4.1 Results

All results were collected in the table shown in 4.1. In table 4.2, we show the difference in empathy between control and experimental group.

BCT	BBT	ACT	ABT	BCC	BBC	ACC	ABC
135	108	137	109	132	105	127	100
107	81	123	98	122	96	120	95
134	107	124	97	116	90	119	95
106	82	97	73	111	86	98	76
112	87	99	78	104	82	100	76
128	101	151	117	119	94	124	97
113	86	117	89	131	105	128	102
135	104	137	107	115	87	106	84
126	99	123	96	130	100	111	86
134	100	112	82	133	105	142	110
116	86	114	86	114	89	125	93
138	106	141	108	121	93	120	93
127	100	120	98	116	89	111	85
114	90	112	88	117	89	109	83
128	107	135	107	124	93	120	92
105	79	108	81	124	100	118	94
122	94	119	93	120	95	121	96
139	111	139	109	138	102	135	99
137	108	132	105	121	94	122	96

117	87	126	95	115	88	113	85
128	103	128	102	143	115	110	84
128	97	127	97	121	97	111	87

Table 4.1: This table shows the complete score and baseline test score from the experiment participants in the test and control group. In this table, BCT = "B"efore Intervention + "C"omplete Test + "T"est Group, BBT = "B"efore Intervention + "B"aseline Test + "T"est Group, ACT = "A"fter Intervention + "C"omplete Test + "T"est Group, ABT = "A"fter Intervention + "B"aseline Test + "T"est Group, BCC = "B"efore Intervention + "C"omplete Test + "C"ontrol Group, BBC = "B"efore Intervention + "B"aseline Test + "C"ontrol Group, ACC = "A"fter Intervention + "C"omplete Test + "C"ontrol Group and ABC = "A"fter Intervention + "B"aseline Test + "C"ontrol Group.

CC	**CB**	**TC**	**TB**
-5	-5	2	1
-2	-1	16	17
3	-5	-10	-10
-13	-10	-9	-9
-4	-6	-13	-9
5	3	23	16
-3	-3	4	3
-9	-3	2	3
-19	-14	-3	-3
9	5	-22	-18
11	4	-2	0
-1	0	3	2
-5	-4	-7	-2
-8	-6	-2	-2
-4	-1	7	0
-6	-6	3	2
1	1	-3	-1
-3	-3	0	-2

1	2	-5	-3
-2	-3	9	8
-33	-31	0	-1
-10	-10	-1	0

Table 4.2: This table is derived from the data shown in table 4.1. Calculation was done by subtracting before test score by after test score where, CC = ACC-BCC, CB = ABC-BBC, TC = ACT-BCT and TB = ABT-BBT. The representation of CC = "C"ontrol Group + "C"omplete Test, CB = "C"ontrol Group + "B"aseline Test, TC = "T"est Group + "C"omplete Test, TB = "T"est Group + "B"aseline Test.

4.2 Box plot and swarm plot results

The first two box plots in the figure 4.1 show the relationship between complete score before experiment and overall score of baseline before the experiment for test group, which is denoted by BCT and BBT. The mean score for BCT and BBT was 124 and 96 respectively; and standard deviation was 11 and 10, respectively. Likewise, we also calculated median and variance, the median achieved for BCT and BBT is 127 and 99, and the variance is 122 and 101, respectively.

Similarly, for the third and fourth box plot of figure 4.1, that is (ACT) complete score after experiment, (ABT) baseline score after experiment for test group, the mean value achieved was 123 and 96 respectively. The standard deviation achieved for the ACT and ABT was 14 and 12. Similarly, the median achieved is 123 and 97, and variance achieved is 184 and 133 respectively for ACT and ABT.

For the control group, the trend followed by the score is described below:

Additionally, box plot BCC and BBC respectively represents complete score and baseline score before the experiment for control group. The mean for BCC and BBC achieved was 122 and 95, and the standard deviation score was 9 and 7, respectively. The median score for BCC and BBC is 121 and 94 and the variance achieved is 85 and 62 respectively.

Likewise, ACC and ABC represent complete score and baseline score after the experiment for control group. The mean for ACC and ABC was 118 and

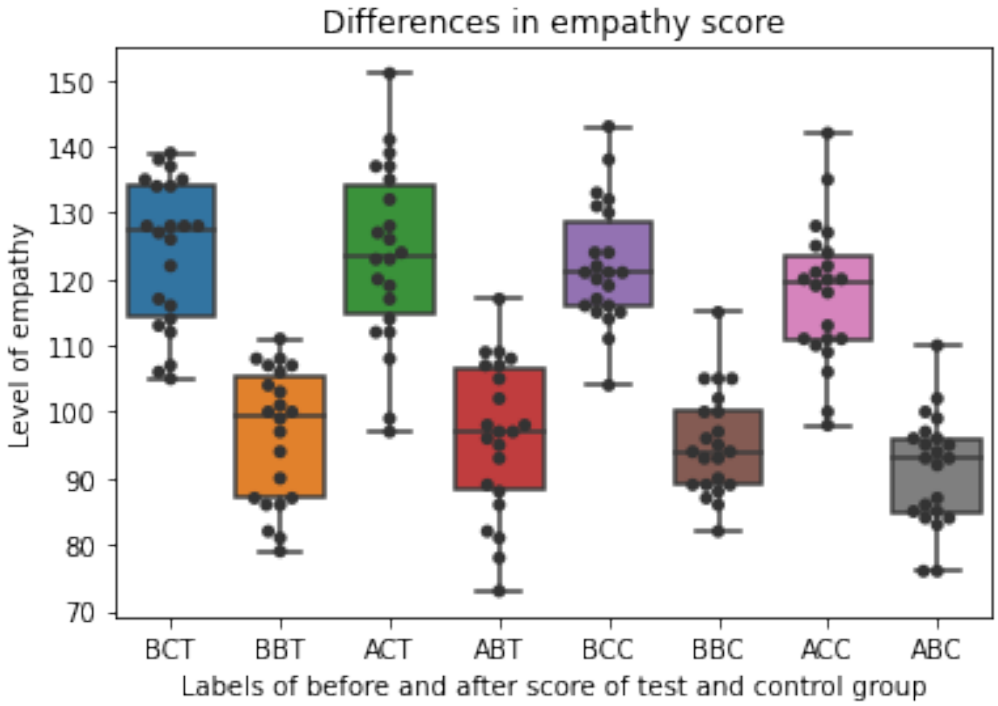

Figure 4.1: The box plot and swarm plot derived from the table 4.1; we have plotted the individual empathy score in the boxplot in the form of a swarm plot.

91, and the standard deviation was 11 and 8, respectively. The median was 119, and 93 and the variance was 112 and 70, respectively.

After plotting the scores, we came to know that there was no significant change in the score after the intervention of the test group compared to the control group, but we observe that the empathy level in the control group has decreased to some extent after the intervention. Whereas, in the test group, we can see an unnoticeable decrease in the scoreline, from which we can hypothesize that the intervention did not affect the empathy score but remained unchanged after the intervention for the test group.

From the figure 4.1, there is a small decrease in mean and median whereas there is an increase in variance and standard deviation in the empathy score of the test group after the experiment. Contrary to this, the average mean and median for the control group have also decreased after the intervention. However, there seems to be no significant change in empathy after the intervention

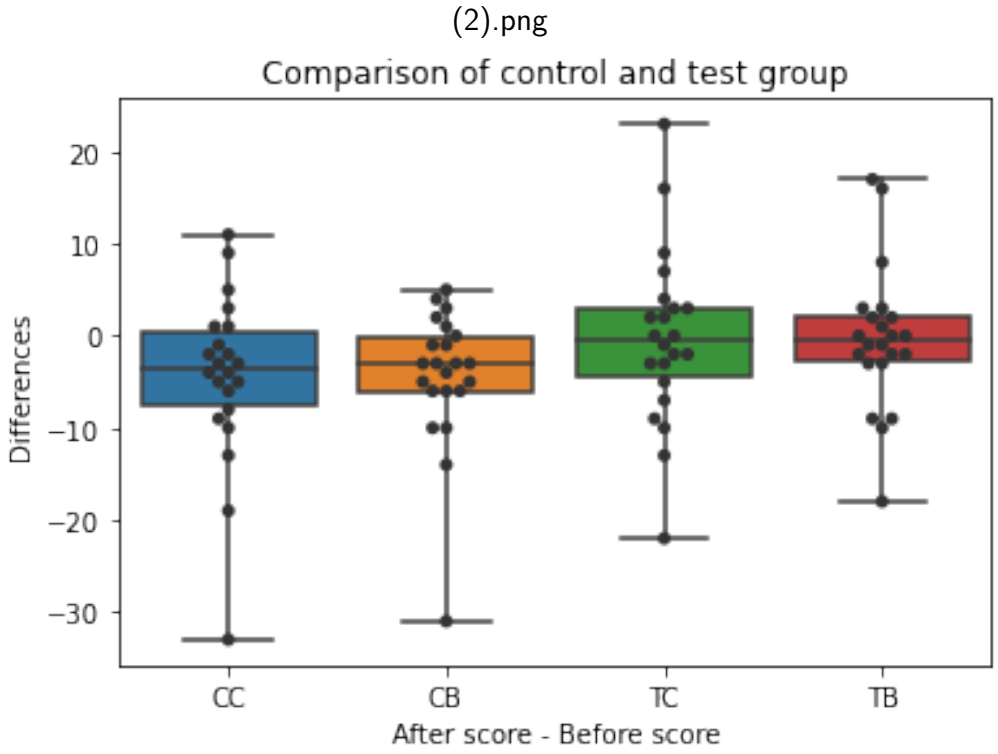

Figure 4.2: The box plot diagram shows the difference that we achieved after the experiment. We compared the results before and after the experiment for both the test and control groups. This boxplot is derived from the data shown in figure 4.2 and it also shows the plot about where the empathy levels stands according to the participants.

for control group.

In figure 4.1, we can also see the complete score of test group after the experiment (ACT) has a significant increase in comparison with a complete score of test group before the experiment (BCT). However, the increase may have been possible only due to some of the outliers. In general, we can see a decrease in the median in ACT compared to BCT. Similarly, the baseline score after the experiment(ABT) and baseline score before the experiment (BBT) also follows the same pattern as ACT and BCT, i.e. there is an increase in the score if we see the stretch of the boxplot, but there is a decrease in the median value. The codes used for creating box-plot with swarmplot is mentioned below:

```
ax = sns.boxplot(data=analysis, whis=np.inf)
ax = sns.swarmplot(data=analysis, color=".2")
plt.title("Differences in empathy score")
plt.ylabel("Level of empathy")
plt.xlabel("Labels of before and after score of test and
            control group")
```

Likewise, in figure 4.1, when we compare the scores before and after the experiment of the control group, the median and the scoreline in the BCC and ACC are almost the same. This means the intervention does not seem to change the level of empathy. Similarly, when we compare the baseline score of the control group, before and after the experiment, we do not see much significant change in the results as well.

Table 4.2 is derived by calculating the difference of after and before test scores for control and test group for both complete and baseline scores.

For instance, CC = ACC-BCC, CB = ABC-BBC, TC = ACT-BCT and TB = ABT-BBT.

For the control group, the sum of the difference in the complete questionnaire consisting 40 questions after and before was -97 and the difference in baseline questionnaire for control group was -86. Similarly, the sum of the difference between before and after the experiment of the complete questionnaire for the test group was -8 whereas, baseline difference was also -8.

While calculating the mean and standard deviation for the differences achieved before and after experiment were: mean value for CC = -4, CB = -4, TC = 0, TB = 0.

Similary, the calculation for standard deviation achieved was CC = 9, CB = 7, TC = 9, TB = 8.

Figure 4.2 shows that there are six individual participant's scores in CC, which are within or above the third quartile and most of the scores are inside the box between Q1 to Q3. Similarly, for CB, most of the scores are within the median or above it. There are seven scorelines that falls below the Q1.

While we look at the test group results in figure 4.2, for TC, there are some of the outliers in TC. Likewise, for TB, it is similar to TC.

In figure 4.2, we can also visualize the total difference in complete score for control group (CC) and total difference in complete score for test group (TC). The broadness of the box is higher in TC, and a similar result is followed while comparing the differences in the control group baseline score(CB) and baseline score difference in the test group (TB).

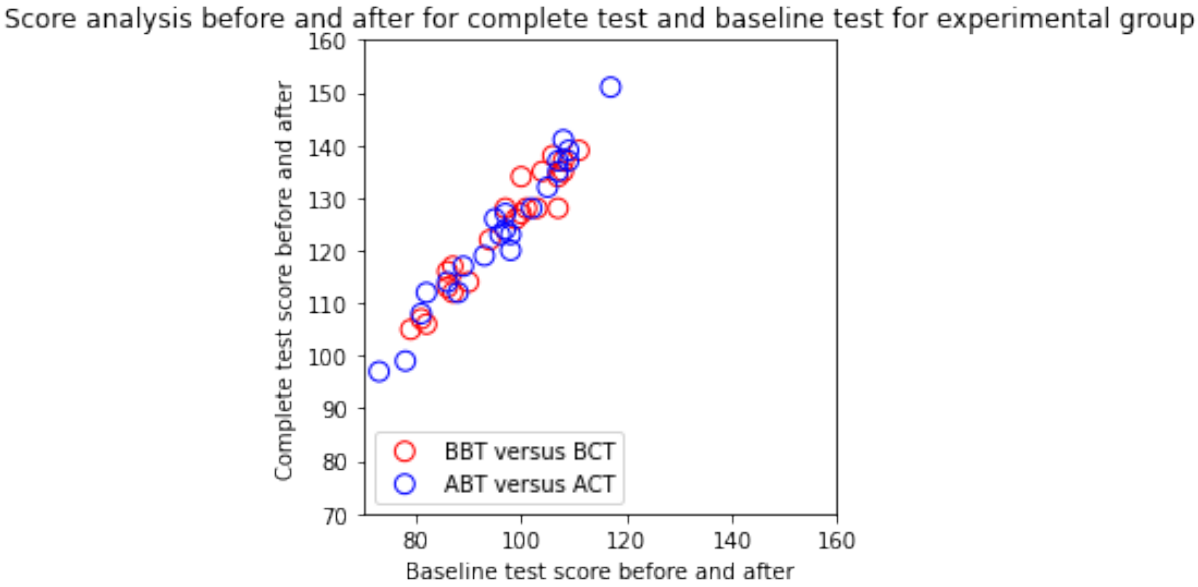

Figure 4.3: The results of only 9 questions that we implemented in our research in addition to QCAE. It shows the plot about where the empathy levels stands according to the participants.

In the test and control groups of complete and baseline empathy results, we observe no significant median change.

4.3 Scatter plot results

In figure 4.3, the x-axis represents the baseline score before and after the intervention and y-axis represents complete score before and after the intervention for the experimental group. The plot shows that there is a positive correlation in baseline and complete score for both before and after intervention. This means the complete score increases when there is an increase in baseline score and decreases when there is decrease in the baseline score.

Figure 4.4, shows the important time intervals involved in the entire experiment, from stage I till the end of the intervention, namely:

- We took the time variable into considerations when the participants answered and submitted the pre-test questionnaire.

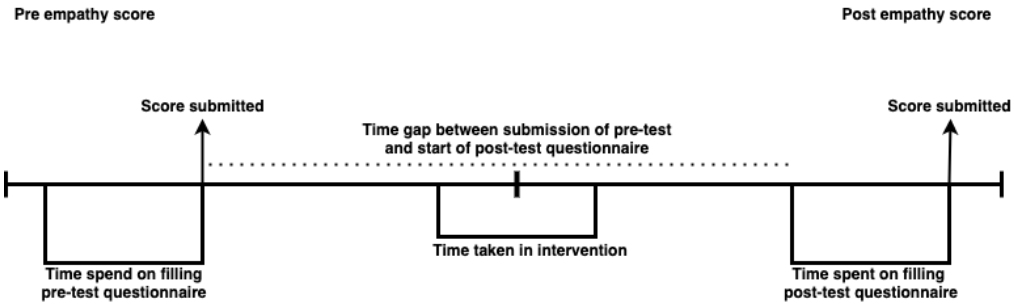

Figure 4.4: The time variable is taken into account for the data analysis using a scatter plot.

- We have also considered the time gap from the time of submission of the pre-test questionnaire to the start of the post-test questionnaire.
- The post-test starts after the participants go through intervention. We have also considered time spent by participants to fill the pre-test and post-test questionnaire for the analysis.
- Furthermore, we have also considered the time taken in intervention.

We have analyzed the above explained time variable with complete empathy score and differences in empathy score before and after the intervention.

We inspected if the time variable changes the empathy of the individual. In order to check whether the time spent by participants while filling the questionnaire affects our results or not, we put the data in a scatter plot.

In figure 4.5, the x-axis represents the time spent by participants while filling the questionnaire in minutes, and the y-axis represents the empathy score. The red, green, blue, and yellow plots represent the scores before experiment for test group, after experiment for test group, before experiment for control group, and after experiment for control group. From the figure 4.5, we can see that most of the participants have filled the questionnaire within the range of 15 minutes for both test and control group and most of the empathy score is above 110 for both the control and test group. We can also see that more participants in the test group have scored more than 125 compared to the control group. Apart from the people who have scored more than 125, we cannot figure out many differences in the test and control group results.

In figure 4.6, the x-axis and y-axis represents the empathy score of complete and baseline questionnaire before the intervention and after the intervention

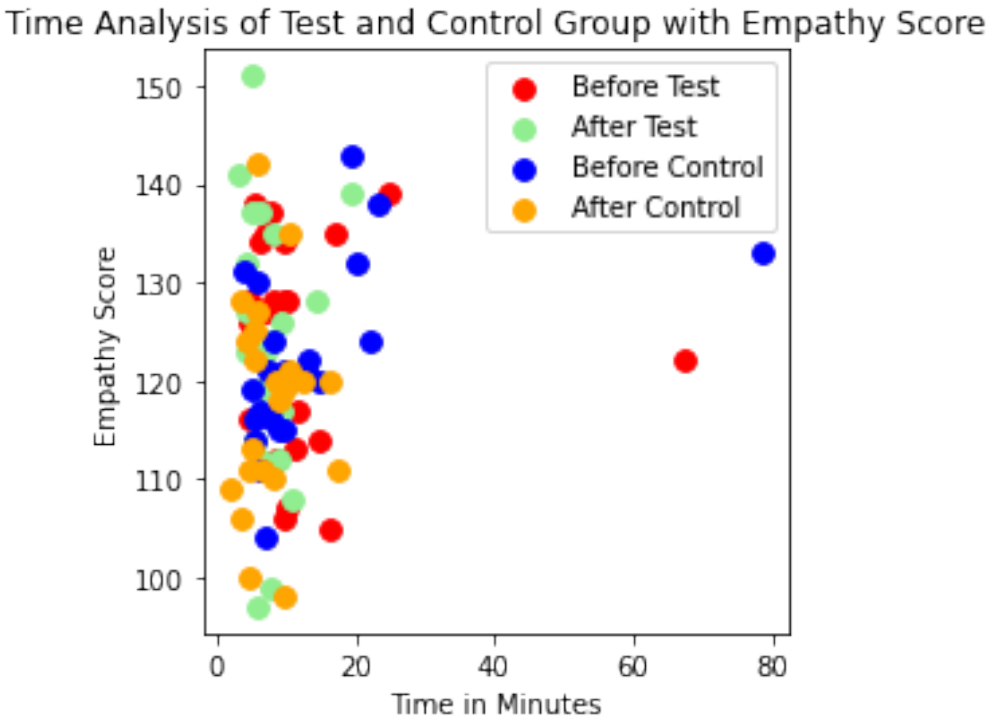

Figure 4.5: Figure shows the plot about where the empathy levels stand according to the participants

respectively. One observes a positive correlation for all the cases, before and after the intervention: more (respectively less) empathic participants tend to remain the most (respectively the least) participants after the intervention. In complete questionnaire, participants have scored more. This is because the complete questionnaire consists of 40 questions and baseline questionnaire of only 31 questions, due to which we can see high score in the scatter plot in the complete test results.

Furthermore, we also compared the interval between pre-test and post-test with differences in empathy score described in Table 4.2, the visualization can be seen in figure 4.7.

In figure 4.7, the x-axis represents an interval of time in days between pre and post-test, and the y-axis represents differences in empathy score. From figure 4.7, we can see that there is a tendency of scoring high in empathy on test group when considering the time differences, but the data is not enough

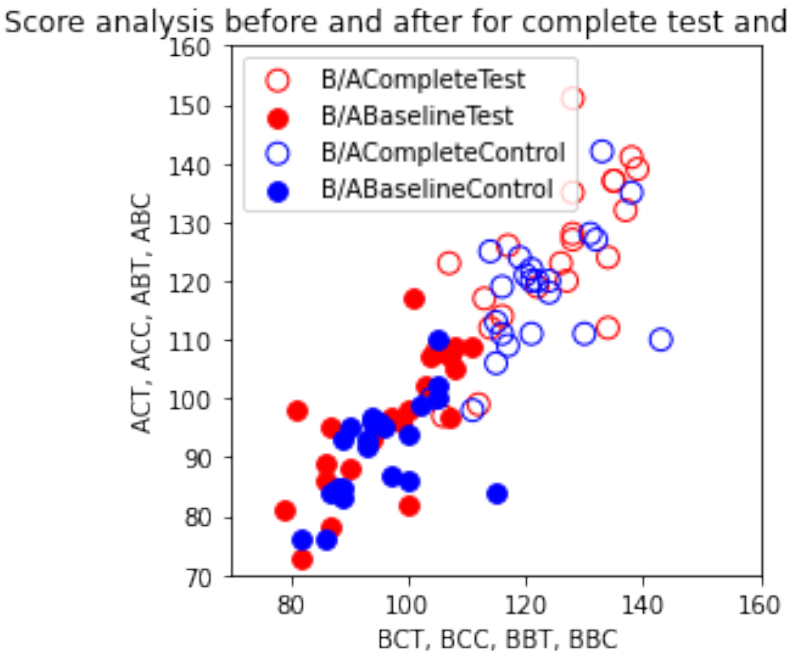

Figure 4.6: Complete and baseline empathy score before for test group and control group versus empathy score after for test and control group.

to claim the statement.

We have also compared the differences in empathy scores with the time spent in intervention for both test and control group. We can see the visualization in figure 4.8.

In figure 4.8, the x-axis represents time spent in intervention in minutes, and the y-axis represents empathy score differences. From figure 4.8, we can visualize that the time taken by test group participants is higher than the time taken by control group participants. The reason behind it was that the control group was supposed to see the image only and describe what they saw within a minute and we showed them four pictures within 4 minutes; on the other hand, for the test group, they were told to perform tasks by making dialogues through eye-tracker.

Between complete questionnaire and baseline questionnaire, the difference is small for both test and control groups. This score suggests that the test group task is showing a positive effect towards an increase in empathy. Also,

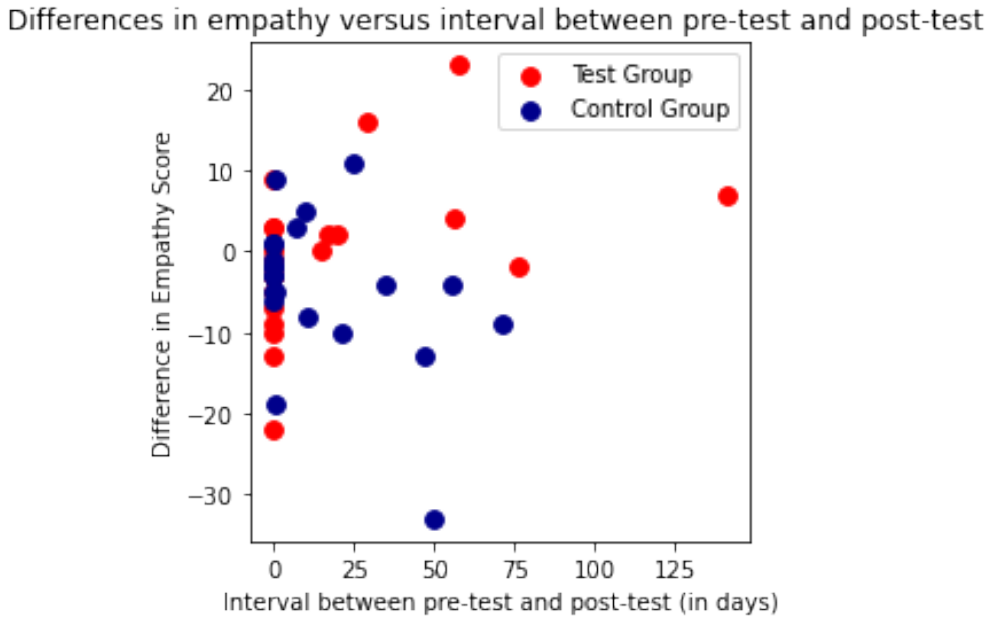

Figure 4.7: The figure shows the plot about where the empathy levels stand in differences in empathy according to the participants

the result achieved from complete questions does not deviate from the result achieved from the baseline questionnaire. However, to get a better statistical significance, we have to perform more experimental sessions.

4.4 Overview of main findings

- While comparing the test group with the control group; we can say that the intervention has an effect corresponding to the second set of values.
- Between complete and baseline tests, complete results seems to get more precision distinguishing variations, which indicates that the selection of questions to adapt the questionnaire by Prof. Dr. Renate Reniers, properly grasps the quantification of empathy in our experiment.
- From figure 4.2, we can see that in all cases there is a tendency to lose empathy. Participants had tendency to get bored or irritated. They may

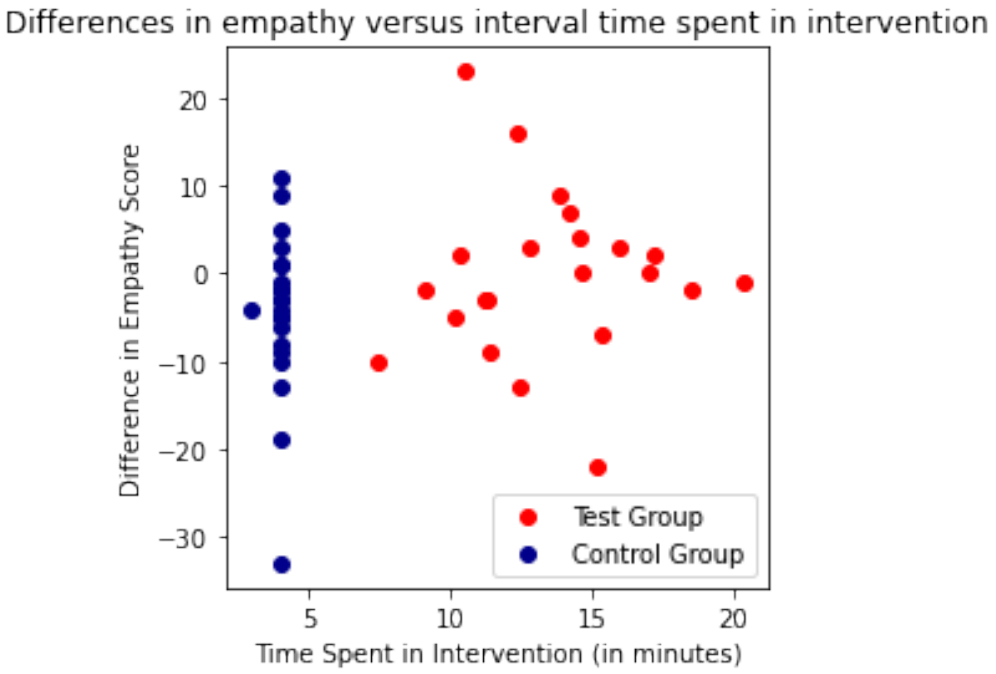

Figure 4.8: Time spent in the intervention versus differences in empathy score.

have been irritated because the second questionnaire (post-test questionnaire) was applied immediately after the intervention.

- Until ten minutes, there is no correlation between time and score but after 10 minutes, there seems to be some positive linear correlation.
- It seems that when both questionnaires (before and after experiment) take the same time, the full spectrum of scores is observed; if there is a difference in the time spent answering the questionnaires before and after, then there seems to be a tendency to lose empathy.
- For the test group, perhaps there is some increase of empathy with the time of intervention, but only till a certain time, beyond which empathy starts to be lost.

A quantitative analysis of how significant these results are, is presented in the next chapter.

Chapter 5
Discussions

5.1 Preliminary interpretation of main findings

From previous chapter and considering how only the difference of the score before and after the intervention (see table 4.2), we can summarize the findings of our research as follows:

- Firstly, intervention for promoting empathy seemed to have no significant effect.

- Secondly, the complete test seems to uncover more significant result towards the rejection of our null hypothesis.

In this chapter we evaluate these two findings in more detail.

The communication interface seems to have positive impact in participants about the struggle of non-verbal people. Most of the participants struggled while making the sentences with communication interface. Even though, we did not get significant result in empathy score, when asked, we found that the participants found it tough to communicate through eyes only. Participants found it difficult to communicate with our developed protocol and explained it, as a really difficult process for the people who rely on only eyes to communicate. Some of the details of experimental session is listed in Appendix D.

While comparing the results in the score of a complete and baseline test, there is a slight increase in the score, particularly for the complete test.

We also see that the extra nine questions added after seem to have impact in the increase in empathy. The complete test's empathy score tends to increase in order with the participants who have scored more in the baseline test also and vice versa. This means the overall score decreases when the baseline test score decreases. Thus, we can draw two conclusions. First, since the complete test correlates positive with the baseline test, i.e with already validated test, our complete test seems to be good tool to assess empathy in our experiment. Secondly, the complete test seems to better uncover possible differences in empathy between control and test group.

5.1.1 Increase in empathy towards specific group of people

While we compare test group score before and after the experiment, few scores fall above the third quartile of the box-plot for both before and after the test. Following the scoreline, we decided to take few interviews of participants who have scored more in the empathy test. We wanted to know the reason behind high scores in a test. Also,this would be beneficial to provide strength to our results.

Indeed, after the interview with the participants, we came to know that one of them was a nurse. During the interview, we also found out that she had some experience working with non-verbal people. People who tend to be involved in the medical field are supposed to be more empathic (Hojat, Fields, & Gonnella, 2003). Nurses, doctors who spend most of their career treating patients, frequently encounter people with difficulties and disabilities (Hojat et al., 2003). Nurses deal with people with different disabilities and are aware with the difficulties faced by them. So, scoring high in the empathy test in this case may be due to affection towards non-verbal people and previous experience and profile.

This participant mentioned, she is working as a midwife; she explained while she was doing her clinical practice. She was caring for one patient with locked-in syndrome for six months, and she knows how difficult it was for the patient to convey a message, even to ask for food. She also mentioned that there was so much misinterpretation in communication. So that may be the reason why she was more empathic than the other participant.

After getting such a response, we again wanted to take few more interviews. On the verge of taking interviews, we found another participant whose best friend survived a motorbike accident, and suffered life time paralysis. After coming to our lab, he realized that there are types of equipment made for people who cannot communicate. His reason behind scoring high empathy score can be due to his friend's incident and struggle.

After knowing the reason behind the increase in empathy, we can also hypothesize that empathy levels vary due to the situation and past incidents.

5.1.2 Reason of decrease in empathy

The complete questionnaire consisted of 40 questions, and participants were requested to do it twice before and after the experiment. Most of the partici-

pants said that they were tired after performing the task. We assume that the participants were uninterested in filling the post questionnaire, so they fill out the post-questionnaire without focusing on the questions. Most participants took less time to fill the questionnaire in post-experiment than the time taken in the pre-test questionnaire. Comparing the time taken to fill the questionnaire before and after the experiment also gives us some evidence that the participants showed less interest in going through the questionnaire again. The explained situation can be a reason for the decrease in empathy score after the intervention.

For the control group, the reason behind the unchanged empathy level might be less time consumed during the experiment. When we compare the time consumed by participants in the test and control group, the average time taken completing the control group task during intervention was 5 till 6 minutes, whereas, for the test group, the average time taken was approximately 25 minutes. Though the time taken was less, but the control group did not feel the stress in making dialogues like the test group, that might be the reason they were less irritated. At the same time, they did not experience the struggle about how difficult it was to make sentences or dialogue using eye-tracker and communication interface.

5.2 Quantitative assessment of experimental results

To evaluate the power needed for our study, since there was no significant level of change in empathy after the intervention, we realised there could be some possible limitations in our study.

We collected a sample of 44 participants, see table 5.2, 22 in each group control and experimental. We believe there can be some changes in the result if we recruit more participants. To better evaluate the significance of our results, we briefly address power analysis. We perform power analysis to find out the sample size that can get the significant result. Representing as (μ_c) and (μ_e) the means of questionnaire scores for the control and experimental groups, respectively, we calculated the estimated difference in mean by subtracting the average mean empathy score of control group (μ_c) by experimental group (μ_e), divided by pooled estimated value ($\tilde{\sigma}$). To get the pooled estimated value we needed the standard deviation of both control and experimental group denoted by (σ_e and σ_e). To perform the power analysis, we used our existing data to calculate the effect size (D), which quantifies how "certain" the null hypothesis

is ($\mu_c = \mu_e$). The effect size is given by:

$$D = \frac{|\mu_c - \mu_e|}{\tilde{\sigma}}\,, \tag{5.1}$$

The pooled estimated value ($\tilde{\sigma}$) in effect size is given by the weighted average of the standard deviations of each group (σ_e and σ_e):

$$\tilde{\sigma} = \sqrt{\frac{\sigma_c^2 + \sigma_e^2}{2}}\,, \tag{5.2}$$

where N_c and N_e are the sample sizes of each group.

	Baseline			**Complete**		
	μ	σ	N	μ	σ	N
Control	**91**	**8**	**22**	**117**	**11**	**22**
Experimental	**96**	**12**	**22**	**124**	**14**	**22**

Table 5.1: Summary of mean, standard deviation and number of samples for control and experimental group for both baseline and complete questionnaire.

From table 5.2, the values in μ, σ, N for both groups and both questionnaires we obtain $D = 0.49$ for the baseline questionnaire and $D = 0.56$ for the complete questionnaire. We used D to quantify the difference between the two questionnaire that is baseline and complete.

We assume that both groups have approximately the same size ($N_e \sim N_c$), since the distribution of participants was done in a way that the control groups has not more than one participant than the experimental group. In this case the size effect D can be used together with the Z-scores of a given significance level α and for a given power $1 - \beta$ can provide an estimate of the needed number of participants, given by

$$N_c \sim N_e \sim \frac{(Z_\alpha + Z_{1-\beta})^2}{D^2}\,. \tag{5.3}$$

	$\tilde{\sigma}$	D	$Z_{1-\beta}$	β	N for $power = 0.8$
Baseline	**10.2**	**0.49**	**0.33**	**37%**	**32**
Complete	**12.6**	**0.56**	**0.66**	**25%**	**26**

Table 5.2: Power analysis results.

While in this way we conclude that our sample sizes were too low to accept with reasonable certainty of the null hypothesis, we can also use it to estimate the probability of having a Type-II error, β, for the given sample sizes we used and the given averages and standard deviations we obtained in each group. Indeed, substituting in Eq. (5.3) the values of N_c, N_e and D and assuming a low probability of Type-I errors (e.g. $\alpha = 0.05$) we obtain a Z-score $Z_{1-\beta} = 0.33$ for the baseline questionnaire and $Z_{1-\beta} = 0.62$ for the complete questionnaire, which correspond to the probability of having Type-II errors $\beta = 37\%$ for the basline questionnaire and $\beta = 25\%$ for the complete questionnaire. For a significance level of $\alpha = 0.05$ (low probability of having Type-I errors) and a power of $1 - \beta = 0.8$ we have $Z_\alpha \sim 1.96$ and $Z_{1-\beta} \sim 0.84$, yielding $N_c = N_e = 32$ for the baseline questionnaire and $N_c = N_e = 26$ for the complete questionnaire.

The power analysis provides quantitative evidence in two aspects:

- For both questionnaires the power is below the usually accepted threshold of 80% to decision for a rejection of the null hypothesis.
- The power obtained for the complete test (75%) is considerably larger than the one for the baseline test (63%), and it is not far from the threshold of 80%.

For the implementation of power analysis we use the following code [1]:

```
# estimate sample size via power analysis
from statsmodels.stats.power import TTestIndPower
# parameters for power analysis
effect = 0.8
alpha = 0.05
power = 0.8
# perform power analysis
```

[1] https://machinelearningmastery.com/statistical-power-and-power-analysis-in-python/

```
analysis = TTestIndPower()
result = analysis.solve_power(effect, power=power,
nobs1=None, ratio=1.0, alpha=alpha)
print('Sample Size: %.3f' % result)
```

5.3 Possible limitations of experimental procedure and future critical remarks

Even if we reject the null hypothesis for the complete questionnaire case, based in a power of 75%, the numerical differences between mean score observed for control and test groups is small. One possible reason for that fact is that there was only one intervention that took place to promote empathy. More interventions can be made, where participants go through other interventions apart from our protocol. We may concluded that one intervention is not enough to promote empathy in individuals. One could have added other techniques such as bibliotherapy. Bibliotherapy can be used in order to train the verbal people about how non-verbal people feel while communicating with the help of stories and poems (Shechtman, 2008).

We also think that, the questionnaire that we used for getting the empathy results might not have been precise enough for our research. The answer from the questionnaire did not reflect a substantial increase in the empathy, but participants explained experiment to be difficult and they also mentioned living with conditions where people can only communicate with their eyes is really difficult. In this situation, we can try other questionnaires from the literature which are more related to our research. For example tests measuring the so called interpersonal reactivity index, which considers four subscales while building the questionnaire (De Corte et al., 2007):

Perspective taking, designed to adopt other's psychological positions involuntarily; Fantasy to access participants tendency to put themselves into the feelings of other people or fictitious characters in movies or plays; Empathic concern, to make feel themselves, individuals in a place of someone unable to do specific tasks, for example, who cannot communicate verbally; Personal distress, to make personally distressed people aware of the feeling of other people (De Corte et al., 2007).

To overcome the limitations, we came across some ideas which we think might help in promoting social inclusion. The points that we discussed as limitations in this chapter can be taken as inspiration to work in this field of

study to promote social inclusion for people with disabilities. While, ways to promote empathy can bring up lots of ideas in society, this study can help us get a clear understanding of the difficulties encountered by non-verbal people.

Chapter 6
Conclusion and future directions

To sum up, we developed a protocol to develop empathy in people who can communicate verbally towards people with special needs. With the protocol we addressed two research questions:

RQ1 How to develop empathy in verbal individuals towards individuals with special needs for communicating?

RQ2 Using an eye-tracker, how can such a framework for developing empathy be implemented in a way that can easily be reproduced?

The protocol seems to have minimal impact changing level of empathy in the test group than in the control group.

When we compare the power that we got for the baseline and complete questionnaire, still the questionnaire we developed have higher power than the the power we got from the baseline questionnaire, which means there is a possibility of rejecting the null hypothesis if we recruit more participants.

A possible explanation for this lack of effect is that one intervention is not enough to develop empathy in general. There are other strategies by which we can perform the experiments. A different approach could be bibliotherapy, exposure therapy. Still, by establishing an experimental protocol, this study has opened a lots of possibilities for research in the field of promoting social inclusion through augmented communication technology.

Indeed, this study opens a panoply of different possible ways to help promoting empathy towards people having special needs to communicate. One new way can be the exposure to the situation to promote empathy. For instance, the situation where the participants can see the struggle that non-verbal people have while communicating. The exposure therapy to promote social inclusion can be done through the use of VR virtual reality. Virtual reality is a technology that creates an artificial environment but seems like real world with the help of technology, such as, VR headsets, googles with gloves and screen. One can create an virtual environment, where people can be involved in others peoples sense of connection, where they get to experience the challenges faced by the people with disabilities through virtual reality (Schutte & Stilinović, 2017).

We can also create a kind of protocol focusing on subjective perception (Endsley, 2016). We develop a protocol by getting the opinions of non-verbal people. The protocol may lead to a more precise framework in developing

empathy by following the same method with the participants that we performed. For instance, by developing a protocol by exercising what non-verbal people think and manage to communicate in society.

Human-inspired artificial intelligence could be another option. We believe understanding human emotions can be helpful making the protocol more effective for designing a framework to promote social inclusion Wang, Yang, Abdul, and Lim (2019).

We can also use our protocol in schools, where it could be applied for longer period of time among the students. Because of that the protocol could likely be more effective.

In this way we were able to complete our thesis and have a understanding of different kinds of augmented communication technologies, and how they can help in making an inclusive society.

Appendix A
Questionnaire of cognitive and affective empathy

In this appendix, we list all questions composing the Questionnaire of Cognitive and Affective Empathy from Reniers et al. (2011) with permission and added the additional questions that we established for the complete questionnaire. These are the questions that will be used in the experiment presented in this thesis.

The questions are:

1. I sometimes find it difficult to see things from the 'other guy's' point of view.
2. I am usually objective when I watch a film or play, and I don't often get completely caught up in it.
3. I try to look at everybody's side of a disagreement before I make a decision.
4. I sometimes try to understand my friends better by imagining how things look from their perspective.
5. When I am upset at someone, I usually try to 'put myself in his shoes' for a while.
6. Before criticising somebody, I try to imagine how I would feel if I was in their place.
7. I often get emotionally involved with my friends' problems.
8. I am inclined to get nervous when others around me seem to be nervous.
9. People I am with have a strong influence on my mood.
10. It affects me very much when one of my friends seems upset.
11. I often get deeply involved with the feelings of a character in a film, play or novel.
12. I get very upset when I see someone cry.

13. I am happy when I am with a cheerful group and sad when the others are glum.
14. It worries me when others are worrying and panicky.
15. I can easily tell if someone else wants to enter a conversation.
16. I can pick up quickly if someone says one thing but means another.
17. It is hard for me to see why some things upset people so much.
18. I find it easy to put myself in somebody else's shoes.
19. I am good at predicting how someone will feel.
20. I am quick to spot when someone in a group is feeling awkward or uncomfortable.
21. Other people tell me I am good at understanding how they are feeling and what they are thinking.
22. I can easily tell if someone else is interested or bored with what I am saying.
23. Friends talk to me about their problems as they say that I am very understanding.
24. I can sense if I am intruding, even if the other person does not tell me.
25. I can easily work out what another person might want to talk about
26. I can tell if someone is masking their true emotion.
27. I am good at predicting what someone will do.
28. I can usually appreciate the other person's viewpoint, even if I do not agree with it.
29. I usually stay emotionally detached when watching a film.
30. I always try to consider the other fellow's feelings before I do something.
31. Before I do something I try to consider how my friends will react to it.

32. **I feel frustrated when I cannot communicate my ideas.**
33. **I complete the sentences of other people when they find it difficult to communicate their ideas to help them.**
34. **If someone is suffering from stutter, I prefer speaking rather than texting.**
35. **I try to be patient when people speak with me very slowly.**
36. **In a situation, when there are handicapped students in a classroom, the pace of the class should be adapted to the handicapped student.**
37. **In a situation, when there are only one handicapped student in a classroom, the pace of the class should be adapted to him/her independently of how many students are in the class.**
38. **I feel impatient when people communicate slowly, so I complete other people's sentences to make the communication more efficient.**
39. **I believe that the handicapped students should catch the pace of the group in a classroom to not delay the progress of other students.**
40. **I feel left out (or excluded) when I cannot participate in a conversation.**

Appendix B
Consent Form

Appendices

Appendix 1: Consent Form

Consent form

The current research is part of the Master Thesis in Universal Design of ICT,

Department of Computer Science, Faculty of Technology, Art and Design,

Oslo Metropolitan University

Student's Information

Name: Samip Bhurtel

Contact details: 96705443 (samipbhurtel7@gmail.com)

Title of the Study: Strategies to promote social inclusion of user of augmented communication technologies.

Purpose of the Study: The aim of this study is to increase empathy in "normal" individuals with respect to individuals having special needs for communicating, through the help of eye tracker

Procedures: In order to participate in this study, participants will have to participate in first meeting. In this meeting, you will have to go through a series of questionnaires and have to answer the questions. This will be done in order to measure how empathic you are in respect to the non-verbal people. After that we will train the individual through the help of eye tracker and ask the similar type of questions again, after or while using the eye tracker. This whole procedure will approximately take 20-30 minutes.

Risks/Discomfort: There is no known medical risks involved in participating this research. If you would feel any discomfort or fatigue during the participation, you can take rest. You will provide several opportunities to take rest in case of discomfort, and additional breaks can also be taken.

Benefit: Your participation may contribute to the result of the study that the eye tracker is helpful to increase empathy in "normal" individuals with respect to individuals having special need for communicating.

Alternative to the participation: Participation in this study is voluntary and you have the right to withdraw or discontinue participation at any time. And withdrawing or discontinuing participation in the study will not have any consequences to you. You have the right to demand for your data to be deleted from the study.

Confidentiality: All the data collected during the study period will be kept highly confidential. There will be no data sharing to any governmental organizations.

I have read and understand the information about the study and all the information in this form is explained to me and I am willing to participate

_____ _____

Signature Date

Supervisor: Gustavo Mello

Contact information:

E-mail:

Visiting address: Pilestredet 50, Oslo

Co supervisor: Pedro Lind

Contact information:

E-mail:

Visiting address: Pilestredet 50, Oslo

Appendix C
Focus group meeting minutes

We gathered together a focus group to talk about the views about the questions that we made for the experiment, for my masters thesis Strategies to promote social inclusion of user of augmented communication technologies.

We invited three people who are students of university pursuing their master's degree, in the field of health and information technology. Participants agreed to take part in the focus group discussion and present their views. The focus group discussion was held in the AI lab at Oslo Metropolitan University.

Firstly we welcomed them and after that, we directly jumped to 10 questions that were prepared for the focus group discussion. The points of each question are mentioned below which were discussed in the meeting.

1. I feel frustrated when I cannot communicate my ideas.

   Everyone agreed with the question.

2. I complete the sentences of other people when they speak slowly.

   As per the discussion, participants found the question ambiguous. The decision was made to break down the question.

3. I prefer speaking rather than texting.

   Question should include the context in which disability is involved.

4. I feel impatient when other people communicate with me very slowly.

   Instead of this question there was a suggestion of questions discussed were, i) I adjust my pace of communication according to need. ii) I try to be patient when people speak with me very slowly.

5. In a classroom the pace of the class should be adapted to the handicapped student.

   There was a long discussion on this question as everyone shared their views differently. We decided to discuss it later as the time was running out.

   After that we decided to eliminate question 6 and 7 and 8 . The decision was made to make these questions 1 out of 3.

6. The pace should be adapted even if the classroom has only one handicapped student.

7. The pace should be adapted independent of how slow the class for the other students is.

8. The handicapped student should catch the pace of the group.

9. I feel left out when I cannot participate in a conversation. For this question everyone said it's valid. There was a discussion about people talking in norwegian instead of using english as a universal language in a workplace or in restaurants.

10. I don't like to speak when the conversation is irrelevant. One of the participants described it as depending on who the people are.

The discussion concluded nicely and we decided to make adjustments to some of the questions.

Appendix D
Experimental sessions

D.0.1 Test group sessions

1. Participant 1: Since, it was a first session, it took about 45 minutes to complete the whole session. During the session there was no any disturbance, so it went really well.

   After completing the session, participant described the experiment to be tiring to eyes. According to the participant, it was quite difficult for her to write messages through the help of e-tran board. It took time for her to convey the message through the e-tran board because it was completely new for her. As explained by the participant, it was quite stressful to write each letter with eye gaze while building the dialogue.

   She also mentioned that the letters were jumbled up which made it hard to find and write each letter. She stated that if its the only way to communicate, then the individual should get proper training before use.

2. Participant 3: The overall time taken by the participant was 30 minutes. The impression of the participant in using the device was that, it was difficult to use. The participant, initially faced some challenges while using this device and found it difficult to get command of the communication board i.e. e-tran. The same problem happened to the participant as told by participant 1: it was quite confusing and his eyes were tired after the experimental session. Participant further explained that, though it is difficult to be precise while building messages, it is a really useful process, if it is the only way to communicate.

3. Participant 5: It was really difficult for the participant to perform the communication through the eye-tracker. His eyes became blurry after searching for the letters in the very initial phase of the experiment. The experience for him while using the devices was not so good. He mentioned that the device would only be useful for someone who cannot move their hands and legs. He suggested that we can replace letters by pictures.

4. Participant 7: Participant felt blurry eyes while communicating through the interface. He told that the bright screen caused strain in his eyes which lead to blurriness. He also felt color confusion while using the interface. Although, he found it difficult to communicate through the interface, he

explained that the device would be really useful for the people who rely on it for daily activities. He went by the phrase, something is better than nothing.

5. Participant 9: For this participant, color combination in communication board was hard to determine and the position of letters was difficult to identify. She faced some challenges doing the word separation.

   She explained that it is a good interface, if it's the only possible way to communicate.

6. Participant 11: Participant was confused in building the messages. To make a second dialog, it took her three takes. She was confused with the color and also how words were formed using e-tran board even after the training.

   She was so impressed about the device as it was the first time she was using it. Later on, she also mentioned that the device is a boon for the people who cannot communicate verbally.

D.0.2 Control group sessions

1. Participant 2: Participant 2 did the control task. After the explanation of everything, i.e the uses of E-tran board and eye-tracker, we told the participant to look at the fig 3.4 for a minute. We repeated the experiment three times more for one minute each. Participant described the image as fuzzy, with colors in it. He told that he couldn't track anything in the image.

2. Participant 4: Participant was asked to look at the image shown in monitor for multiple times. And was asked what he saw. He also found the image to be confusing and did not have any specific meaning behind.

All the remaining participants described the image, same as described by participant number 2 and 4. They stated it was blurry and they didn't figured out anything. The control group experiment was carried out in a similar manner for all the participants.

Appendix E
NSD format consent

Are you interested in taking part in the research project

"(Strategies to promote social inclusion of user of augmented communication technologies)"?

This is an inquiry about participation in a research project where the main aim *of this study is to test a framework for developing empathy in people in general, towards individuals having special needs for communicating. It does not involve participants with some sort of disability or special need. Participants are separated in two groups randomly (test and control group). The training process is done through the help of eye-tracking systems and communication interfaces. We are collecting the data by requesting the participants to fill in one questionnaire in nettskjema system. The questionnaire is composed of 40 questions (see Attachment 1 in consent. The questionnaire will be filled twice by each participant, before and after experimental intervention. We are doing it twice, in order to compare the empathy level of a person i.e before and after the intervention. The interventions concern a session of 30 minutes where the participant will experience how to communicate with the expert through an eye-tracking device*. In this letter we will give you information about the purpose of the project and what your participation will involve.

Purpose of the project
The aim of this study is to increase empathy in verbal individuals with respect to individuals having special needs for communicating, through the help of eye tracker. The study will help identify whether the protocol that we have developed helps improving the empathy level of the individuals.

Who is responsible for the research project?
Oslo Metropolitan University is the institution responsible for the project.

Why are you being asked to participate?
The participants are students from different faculties and departments at Oslo Metropolitan University, UiO and NTNU. There will be no exact knowledge about the age distribution of the sample collected. The information below about age is an estimate. The Participants provides email, and full name.

What does participation involve for you?

- *If you chose to take part in the project, this will involve that you fill in an online survey twice that is before the intervention and after the intervention. The invention takes about 20-30 minutes. No other person is aware of the participants' identity. The personal data collected from each participant includes name, email address of the participants and record of the eye-gaze trajectory of the participants in the control group. Moreover, no data showing iris patterns (such as pictures or videos) will be collected, when using the eye-tracker. The eye-tracker only flags the position of the eye and eye-gaze. It will take approx. 45 minutes. The survey includes questions about* ***questionnaire of cognitive and affective empathy****. Your answers will be recorded electronically.*

Participation is voluntary: Participation in the project is voluntary. If you chose to participate, you can withdraw your consent at any time without giving a reason. All information about you will then be

made anonymous. There will be no negative consequences for you if you chose not to participate or later decide to withdraw.

Your personal privacy – how we will store and use your personal data
We will only use your personal data for the purpose(s) specified in this information letter. We will process your personal data confidentially and in accordance with data protection legislation (the General Data Protection Regulation and Personal Data Act).

- *None of the information will be shared beyond the researcher and nothing will be attributed by name. The knowledge generated from this research will be shared with you and your community before it is made available to the general public. Each participant will be provided with a summary of the results.*
- *So long as participants can be identified in the collected data, they have the right to: access the personal data that is being processed about them, request that their personal data is deleted or that incorrect personal data about them is corrected/rectified, receive a copy of personal data (data portability), and to send a complaint to the Data Protection Officer or The Norwegian Data Protection Authority regarding the processing of your personal data.*

- *We will replace the participants name and email with a code that is **participants id**. The list of names, contact details and respective codes will be stored separately from the rest of the collected data, the name and email address of the participants is stored in nettsjema. The data collected for the project is accessible to me (Samip Bhurtel), UIO/nettskjema as data processor and supervisors (Pedro lind and Gustavo mello). The personal data collected from each participant includes name, email address of the participants and record of the eye-gaze trajectory of the participants in the control group. Moreover, no data showing iris patterns (such as pictures or videos) will be collected, when using the eye-tracker. The eye-tracker only flags the position of the eye and eye-gaze.*

Participants will not be recognizable in publications. We are going to analyse only the quantitative data received from the participants and some qualitative data explained by the participants but the participants is not identifiable.

What will happen to your personal data at the end of the research project?
The project is scheduled to end *15th of May 2021. All the personal data will be destroyed after the data analysis that will do before the submission of project dated 15th of May 2021.*

Your rights
So long as you can be identified in the collected data, you have the right to:

- access the personal data that is being processed about you
- request that your personal data is deleted
- request that incorrect personal data about you is corrected/rectified
- receive a copy of your personal data (data portability), and
- send a complaint to the Data Protection Officer or The Norwegian Data Protection Authority regarding the processing of your personal data**.**

What gives us the right to process your personal data?
We will process your personal data based on your consent.

Based on an agreement with *Oslo Metropolitan University*, NSD – The Norwegian Centre for Research Data AS has assessed that the processing of personal data in this project is in accordance with data protection legislation.

Where can I find out more?
If you have questions about the project, or want to exercise your rights, contact:

- *Oslo Metropolitan University* via *Samip Bhurtel Email: s339954@oslomet.no, Phone: +4796705443*
- *Supervisor: Pedro G Lind , Email: pedrolin@oslomet.no,*
- *Supervisor: Gustavo Mello, Email: gustavom@oslomet.no,*
- Our Data Protection Officer: *Ingrid jacobsen, Email : ingrid.jacobsen@oslomet.no*
- NSD – The Norwegian Centre for Research Data AS, by email: (personverntjenester@nsd.no) or by telephone: +47 55 58 21 17.

Yours sincerely,

Project Leader
(Samip Bhurtel)

Consent form

I have received and understood information about the project *Strategies to promote social inclusion of user of augmented communication technologies* and have been given the opportunity to ask questions. I give consent to participate in two online survey and one experiment.

I give consent for my personal data to be processed until the end date of the project, approx *15 May 2021*.

(Signed by participant, date)

Appendix F
NSD assessment

NSD NORSK SENTER FOR FORSKNINGSDATA

NSD sin vurdering

Prosjekttittel

Strategies to promote social inclusion of user of augmented communication technologies.

Referansenummer

119586

Registrert

09.02.2021 av Samip Bhurtel - s339954@oslomet.no

Behandlingsansvarlig institusjon

OsloMet – storbyuniversitetet / Fakultet for teknologi, kunst og design / Institutt for informasjonsteknologi

Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)

Gustavo de Melo , gustavom@oslomet.no, tlf: +4767237964

Type prosjekt

Studentprosjekt, masterstudium

Kontaktinformasjon, student

Samip Bhurtel, s339954@oslomet.no, tlf: 96705443

Prosjektperiode

12.08.2020 - 15.05.2021

Status

28.04.2021 - Vurdert

Vurdering (1)

28.04.2021 - Vurdert

Our assessment is that the processing of personal data in this project will comply with data protection legislation, so long as it is carried out in accordance with the Notification Form and attachments, dated 28.04.21, as well as in correspondence with NSD and our assessment. Everything is in place for the processing to begin.

TYPE OF DATA AND DURATION
The project will be processing general categories of personal data, and special categories of personal data regarding health until 15.05.2021.

LEGAL BASIS

The project will gain consent from data subjects to process their personal data. We find that consent will meet the necessary requirements under art. 4 (11) and 7, in that it will be a freely given, specific, informed and unambiguous statement or action, which will be documented and can be withdrawn.

The legal basis for processing general categories of personal data is therefore consent given by the data subject, cf. the General Data Protection Regulation art. 6.1 a).

The legal basis for processing special categories of personal data is explicit consent given by the data subject, cf. art. 9.2 a), cf. the Personal Data Act § 10, cf. § 9 (2).

PRINCIPLES RELATING TO PROCESSING PERSONAL DATA
NSD finds that the planned processing of personal data will be in accordance with the principles under the General Data Protection Regulation regarding:

- lawfulness, fairness and transparency (art. 5.1 a), in that data subjects will receive sufficient information about the processing and will give their consent
- purpose limitation (art. 5.1 b), in that personal data will be collected for specified, explicit and legitimate purposes, and will not be processed for new, incompatible purposes
- data minimisation (art. 5.1 c), in that only personal data which are adequate, relevant and necessary for the purpose of the project will be processed
- storage limitation (art. 5.1 e), in that personal data will not be stored for longer than is necessary to fulfil the project's purpose

THE RIGHTS OF DATA SUBJECTS
NSD finds that the information that will be given to data subjects about the processing of their personal data will meet the legal requirements for form and content, cf. art. 12.1 and art. 13.

Data subjects will have the following rights in this project: access (art. 15), rectification (art. 16), erasure (art. 17), restriction of processing (art. 18), notification (art. 19) and data portability (art. 20).

We remind you that if a data subject contacts you about their rights, the data controller has a duty to reply within a month.

FOLLOW YOUR INSTITUTION'S GUIDELINES
NSD presupposes that the project will meet the requirements of accuracy (art. 5.1 d), integrity and confidentiality (art. 5.1 f) and security (art. 32) when processing personal data.

UIO (Nettskjema) is a data processor for the project. NSD presupposes that the processing of personal data by a data processor meets the requirements under the General Data Protection Regulation arts. 28 and 29.

To ensure that these requirements are met you must follow your institution's internal guidelines and/or consult with your institution (OsloMet).

NOTIFY CHANGES
If you intend to make changes to the processing of personal data in this project it may be necessary to notify NSD. This is done by updating the information registered in the Notification Form. On our website we explain which changes must be notified. Wait until you receive an answer from us before you carry out the changes.

FOLLOW-UP OF THE PROJECT
NSD will follow up the progress of the project at the planned end date in order to determine whether the processing of personal data has been concluded in accordance with what is documented in the Notification Form.

Good luck with the project!
Contact person at NSD: Inga Brautaset

References

Alves, C. C. d. F., Monteiro, G. B. M., Rabello, S., Gasparetto, M. E. R. F., & Carvalho, K. M. d. (2009). Assistive technology applied to education of students with visual impairment. Revista Panamericana de Salud Pública, 26, 148–152.

Arai, K., & Mardiyanto, R. (2011). Eye-based hci with full specification of mouse and keyboard using pupil knowledge in the gaze estimation. In 2011 eighth international conference on information technology: New generations (pp. 423–428).

Aslin, R. N. (2012). Infant eyes: A window on cognitive development. Infancy, 17(1), 126–140.

Barfoot, K. M., Matthew, M., & Callaway, A. J. (2012). Combined eeg and eye-tracking in sports skills training and performance analysis. In World congress of performance analysis of sport ix university of worcester.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. Journal of autism and developmental disorders, 34(2), 163–175.

Bauer, G., Gerstenbrand, F., & Rumpl, E. (1979). Varieties of the locked-in syndrome. Journal of neurology, 221(2), 77–91.

Bulling, A., Roggen, D., & Tröster, G. (2009). Wearable eog goggles: eye-based interaction in everyday environments. In Chi'09 extended abstracts on human factors in computing systems (pp. 3259–3264).

Buvanswari, B., & Reddy, T. K. (2015). Eye scrutinized wheel chair for people affected with tetraplegia. International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), 5(2).

Calvo, A., Chiò, A., Castellina, E., Corno, F., Farinetti, L., Ghiglione, P., ... Vignola, A. (2008). Eye tracking impact on quality-of-life of als patients. In International conference on computers for handicapped persons (pp. 70–77).

Carpenter, W. T., Kirkpatrick, B., & Buchanan, R. W. (1990). Conceptual approaches to the study of schizophrenia. In Recent advances in schizophrenia (pp. 95–113). Springer.

Caruso, D. R., & Mayer, J. D. (1998). A measure of emotional empathy for adolescents and adults.

Cheng, D., & Vertegaal, R. (2004). An eye for an eye: a performance evaluation comparison of the lc technologies and tobii eye trackers. In Proceedings of the 2004 symposium on eye tracking research & applications (pp. 61–

61).

Chinnery, B. (1990). Disabled people get the message: Non-verbal clues to the nature of social work. Practice, 4(1), 49–55.

Cohen, D., & Strayer, J. (1996). Empathy in conduct-disordered and comparison youth. Developmental psychology, 32(6), 988.

Crane, H. D., & Steele, C. M. (1985). Generation-v dual-purkinje-image eyetracker. Applied optics, 24(4), 527–537.

Dadds, M. R., Hunter, K., Hawes, D. J., Frost, A. D., Vassallo, S., Bunn, P., ... El Masry, Y. (2008). A measure of cognitive and affective empathy in children using parent ratings. Child psychiatry and human development, 39(2), 111–122.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. Journal of personality and social psychology, 44(1), 113.

Davis, M. H., et al. (1980). A multidimensional approach to individual differences in empathy.

Davis, N. J. (1992). Teaching about inequality: Student resistance, paralysis, and rage. Teaching Sociology, 232–238.

Decety, J., & Fotopoulou, A. (2015). Why empathy has a beneficial impact on others in medicine: unifying theories. Frontiers in behavioral neuroscience, 8, 457.

De Corte, K., Buysse, A., Verhofstadt, L. L., Roeyers, H., Ponnet, K., & Davis, M. H. (2007). Measuring empathic tendencies: Reliability and validity of the dutch version of the interpersonal reactivity index. Psychologica Belgica, 47(4), 235–260.

Ekstein, R., & Caruth, E. (1969). Levels of verbal communication in the schizophrenic child's struggle against, for, and with the world of objects. The Psychoanalytic study of the child, 24(1), 115–137.

Elbert, T., Lutzenberger, W., Rockstroh, B., & Birbaumer, N. (1985). Removal of ocular artifacts from the eeg—a biophysical approach to the eog. Electroencephalography and clinical neurophysiology, 60(5), 455–463.

Endsley, M. R. (2016). Designing for situation awareness: An approach to user-centered design. CRC press.

Evans, D. G., Drew, R., & Blenkhorn, P. (2000). Controlling mouse pointer position using an infrared head-operated joystick. IEEE Transactions on rehabilitation engineering, 8(1), 107–117.

Field, M. J., Jette, A. M., of Medicine (US) Committee on Disability in America,

I., et al. (2007). Transportation patterns and problems of people with disabilities. In The future of disability in america. National Academies Press (US).

Furman, R. (2005). Using poetry and written exercises to teach empathy. Journal of Poetry Therapy, 18(2), 103–110.

Grisold, W. (2004). Locked-in syndrome. Palliative Care in Neurology, 69, 135.

Gutiérrez, A., & Miravet, D. (2016). The determinants of tourist use of public transport at the destination. Sustainability, 8(9), 908.

Hojat, M., Fields, S. K., & Gonnella, J. S. (2003). Empathy: an np/md comparison. The Nurse Practitioner, 28(4), 45–47.

Isokoski, P. (2000). Text input methods for eye trackers using off-screen targets. In Proceedings of the 2000 symposium on eye tracking research & applications (pp. 15–21).

Janssen, J. H. (2012). A three-component framework for empathic technologies to augment human interaction. Journal on Multimodal User Interfaces, 6(3-4), 143–161.

Johnson, J. A., Cheek, J. M., & Smither, R. (1983). The structure of empathy. Journal of personality and social psychology, 45(6), 1299.

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the basic empathy scale. Journal of adolescence, 29(4), 589–611.

Kerr, W. A., & Speroff, B. J. (1954). Validation and evaluation of the empathy test. The Journal of General Psychology, 50(2), 269–276.

Kim, J., & Lee, S. J. (2010). Reliability and validity of the korean version of the empathy quotient scale. Psychiatry investigation, 7(1), 24.

Kim, J., Park, H., & Ghovanloo, M. (2012). Tongue-operated assistive technology with access to common smartphone applications via bluetooth link. In 2012 annual international conference of the ieee engineering in medicine and biology society (pp. 4054–4057).

Klaussen, T., & Rasmussen, L.-M. P. (2013). Psychometric validity and reliability of the social skills improvement system-rating scales (ssis-rs) (Unpublished master's thesis). Universitetet i Tromsø.

Krohn, O., Varankian, V., Lind, P., & Mello, G. (2020). Implementation of an inexpensive eye-tracking system for educational purposes.

Kupetz, D., Wentzell, S., & BuSha, B. (2010). Head motion controlled power wheelchair. In Proceedings of the 2010 ieee 36th annual northeast bioengineering conference (nebec) (pp. 1–2).

Kupetz, M. (2014). Empathy displays as interactional achieve-

ments—multimodal and sequential aspects. Journal of Pragmatics, 61, 4–34.

Liu, P., Sun, J., Zhang, W., & Li, D. (2020). Effect of empathy trait on attention to positive emotional stimuli: evidence from eye movements. Current Psychology, 1–11.

Majaranta, P., & Bulling, A. (2014). Eye tracking and eye-based human–computer interaction. In Advances in physiological computing (pp. 39–65). Springer.

Majaranta, P., & Räihä, K.-J. (2007). Text entry by gaze: Utilizing eye-tracking. Text entry systems: Mobility, accessibility, universality, 175–187.

Matthews, D. E., & Farewell, V. T. (2015). Using and understanding medical statistics. Karger Medical and Scientific Publishers.

Mehrabian, A. (1996). Manual for the balanced emotional empathy scale (bees). Available from Albert Mehrabian, 1130.

Miskelly, F. G. (2001). Assistive technology in elderly care. Age and ageing, 30(6), 455–458.

nettskjema. (2020). nettskjema. Retrieved 2020-11-10, from `https://www.uio.no/tjenester/it/adm-app/nettskjema/`

Nielsen, J., & Pernice, K. (2010). Eyetracking web usability. New Riders.

Ostle, B., et al. (1963). Statistics in research. Statistics in research.(2nd Ed).

Paolo Senese, V., De Nicola, A., Passaro, A., & Ruggiero, G. (2018). The factorial structure of a 15-item version of the italian empathy quotient scale. European Journal of Psychological Assessment, 34(5), 344.

Prabhu, V., & Prasad, G. (2011). Designing a virtual keyboard with multi-modal access for people with disabilities. In 2011 world congress on information and communication technologies (pp. 1133–1138).

Reniers, R. L., Corcoran, R., Drake, R., Shryane, N. M., & Völlm, B. A. (2011). The qcae: A questionnaire of cognitive and affective empathy. Journal of personality assessment, 93(1), 84–95.

Reynolds, R. G., & Peng, B. (2005). Cultural algorithms: computational modeling of how cultures learn to solve problems: an engineering example. Cybernetics and Systems: An International Journal, 36(8), 753–771.

Robinson, D. A. (1963). A method of measuring eye movemnent using a scieral search coil in a magnetic field. IEEE Transactions on bio-medical electronics, 10(4), 137–145.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothe-

sis. Psychonomic bulletin & review, 16(2), 225–237.
Sasson, N. J., & Elison, J. T. (2012). Eye tracking young children with autism. JoVE (Journal of Visualized Experiments)(61), e3675.
Scherer, M. J. (2004). Connecting to learn: Educational and assistive technology for people with disabilities. American Psychological Association.
Schutte, N. S., & Stilinović, E. J. (2017). Facilitating empathy through virtual reality. Motivation and emotion, 41(6), 708–712.
Shaw, L. R., Chan, F., & McMahon, B. T. (2012). Intersectionality and disability harassment: The interactive effects of disability, race, age, and gender. Rehabilitation Counseling Bulletin, 55(2), 82–91.
Shechtman, Z. (2002). Cognitive and affective empathy in aggressive boys: Implications for counseling. International Journal for the Advancement of Counselling, 24(4), 211–222.
Shechtman, Z. (2008). Treating child and adolescent aggression through bibliotherapy. Springer.
Slykhuis, D. A., Wiebe, E. N., & Annetta, L. A. (2005). Eye-tracking students' attention to powerpoint photographs in a science education setting. Journal of Science Education and Technology, 14(5-6), 509–520.
Smith, E., & Delargy, M. (2005). Locked-in syndrome. Bmj, 330(7488), 406–409.
Spreng, R. N., McKinnon, M. C., Mar, R. A., & Levine, B. (2009). The toronto empathy questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. Journal of personality assessment, 91(1), 62–71.
Swift, S. M. (2012). Low-tech, eye-movement-accessible aac and typical adults.
Tobii. (2020). Tobii eye-tracker guidelines. Retrieved 2020-05-27, from `https://www.tobiipro.com/\protect\discretionary{\char\hyphenchar\font}{}{}siteassets/tobii-pro/brochures/\protect\discretionary{\char\hyphenchar\font}{}{}tobiipro-x3-120-brochure.pdf/\protect\discretionary{\char\hyphenchar\font}{}{}?v=1.0`
Touchette, P. E., MacDonald, R. F., & Langer, S. N. (1985). A scatter plot for identifying stimulus control of problem behavior. Journal of applied behavior analysis, 18(4), 343–351.
Vallat, R. (2018). Pingouin: statistics in python. Journal of Open Source Software, 3(31), 1026.
van der Merwe, E., & Alant, E. (2004). Associations with minspeak™ icons.

Journal of communication disorders, 37(3), 255–274.

Wang, D., Yang, Q., Abdul, A., & Lim, B. Y. (2019). Designing theory-driven user-centric explainable ai. In Proceedings of the 2019 chi conference on human factors in computing systems (pp. 1–15).

Williamson, D. F., Parker, R. A., & Kendrick, J. S. (1989). The box plot: a simple visual method to interpret data. Annals of internal medicine, 110(11), 916–921.

Wispé, L. (1986). The distinction between sympathy and empathy: To call forth a concept, a word is needed. Journal of personality and social psychology, 50(2), 314.

Zhang, X., Kulkarni, H., & Morris, M. R. (2017). Smartphone-based gaze gesture communication for people with motor disabilities. In Proceedings of the 2017 chi conference on human factors in computing systems (pp. 2878–2889).