Affect Conveying Instant Messaging

Master of Science Thesis
University of Turku
Department of Computing
Interaction Design
2021
Rasmus Kyyhkynen

UNIVERSITY OF TURKU
Department of Computing

Rasmus Kyyhkynen: Affect Conveying Instant Messaging

Master of Science Thesis, 76 p., 2 app. p.
Interaction Design
August 2021

Instant messaging applications cannot convey non-verbal communication through text-based messages. That can lead to an unpleasant misunderstanding between dyads when the discussion is held on a computer or smartphone. This study aims to determine if the affect conveying instant messaging applications has any usage within users who are daily users of instant messaging applications. Furthermore, does the application benefit the test users in the real variant group compared with the control group users? The tests were conducted with an instant messaging prototype application developed just for this experiment.

To test the affect conveying instant messaging prototype, we gathered a test group, which were randomly divided into two different groups, those that tested the correct version and the control group. Both test groups tested the same application but with different affect conveying module or variant. The real group tested the real variant, and the random or control group tested the variant, which randomly chooses the conveyed affect or emotion. The affect is conveyed with emojis in both of the variants. After the tests were done, testers had to answer nine different interview questions. Finally, for three interview questions, testers give a grade on how satisfied they were with that particular function. The grades were analyzed with descriptive statistical methods, and the verbal interview answers were analyzed by gathering recurring themes across the answers.

The study results show that the real variant of the affect conveying instant messaging prototype performed better overall than the random variant. Test users also think that the prototype and its affect conveying functionality was fun. However, they did not see any exact situations where they would use the affect conveying functionality in an instant messaging application. Testers thought that they would use it with friends and family rather than in professional life. Generally, the way emotions were conveyed in the prototype was well-received. Test users did not see any significant issues in it or if the same functionality would be used in applications like games.

Keywords: affective, computing, instant, messaging

TURUN YLIOPISTO
Tietotekniikan laitos

Rasmus Kyyhkynen: Affect Conveying Instant Messaging

Pro gradu -tutkielma, 76 s., 2 liites.
Vuorovaikutusmuotoilu
Elokuu 2021

Pikaviestintäsovellukset eivät oletusarvoisesti välitä sanatonta viestintää tekstimuodossa olevien viestien mukana. Se on omiaan aiheuttamaan epämiellyttäviä väärinymmärryksiä keskusteluosapuolien välillä, kun keskustelu käydään tietokoneiden tai älypuhelimien avulla. Tämän tutkielman tavoitteena on selvittää, onko tunteita automaattisesti välittävällä pikaviestintäsovelluksella käyttöä niiden käyttäjien keskuudessa, jotka käyttävät pikaviestintäsovelluksia päivittäin omassa elämässä. Lisäksi selvitämme, tuoko testisovellus hyötyä sovelluksen oikeaa versiota testanneille verrattuna kontrolliryhmään. Testit ovat suoritettu sille erikseen luodulla pikaviestintäsovelluksella.

Testatakseen testiin erikseen tehtyä pikaviestintäsovellus prototyyppiä, kokosimme testiryhmän, joka satunnaisesti jaettiin kahteen eri ryhmään, oikeaa versiota testanneisiin ja kontrolliryhmään. Molemmat testiryhmät testasivat samaa sovellusta, mutta joissa olivat kuitenkin eri affektiivinen moduuli. Todellisen testin suorittaneet testasivat niin sanottua oikeaa versiota, jossa tunteiden välitys toimi kuten sen oli tarkoitus toimia ja kontrolliryhmä testasi sovelluksen versiota, jossa tunteet, joita välitettiin, olivat sovelluksen satunnaisesti valitsemia tunteita, eivätkä oikeasti sovelluksen analysoimia tunteita käyttäjästä. Sovellusprototyypin testin jälkeen testaajat vastasivat yhdeksään haastattelukysymykseen, joista kolme kysymystä olivat sellaisia, joihin testaajan tuli antaa numeerinen arvosana siten kuinka tyytyväisiä he olivat kyseiseen toiminnallisuuteen. Numeeriset arvosanat analysoitiin kuvailevia tilastointimenetelmiä käyttäen ja sanalliset haastatteluvastaukset analysointiin teemoittamalla haastattelut ja löytämällä sieltä toistuvia teemoja.

Tutkimustulokset osoittavat, että pikaviestintäsovellus prototyypin niin sanottu oikea versio toimi yleisesti paremmin mitä satunnaistettu tunteiden välitys. Testaajat myös ajattelivat, että prototyyppi ja sen automaattinen tunteidenvälitys oli hauskaa ja mielekästä seurattavaa keskustelun aikana. Testaajat eivät kuitenkaan haastattelun aikana löytäneet mitään tarkkaa reaalielämän käyttöä sovellukselle. Testaajat käyttäisivät sovellusta mieluimmin läheisten ja jo tuntemien ihmisten kanssa kuin esimerkiksi työelämässä. Testaajat yleisesti pitivät pikaviestintäsovelluksen tunteiden välityksestä, ja he eivät nähneet mitään suurta ongelmaa sen toiminnan kanssa tai että samanlainen toiminto olisi vapaaehtoisesti käytössä joissain muissa sovelluksissa, kuten videopeleissä.

Asiasanat: affektiivinen, tietojenkäsittely, pikaviestintä

Contents

List of Figures

List of Tables

1 Introduction

Instant message applications have evolved and improved during the years when they were first introduced to desktop computers. For example, many instant messaging applications today have video and audio calls to communicate with each other. However, text-based messaging has not changed dramatically. Devices have more computing power and resources to calculate more complex calculations, giving the applications the ability to have advanced functions that could help humans, such as how they communicate between internet-connected devices. Also, the more complex calculation can be done in the cloud systems to ease the load from the end-user devices with the help of high-speed internet connections. Traditional text-based conversation held between individuals in a dyad using, for example, a computer or smartphone does not convey human non-verbal communication. That is one of the reasons why participants in the conversation might not understand the meaning of one's saying, or it may also induce unpleasant misunderstandings [1]. The communication between two people in instant messaging applications has gone in the direction of where people in the conversation uses, for example, emojis to communicate emotions, feelings, or mood to supplement the initial text-based message. However, the emoji user inserts to the message might not be the affective state the user is experiencing but something that they want to signal.

Psychologists around the world have studied how we humans form emotions and especially how we communicate emotions with our body [2, pp. 133–138]. For exam-

ple, American psychologist Paul Ekman concluded after years of studying human emotion expression around the world that human facial expressions when experiencing specific emotions are the same no matter what culture the person is coming from. That research has led to a system called Facial Action Coding System (FACS), which determines the movement points of the face when a human is, for example, experiencing happiness or sadness [2, p. 263]. Moreover, FACS is used in facial expression analyzing systems to predict emotions. Studies around emotions have also brought interest in the technical field and its contributors to developing computer applications that can analyze and predict emotion from the picture of the human face. One of the best-known applications is Affectiva's software development kit (SDK) that can recognize emotions from human facial expressions in real-time. Nowadays, facial expression recognition systems use machine learning techniques to predict emotions.

In this thesis, we have developed an instant messaging application that can automatically convey emotions with emojis along with send text messages. With the prototype, we aim to reduce the possibility of misunderstanding text-based messages and make the discussion in the application feel more like face-to-face dialogue. Furthermore, minimize the possibility of misunderstanding between individuals in a dyad. The prototype application reads the user's facial expressions with a computer's web camera, and the affective module analyses the facial expressions during the conversation held in the application. It then conveys the emotion with the corresponding emoji. Some research is done from this topic but not specifically with emojis and that the affect conveying is more continuous and not reaction-based conveying [3], [4]. In reaction-based conveying, the emotion is conveyed after the affective module detects the emotion, and it is predicted strong enough to be sure that the predicted emotion is the real emotion the user is experiencing. We combine the traditional instant messaging application, which cannot convey non-verbal com-

munication in text-based messaging, and the facial expression recognition module. Test users are separated into two different groups to test the prototype. The prototype application has two versions of facial expression recognition to test whether the real variant performs better than the control variant called random variant.

The main contribution of this thesis is to find what kind of feelings and thoughts the daily instant messaging application user gets when using the prototype. Furthermore, does the application's real variant benefit the real variant test users compared with the control group? For example, are the messages easier to understand if the affect is conveyed automatically with emojis? Finally, do they consider it useful in daily use, and are there any potential use cases of this application and its affect conveying functionality in their own life.

This thesis is constructed from nine different chapters, first of them being Chapter 1, which introduces the subject and the study. In Chapter 2, we go through the basics of emotion theory. Chapter 3 consists of the idea behind affective computing. Chapter 4 explains what instant messaging applications are and how they are used today. Chapter 5 explains the research methods of this thesis and the explanations of why they were chosen. Chapter 6 goes into detail about how the affect conveying instant messaging prototype is made and what technologies are used in the application. Chapter 7 presents the results of the tests conducted in this research, followed by Chapter 8, which includes the analyses of the results gathered in this thesis. Lastly, Chapter 9 concludes this thesis by presenting the conclusion of the results and the main findings of this work.

2 Emotions in Affective Computing

Affective computing is an interdisciplinary field that includes, for example, computer science, psychology, and engineering. The idea of making computational devices more human-like by implementing emotion measurements and analyzing methods or simulating emotions in them insists that emotions and affects need to be understood before implementing them in systems. Therefore, in this chapter, we are going through what emotions are, the basic theories of emotions, and how humans convey emotions. Let us start by looking at what emotions are.

2.1 Basics of Emotions

Emotions and affective states play a crucial part in affective computing. One part of affective computing is that it tries to make Human-Computer Interaction (HCI) more human-like by implementing emotions or human-like emotions into systems or computers [5]. On the other side, it tries to develop systems with emotion detection, for example, for educational or human communication enhancing purposes. Emotions in human communication provide an understanding of one another, and without emotions interacting would feel artificial. For example, if someone tells a joyful story, and the storyteller does not express any emotions, it could feel strange and unnatural. Emotions are also a big part of human decision-making and learning [6, p. 41]. Deciding the outcome from a large number of decisions would be hard without emotions reducing the number of possible outcomes [7]. This reduction of

decisions depends on what the decision-maker is feeling and the needs and goals of an individual. Scientists have not been able to come with a scientific definition of emotion [7]. The widely agreed definition of emotion is that emotion is a reaction to events relevant to the needs and goals of an individual.

2.1.1 Affect

Affect or affective states are considered a wider term for emotions in general [5, pp. 24–25], [2, pp. 89–117]. Affect is an umbrella term for emotions, moods, and feelings. Emotions are, for example, basic emotions like anger, fear, disgust, happiness, sadness, and surprise. These types of emotions are strong and do not last very long. Exact emotions are labels for the subjective feeling. On the other hand, moods and feelings are not exact emotions, but they are more like emotions-like states and feelings. A mood is a longer-lasting state than the emotion [2, pp. 89–117]. A human can, for example, be in a good mood, which can last from a short period to days, and it is not so intense or strong as the emotions like fear. A feeling is an underlying subjective experience of emotion. Moods often influence human decision-making. For example, when a human is in a bad mood, humans tend to share things more evenly, but they are more selfish when a human is happy. Sad people more often do not make decisions that are influenced by biases, but happy people do. A happy mood can cause irrelevant judgment of other's age and gender when judging, for example, how well someone has written the article.

2.1.2 Basic Emotions

Psychologists William James came with the idea of basic emotions [6, pp. 65–70]. James argued that all humans have at least seven different basic emotions: anger, fear, joy, grief, love, hate, and pride. However, he did not pursue this idea, and it remained as a sketch to the point where psychologists William McDougall took

the idea for wider consideration. McDougall created the theory, which he called discrete basic emotions. McDougall argues that humans have emotion modules that are "hardwired." All modules have a detector that takes the emotion pulse when intended for it. After receiving the emotion pulse, it activates the emotion pattern. Most know basic emotions theory is from Paul Ekman, which states that these basic emotions trigger the same facial expression for all humans regardless of culture [2, pp. 133–138]. This approach is critized by other scientists [8], [2, pp. 123–138]. For example, Russel argues that facial expressions are often voluntary and controlled by social rules. There is evidence that facial expressions are culturally different when expressing emotional states, shown in Rachel Jack's research in 2012. Russel would like to change the Basic emotion theory to the "minimal universality of an onlooker's interpretation of facial expressions" because humans tend to perceive others' affective states universally, for example, with dimensions, valence, and arousal.

Other researchers, for example, Sauter and Eisner disagree with Jack's research findings and conclusions, and there is still debate between researchers about whether basic emotions are universal between humans [9, pp. 166–168], [10].

2.2 Theories of Emotion

In this section, we have two main trends of emotion theories, behaviorist and cognitive. We will introduce James-Lange Theory as an example of behaviorist emotion theory and appraisal theory as an example of cognitive emotion theory.

2.2.1 James-Lange Theory

Psychologists have tried to develop emotion theories that explain how emotions form. One of those theories is the James-Lange theory, which was developed independently between two psychologists William James and Carl Lange. They separately came

with an emotion theory that stated and described humans' process when emotions are felt [6, pp. 52–71]. For example, some event is recognized with eyes, which in this example is a fearful event. Pulse travels from the eyes to the sensory cortex, which confirms the fearful object or event's sensory. From there, the pulse starts to activate fear reaction in the body. This fear reaction activates physiological bodily changes, skin, heart, muscles, and then pulse travels back to the cortex. When the pulse is back in the cortex, then humans feel fear. According to James, these bodily reactions activate emotion. A classic example of this is when a human is in the forest, and a bear appears in the vision of a human. Human reacts with physiological arousal; for example, heart rate increases, indicating that humans fear the bear. This emotion theory lasted until the 1960s when cognitive emotion theory replaced the behaviorist emotion theory like James' theory. Other behavioral emotion theories are like the Cannon-Bard theory or Schachter-Singer theory, which we do not get into here.

2.2.2 Appraisal Theory

The two most known cognitive emotion theorists Magna Arnold and Richard Lazarus, came with more cognitive emotion theory [6, pp. 57–60]. For example, Lazarus's theory of emotion is called the appraisal theory. Appraisal theory states that human has to evaluate the object or event before the emotion is activated according to this evaluation. Lazarus argued that humans have two appraisals: primary and secondary. Primary appraisal is a quick and often unconscious evaluation of an event. The secondary appraisal is how one can cope with the stressful event by reassessing the situation, which lowers the initial event's negative effect. Emotion models made of the appraisal theory then are considered to be the most promising way of generating affect in computing devices [2, pp. 383–436]. One of the best known model is called the OCC model developed in 1988, which name comes from the names of theorists Ortony, Clore, and Collins [6, pp. 164–165]. OCC model and some variation

of it is used today by researchers to implement and generate human-like emotion in machines.

2.3 Emotions in Human Communication

Humans typically convey emotions in a face-to-face conversation with facial expressions, body posture, and the pitch of the human voice [1]. Traditional online chats, in which the only communication channel is written text, cannot convey these non-verbal communication elements to the other chat participants because they cannot see or hear each other [3]. Typically, this can lead to a misunderstanding of the message. Humans can identify the affects from the written text. However, while this is possible, there is still something that humans are naturally missing when communicating with the text-based communication channel. This something is often non-verbal communication. The problem we are solving in this thesis is how we can convey non-verbal communication and, more specifically, emotions and affects to other participants of the online chat when users can not see or hear others.

The next chapter will introduce how emotions can be measured and analyzed with a computer and the most used and reliable ways of measuring emotions. We will also see what affective computing is and the applications where affective computing can be useful.

3 Affective Computing

Coupling emotions and computers have been considered a subject from scientific movies or television series. That started to fade away when Rosalind Picard released book called *Affective Computing* [5] in 1997 which introduced findings from the research in affective computing and introduced methods to measure and simulate emotions in computer systems. After that, affective computing has been a growing research area alongside artificial intelligence and machine learning in the 2000s [6, pp. 52–53]. What also raised the interest in developing machines with emotions or emotion-sensing capabilities was when emotion research got to the point where emotions were considered beneficial. Before that emotions have regarded as an obstacle to rational thinking and decision-making [6, pp. 52–53].

In this chapter, we are going through the idea of affective computing, for example, measuring and simulating emotions in computational devices and the most promising applications of affective computing today and in the near future.

3.1 What is Affective Computing?

The term "affective computing" was coined by Rosalind Picard in 1997 [5]. The book started a new research and development area in the field of computer science [5]. The term and idea of affective computing are originated from the technical report paper by Picard in 1995 [6, p. 39], [11]. Affective computing is a multidisciplinary field mixing computer science, engineering, psychology, neuroscience, and many other [6,

p. 23]. In the book's review, Sloman [12] writes that the *Affective Computing* had no competitors in the field at the time book was published, and it was one of the only book relating to how computers could use emotions to create more natural human-computer interaction. Today, affective computing is a relatively mature research field with journal *IEEE Transactions on Affective Computing* [6, p. 23].

In affective computing, devices and systems are developed to recognize, interpret, and simulate human affects [5, p. 3]. In the words of Picard, affective computing is "computing that relates to, arises from, or deliberately influences emotion". Affect is used as a synonym of emotion in affective computing [5, pp. 24–25]. The affective computing objective is to develop systems that can simulate human empathy and make computers emotionally intelligent or act emotionally intelligent [5, pp. 75–79]. Simulating human empathy means that computers try to read a user's emotional state and adapt behavior accordingly. For example, an affective virtual tutor in an e-learning platform could recognize student frustration and give hints to help the student continue with the assignment. With these hints, emotionally intelligent software can lower the frustration students feel, making learning more effective. Picard thinks affective computing is a key element to make computers or systems intelligent [6, p. 53]. Her thought is originated from Antonio Damasio's research in neuroscience, which has shown that emotions are a key part of human decision-making.

Humans tend to express emotions at computers on their own [5, p. 248]. People express their frustration at computers, but computers have no ability to do anything about it. If computers could see when the user is frustrated, designers could use it to design better software. This could be helpful, for example, in usability testing. People who are managing the test could forget or do not recognize the user's emotion. However, the affective computer could help the testers and save all emotions that are recognized by the machine.

3.2 Measuring Emotions

Affective systems or devices sense the emotional state of the user. This recognition of a user's emotional state can be done via camera, microphone, or low-level sensors.

3.2.1 Camera

The camera, coupled with appropriate emotion measuring software, can read users' facial expressions or body posture and interpret the emotional state the user is experiencing. Psychologist Paul Ekman studied human non-verbal communication and found that facial expressions do not vary between people from different cultures [5, pp. 25–26]. According to Ekman, humans have the same facial muscular movement and characteristics when they are expressing basic emotional states like joy, sadness, anger, or disgust. However, the latest meta-analyses have shown that facial expressions vary between cultures, and facial movements do not always imply exact emotions [2, pp. 135–137]. James A. Russell proposes shifting the Basic Emotion Theory to a more universal theory [2, pp. 123–138]. This is because humans themselves tend to perceive others' emotional states more universally, like recognizing valence and arousal levels. Studies have also shown that humans might not express the emotion that they are experiencing. For example, one does not always express happiness when feeling happy. American company Affectiva has an application that senses in real-time basic human emotions from the face with the device's internal camera. The application uses known characteristics of human faces in an emotional state and represents the level of emotion on screen. Recognizing affects from facial movement is done by sensing how facial muscles are moving the face [2, p. 263]. According to the Facial Action Coding System (FACS), humans have 46 possible facial action units. These action units are then associated with emotions. For example, happiness consists of two different actions, units 6 and 12. Action unit 6 is a cheek raiser, and 12 is a lip corner puller.

Recognizing affective states or emotions from face movement has proven to indicate the emotional state the human is feeling. What has not been much studied or taken into account is, can we use eye-tracking for recognizing the emotional states of humans [13], [14], [7]. For example, pupil dilation size has proven to change when expressing positive or negative emotions. That could be used as a complementary method for face recognition. With eye-tracking, the user's eyes are measured with a monitoring system that follows the point of gaze or where the user's eye is focused. These types of cameras can also measure pupil diameter. The eye is typically measured by reflecting infrared from the eye and sensing the light with a camera lens capable of doing that. The possible eye-tracking device is an external device with an eye-tracking camera and projector, like Tobii Eye Tracker 5 mounted under a PC monitor. Furthermore, some of the Virtual Reality glasses have eye-tracking technology, like in HTC Vive.

3.2.2 Microphone

In affective systems, a microphone can be used to analyze the user's emotional state from speech [5, pp. 178–184]. Recognizing emotional states from users' speech can be done by comparing real human speech with the database's contents [7]. The database often has laboratory-made examples of different emotional speech recorded from humans. The comparison can be made with suitable algorithms, for example, the k-nearest neighbor algorithm. The algorithm compares pitch and intonation values of speech and selects the nearest option.

3.2.3 Sensors

In addition to cameras and microphones, basic emotions can be recognized from low-level sensor data, but not directly [5, pp. 141–164]. Humans react with different physiological signs in different emotional states. These physiological signs come from

the human autonomic nervous system, which regulates bodily signs. Physiological signs that are usually measured are muscle tension, heart rate, skin conductivity, skin temperature, respiration, and brain activity. Measurement of these physiological signs can be done with several different wearable devices with sensors. Low-level data from sensors do not indicate any emotional state. However, when different data patterns are analyzed and interpreted, signs of joy, anger, fear, or other emotional states can be found [2, p. 256]. Analysing of data can be accomplished with different machine learning and data mining algorithms.

3.2.4 Multimodal Sensing

However, recognizing and sensing emotions from one affect sensing device is often not enough [2, pp. 256–258]. That is why it is preferred to use multiple data sources to recognize affects from a human. Affect sensing from multiple sources is called multimodal affect recognition. Multimodality requires adding a third step, which is called data integration. Data integration is either done by integrating the raw data from sensing devices and then making the pattern recognition from the integrated data, or every sensing device's data is analyzed separately before the analyzed data is integrated.

Affective computing does not always mean recognizing exact emotions [7], [15]. Exact emotions are hard to recognize from humans because humans tend to feel emotions in different ways. People's faces may be the same when experiencing one of the basic emotions, but recognizing more cognitive types of emotion like grief is much harder. It is easier to recognize, for example, grief if you know the person already, but computers do not read and learn users' emotions by default [5, pp. 53–54]. It could be easier to measure users' arousal and valence levels than discrete classes of emotions in some situations [7]. For example, using a two-dimension model Circumplex model to predict users' affect is a way of interpreting users' emotions. Each emotion is

placed in a two-dimensional coordinate system. The model uses arousal and valence as dimensions. Arousal is the level of how calm or exited the user is. Valence measures how negative or positive the situation is for users. Dimension models of emotions create a more flexible representation of emotions. One example of using the dimensional model in games is where Makantasis et al. [15] analyzed gameplay videos from different players and try to predict players' arousal level from these videos. Furthermore, a third dimension: dominance, can be added to this model. This helps to differentiate emotions with the same arousal and valence levels, e.g., fear and anger. Dominance illustrates how dominant or submissive the user feels, and that helps. The emotional dimension model with three dimensions is called the PAD model, PAD coming from Pleasure (valence), Arousal, and Dominance [2, p. 258].

3.3 Simulating Emotions

Affective computing consists of more than recognizing user's emotions. An affective computer or system could also express emotions or ultimately have emotions. Expressing emotions from the computer side could be a speech with emotions coloring it or showing a face that indicates what emotional state the computer is experiencing. Computers do not have human-like emotions, but they may simulate emotions. Expressing emotions can be thought of as empathy towards the user and intelligent emotional response to a user. [2, pp. 15–17] Simulating or generating human-like emotions in systems is often done by emotion model theory. The most used emotion model theory in affective computing is the OCC model (See Section 2.2.2).

3.4 Applications

Where can affective computing be used? And what are the applications of affective computing? In this section, there are examples of using affective computing in a couple of potential areas.

3.4.1 Healthcare

One significant usage of affective computing is in healthcare [2, pp. 383–436]. The healthcare field is interested in using technology to heal or train patients with, for example, depression and autism. In psychology treating autism and teaching, for example, autistic children to be more capable of expressing emotions and read other's emotions. Affective robots and avatars are considered to be helpful in training. Especially autistic children tend to be very interested in robotics and other mechanical devices, which helps keep the children more interested in learning. Some studies have found e-therapy or cyberpsychology to be a good alternative or complementary way of doing therapy sessions [6, pp. 817–833]. One side of cyberpsychology tries to find ways of using technology to help people improve their well-being. These applications could be used in remote places where conventional face-to-face therapy sessions can not be arranged. E-therapy does not replace traditional face-to-face therapy sessions but can be an excellent way to train emotional skills at home. Home training sessions can be done with gamified online therapy applications.

3.4.2 Education

E-learning is also considered the main application where affective computing can be used [6, pp. 629–649], [16]. There could be an e-learning platform where the platform recognizes that the student is frustrated and can not continue with the assignment. An affective e-learning platform could then automatically give a hint

to student on how to proceed, or the platform could message the teacher that one of the students has problems in this part of the assignment and he or she might need help. E-learning platforms have two possible ways to implement affective computing. There are proactive and reactive learning tutors. The proactive way tries to induce the learner's right affective state, engage the learner, and make the most out of the situation. In the reactive model, the system tries to respond to the learner's affective states correctly. For example, if the student is experiencing frustration when doing tasks, the system tries to help the student by giving hints or saying that they need a break.

3.4.3 Robotics and Social Companion

We are in a world where the global population is aging fast [17]. This creates a problem, how we can take care of the aging population when there are too many older persons to be taken care of compared to the younger population. There are more to take care of than people that take care. One increasing need could be the need for social contacts. One of the solutions for making social contacts could be social robots [2, pp. 359–375]. Human-robot interaction (HRI) is its own branch of interaction design. However, it is also closely related to Human-Computer interaction (HCI). In HRI, interaction designers try to understand how the robot should behave and what type of interaction users expect from it. Designing Human-Robot interaction may be more challenging than designing, for example, desktop applications because they are embodied in the environment and interact with the physical world. For good interaction to take place, affective capabilities could be the solution. Making social robots understand human affective states, express emotions, and use affective knowledge to be a more natural companion to older people.

Robots are also considered to do some of the work that humans have been doing to this date, like working on the hotel reception, answering questions at the help

desk, or do the unpleasant conveyer belt work [2, pp. 359–375]. Robots in customer service could be performing better with affective knowledge or affective intelligence. Like with the social companion robots, robots in customer service might notice that the customer is frustrated. Robots could be programmed to behave just like humans might behave in that situation and empathize with the customer.

3.4.4 Retail

Maximizing sold items, making as much income as possible, and generating profits for its owners is the company's ultimate goal. That is why companies are often in search of technology or other methods to help with it. Some have suggested that retail shops could mount cameras or sensors to analyze the emotions of the customers [18]. The easiest way to do this is to mount cameras to places where customers face when interacting with store objects. The cameras can send the image of the customer's face to the computer, which then analyses the customer's emotions from that image. Emotion data can tell the storekeeper, for example, that the customer in that exact position was frustrated or not having a good time. That could indicate to the staff that this shelf or rack should be changed somehow.

3.4.5 Games

Playing a game is an emotional event [19], [7]. Games induce emotions in different game-related situations, such as stressful events or events that trigger empathetic feelings towards the game character. Interactive stories or stories, in general, are the right way of inducing different emotions for the player to experience. What games are lacking is recognizing and interpreting these emotions. For example, we can recognize a player's happiness and mimic or mirror the player's expression to the non-player-character (NPC). Mimicking gives the player the feeling that the NPC understands what the player feels, and the NPC is empathetically capable [6,

pp. 453–470]. This can lead to a more immersive and personal experience. The same can be done in situations where a virtual assistant is necessary. For example, we could have virtual reception in the hotel where there are computers to replace the reception workers behind counters. These virtual customer servants could have cameras in them, and it continuously reads the customer's facial expressions and mimics the expressions of the customer on the face of the virtual avatar.

3.4.6 Entertainment

The applications today try to be highly personalized. Mobile application especially asks personal preferences, such as what type of music the user listens to, age, gender, or how much the user exercises in the day. Personalization of mobile applications could be more popular because mobile devices are highly personal devices, and there is usually just one person who is using it. The questions asked are affected by what kind of application is involved. Sports applications ask how the user exercises and how old are they, and music applications ask what genre is their favorite. Affective computing is one way to make an application feel more personalized, making the interaction with the application more enjoyable [20]. An affective application might notice if the user who is watching videos is in fear and asks the user if they would like to change the video to a more positive one. Same works with music applications [21]. The application recognizes the user's affective state and asks if the user wants to change the song or music category if they are in a negative state. In the future, this kind of application could then remember that the user does not like to watch videos of this kind and warn the user that the next video is something that the user did not like earlier. The warning could come before the user is pointed to the next video, or the video does not play after the user confirms that they understand the warning.

In the next chapter, we will go through the instant messaging applications more

closely. The next chapter contains basic information about what instant messaging applications are, how they have changed compared to 15 years ago, and what emojis are, and how they are used in these applications. Moreover, what kind of affect conveying instant messaging applications are there today.

4 Instant Messaging

Instant messaging applications (IM) are common in today's world. Most adults and younger people have smartphones, which contain at least one or multiple different instant messaging applications. Some of the most used instant messaging applications globally are WhatsApp, Facebook Messenger, Telegram, and WeChat [22]. People use instant messaging applications for fun and personal use to communicate with friends and family or for professional use in work between colleagues. Instant messaging applications are online real-time chat applications that work over the internet. Communication in these applications is text-, voice- and video-based between two or multiple participants in group chat. Users can also share documents, pictures, locations, or contacts with each other. WhatsApp, which is sometimes referred to as a mobile instant messaging application (MIM), also has web and desktop versions available, but a primary platform is mobile. There are also other MIMs, such as WeChat, which is used in China, and Telegram, an open-source variant for MIMs. According to Statista [23] WhatsApp was the most popular global mobile messenger application worldwide in 2019, with 1.6 billion monthly active users. It is also the third most popular social media application. The userbase of these applications is growing every year, and growth does not appear to end in the coming years [22]. That leads to a great opportunity for companies and designers to freely ideate new and future technological functionality the applications could need.

In this thesis, all these applications, like what MSN Messenger was or WhatsApp

and WeChat are today, are called instant messaging applications for clarity and the similarity of the applications. Also, by the fact that these applications have desktop and web versions available so there is no need to categorize them into a specific platform.

In this chapter, we are going through a brief comparison of modern instant messaging applications, emojis used in these applications, and what we could see in instant messaging applications in the future.

4.1 Comparison

Before smartphones, users' desktop computers had instant messaging software like MSN Messenger, AOL Instant Messenger, or Skype, which worked and functions as the new instant message software today from the user's point of view. As shown in Figures 4.1, 4.2, 4.3, the MSN Messenger, or later called Windows Live Messenger (WLM), had almost the same emojis, functionality, and user interface layout as today's instant messaging applications have like WhatsApp and Skype. Although the appearance of emojis and functionality then were similar to today, the functionality especially was not that advanced. For example, new applications use end-to-end encryption, which was not used in older instant messaging applications. Windows Live team announced in 2009 that they had about 330 million active users, and the application was the most popular instant messaging application that day [24]. Microsoft later bought Skype in 2011 and ended the Windows Live Messenger development shortly after. Microsoft's Skype is still in use by consumers and companies [22].

Before WhatsApp and smartphones, text messages were sent between users using paid services like short message service (SMS) and multimedia messaging service (MMS) [22]. When WhatsApp was released in 2009, WhatsApp told users that they do not need to pay for using it and can use it just with an internet connection.

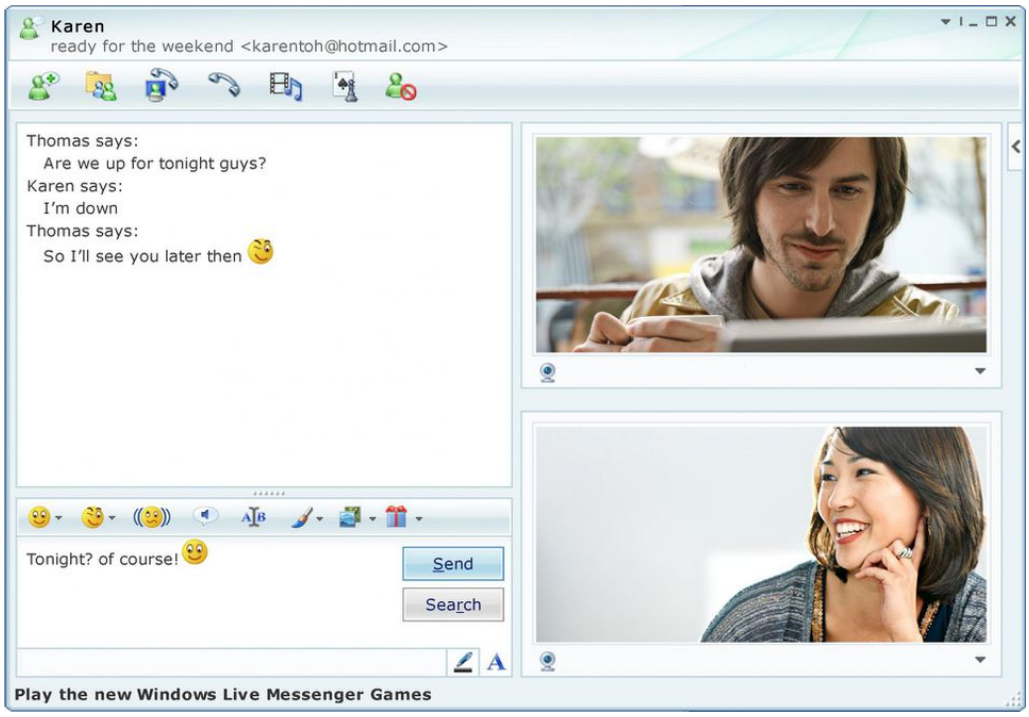

Figure 4.1: Windows Live Messenger. [25]

It drew the users away from the paid SMS and MMS text message services. The transition from the old type of phones to smartphones enabled the new type of instant messaging applications to be popular.

Today, many instant messaging applications work with phone numbers, making application usage on mobile devices more convenient. WhatsApp, for example, recognizes the users that have made the WhatsApp account from the user's contact list by mobile phone number, and it offers the available contacts through the application. However, that has not been the case with all the modern instant messaging applications. For example, to use Skype, users still have to make an account in order to use the application and add users to the contact list with e-mail or Skype name to start messaging with the other users. That is probably due to the fact that its primary usage is on desktop computers. Similarly, Facebook Messenger uses the Facebook account to connect with friends inside the Facebook service.

Instant messaging applications have not changed much after the early 2000s ver-

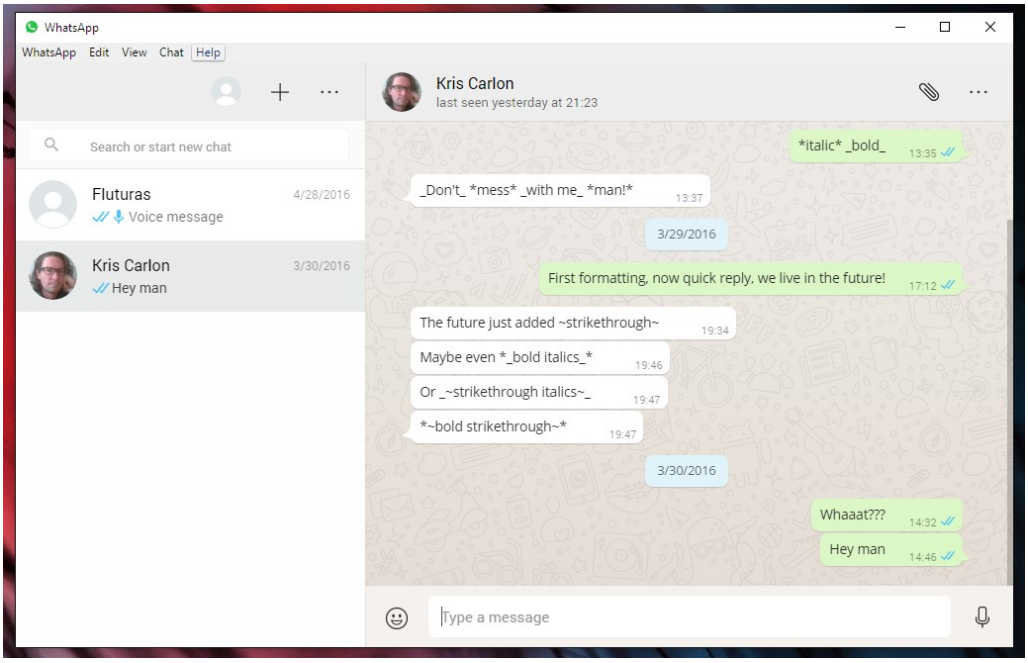

Figure 4.2: WhatsApp Desktop. [26]

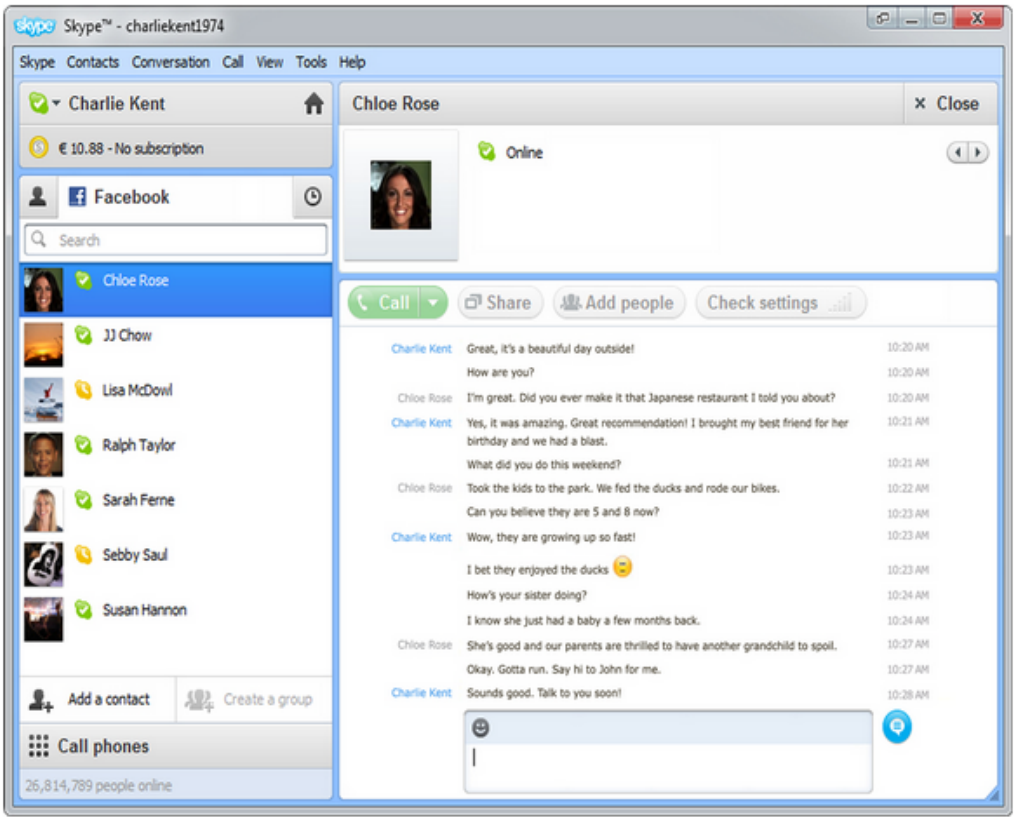

Figure 4.3: Skype. [27]

sus now in 2021. However, the developers have modernized the platform's look, feel, ease of use, and convenience as time passes. The user interface (UI) layout has stayed the same: texts appear vertically below each other, and the text input field has stayed in the same position below the field where text messages appear. The Windows Live Messenger already had a web camera and voice message functionalities in the updated versions, which are almost default functions in today's instant messaging applications. The instant message applications have still kept their popularity to this date without introducing any revolutionary functionality that could have changed how we communicate with each other. Video calls and voice messages could be getting more popular, reducing text-based messaging as time goes by. So if we want to introduce new ways of traditional text messaging, for example, conveying an affect automatically with the text message, it may be too late. Text messaging is still faster and more convenient for sending short messages than video calls where both users need to be in a place where privacy is not a concern, and both can speak at that time.

4.2 Emojis

Emojis are smileys or ideograms that users use on digital platforms [28]. They are used to convey feelings or increase playfulness and enliven the conversation [29]. One can place a smiling emoji at the end of the sentence or where ever user wanted it to convey happiness or a sparkling wine bottle to enliven the birthday wish. Emojis enable the creativity of text messaging, and users are not restricted to use only text. Emojis can also replace some of the text. For example, the user can replace the "hahaha" message with one emoji called "face with tears of joy" when replying to a funny message from a friend. Before emojis and even today, some users use emoticons, text-based versions of emojis, for example, ":-)" or ":-(". Emojis have slowly replaced emoticons.

Japanese designer Shigetaka Kurita is believed to be the developer of emojis [30], [31]. He created the emojis in the late 1990s for the news application that the mobile phone operator company Docomo developed at the time. Kurita himself also worked for the company. He believes that long letters sent in Japan can cause misunderstanding on the receiving end because non-verbal communication, like facial expressions, could not be transferred through written letters. That is why the emojis were initially created, helping to express human affects. Emoji's first appearance on a major platform was in 2011 on Apples' iOS version 5 and have stayed in since.

Emojis were first added to Unicode version 6.0, released in 2010 [28]. Unicode, maintained by Unicode Consortium, is a universal character encoding standard for information technology systems. Unicode includes all the characters from the world's writing systems, and most of the world's digital devices use and support the Unicode standard for characters, symbols, and emojis. Each of the characters in Unicode has its code for information systems to recognize them. Before Unicode standardized emojis, digital platforms and device manufacturers like Apple had their own emojis and emoji designs. Today, however, when Unicode standardized the emojis, the general appearance is the same, but the design and meaning can vary between devices, languages, and cultures. If we think of emojis, for example, in MSN Messenger and WhatsApp today, the graphics and designs have changed from the early 2000s, but the general appearance has stayed the same. There are still differences in the people who use emojis and how they interpret the meaning of emoji [32]. Part of the interpretation differences comes from the fact that emojis are different between, for example, mobile platforms Android and iOS. The "face with tears of joy" emoji was selected as the "word" of 2015 by Oxford University Press, which describes their popularity today [33]. Unicode's latest released emoji version 13.1 has a total of 3,521 emojis, according to Emojipedia [34].

4.3 The Future of Instant Messaging

As mentioned in this chapter, instant messaging applications have not changed dramatically during the past 15 to 20 years from the user perspective. What new could we implement in the instant messaging application today? One of the instant messaging applications' problems is that they cannot convey non-verbal communication through text-based messages [3]. That has been recognized since computer-mediated communication (CMC) was introduced. There are many attempts to implement different ways to show what emotions the user at the other end of the network express. Studies have implemented emotion conveying to the application with different ways to convey them, such as changing color, kinetic typography, or animated avatars [3].

Several studies in the past focus on the claim that instant messaging applications used in workplaces lower the performance of the employees [35]. That has begun to turn around lately with new studies being published. Today, they are considered effective and have many benefits around professional use compared to an asynchronous e-mail, at least in the cases of multi-tasking, time savings, and better social relationships [3].

Instant messaging applications considered harmful in workplaces might be one reason why they have not received any new groundbreaking technological features for communication between users. The most likely reason is that they already have video calls, and users can also send recorded voice messages quickly. Furthermore, users can put emojis in the text by themselves to convey affects. However, they are not the solution to the affect conveying problem of text-based messaging where affect is conveyed almost automatically in the way it is in a face-to-face conversation. Users still use text-based instant messaging for asking quick questions, communicate with multiple users in a group chat or in situations where conversation cannot be held privately [35]. In other types of applications like the team messaging app Slack, which is primarily used in the workplace between colleagues, the primary

communication method is written text in the channels created inside the application [4]. For example, the company that creates software could have separate channels for programming, design, marketing, and management, where the feature of conveying non-verbal communication could be helpful.

4.4 Affect Conveying Instant Messaging

There have been instant messaging application prototypes with affective capabilities in it, at least from the year 2000 when "affective computing" term was coined and brought to the audience [3], [36], [37], [38]. The IM applications that used to be used, such as MSN Messenger, were not made for affective interaction. However, instant messaging applications had emoticons or emojis in them, which are one way of conveying affective states to another participant in the conversation. There are newer affective instant messaging applications and different variations on how the emotion conveying should be done, like mentioned in Section 4.3. There is also one prototype application that uses emojis as an affect conveying method to the chat partner. Liu et al. [4] developed a ReactionBot application on top of the team messaging application Slack. ReactionBot conveys emotion alongside the message by attaching reaction using emoji to the message according to the emotion recognized from the webcam image. It can also attach reaction using emoji to received messages. Reaction emojis in Slack appear below the message.

Poguntke et al. [3] have made a review of what type of instant messaging application with affect sensing capabilities is there and how they try to convey the emotion. They also formed four prototypes of the affective instant messaging application built on the WhatsApp Web application. Four prototypes included a color change of the chat bubble and chat window itself, font change, and profile picture change. Messages sent between users in the instant messaging application could be in different colored boxes, indicating the user's feeling, or the appearance or font

of the text sent by the user changes according to the user's feelings. Alternatively, the users in conversation could have an animated avatar or profile picture in the applications' window, which changes automatically whenever the user's feelings are changed or even filtering the emojis depending on what affect is analyzed from the user's face. These were selected based on surveys.

We can note from the affect conveying instant messaging literature that test user who has tested these prototype applications tend to have mixed feelings about the affect conveying capabilities [3], [4]. A big part of them says that they got help from the affect conveying by being a nice to have feature, and others say it has no value whatsoever. Positive feedback is, however, dominant compared to negative.

The problem with conveying affective states with emojis to another participant in the conversation is the freedom of giving whatever affective state the user wants regardless of what real affective state the user is experiencing [3]. For example, the user might be in a sad state but give the other participants a signal that they are all right and in a happy state. Another problem is that with text, it is hard to convey the affect, but not impossible. Users can detected affective states from chat partners text in instant messaging applications [39], [40]. The affect is there, but the receiver's possibility of misunderstanding the message is high [1]. While it is found that users can identify affects from the text, users still think that something is lost when writing messages compared to communicating face-to-face [3].

This chapter ends the literature review section of this thesis. We have gone through the basics of emotions, the idea behind affective computing, and what are instant messaging applications. For now, we have a good knowledge base to continue to the following sections. The main contribution of this thesis is the affect conveying instant messaging prototype, the experiment made for the prototype, and the interviews. From this point forward, this thesis explains the research methods and testing done from the application, the basic functioning of the application, and

what we found when testing the application with real users. In the next chapter, we go through the research methods of this thesis.

5 Research Methods

When reviewing the literature regarding affect conveying instant messaging application, the results usually come from user experience, and usability tests [4], [3]. Test users are asked to test the software for a while, varying from a few minutes to several short tests sessions over the course of weeks. After the initial experiments with the application are done, users typically answer a questionnaire and are interviewed. These usability test results form subjective interview data from the test users and numerical satisfactory ratings. We follow the same structure in this thesis.

The research structure is listed here:

1. We first identify the problem.
2. A prototype is developed to solve the problem.
3. Create relevant research questions.
4. Choose the research methods. In this thesis, the methods are a usability test, interview, and questionnaire.
5. Run the tests with test users.
6. Write down the results.
7. Analyze the satisfactory rating data with descriptive statistics and the interview data with coding and theming.
8. Interpret the results.

In the following sections, we go through more closely how we conducted the affect conveying instant messaging applications' experiments in this study and what research methods we used, and why we choose the exact research methods for this thesis. Furthermore, we also go through the research questions, the plan for making the experiments, data gathered during the experiments, and the analyses we did for data.

The idea of the experiments done for this thesis was to give the user an experience of what the affect conveying instant messaging application is and how the affect conveying is implemented in the application. After the initial testing of the application is done, users are interviewed to get opinions on the application, its affect conveying implementation, and where it could be used, or do the interviewees see any usages for this kind of application.

The tested instant messaging application prototype in this thesis has a randomized emotion conveying chooser. When the server which transfers the information send by the testers is launched, it picks real or random emotion conveying variant. The test has been carried out in a double-blind test fashion. Thus, neither the user nor the test facilitator knows which test variant is chosen before the test. In the real variant, emotions are analyzed from the tester's face and conveyed to the other test participant in the application with specific logic. This logic is explained thoroughly in Chapter 6. In the random variant, emotions are conveyed randomly. The application picks the emotion randomly from the list and then conveys it to the conversation partner. Emotions are not analyzed from the testers' faces in the random variant.

5.1 Research Questions

A couple of papers consider affect conveying instant messaging applications where emotions are conveyed with, for example, colors or avatars [3]. Emojis have not

been considered that much. Some prototypes have emoji filtering depending on the kind of affective state the facial expression analysis has analyzed. The filtering then chooses the most suitable set of emojis for the analyzed affect. Moreover, one study has done a software, which works with emojis and the affect conveying is implemented as a reaction to the messages [4]. In that study, the affective state is then shown with emoji below the messages. Also, none of the found papers do not really consider the situation where the emotion conveyed is not the affective state of what the user is experiencing and how the users feel when the affect or emoji showed in the application does not represent the feeling of the user.

In this experiment, we compare the results between users who tested the real variant and the random variant of the applications' affect conveying implementation. The testers who randomly test the random variant are considered the control group in this test. We have made an application that uses emojis as conveying method, but the implementation is more continuous than reaction-based conveying. In this thesis, we have formed two research questions to solve and answer. One of which asks whether the affect conveying with emojis gives benefits to the users compared to a system where the affect conveyed is randomized. The second one asks, do we have any use cases for this kind of application, and do the users feel like this could be something that we need in everyday life. The research questions are introduced below. The first research question is:

> *RQ1: Does automatic affect conveying with emojis benefit the instant message application users compared to a control group where affect conveying is randomized?*

RQ1 asks how well the emotion conveying method works in this thesis's prototype application. It is also strongly related to the underlying test mechanic and testing of this application. We wanted to know whether the real emotion conveying variant gives better results on understanding the messages and whether the emotion

conveying implementation performs overall better than the randomized emotion conveying variant. RQ1 also asks whether emojis are a good way of conveying affect in text-based conversation. Emojis are used by the users in the messaging application for conveying affect, but does it differ from the automated way of conveying them. Alternatively, we want to know whether the test users change their own emotions depending on how the prototype displays their emotions. For RQ1, the interview consists of three questions that provide measurable data for the analysis. The second research question is:

> *RQ2: Does automatic emotion conveying have a place in an instant messaging application when the application is used in the free time?*

For answering the RQ2, our study utilizes an interview and its questions where we find answers for it. With RQ2, we find out the possible use cases the test users might see for this application and is the automatic emotion conveying something that the test users could see using in the existing applications, such as in WhatsApp. The most recent studies focus on the affect conveying method and how well the method works than the use case of this kind of application. One of the interesting things here is that do the answers which consider the usefulness of the application in everyday life differ from the users who tested the system where the affect conveying was randomized and with users who tested the real affect conveying implementation. That and other relations are presented in Chapter 7.

5.2 Experiment and Testers

The tests planned to give the testers an idea of how the affect conveying is implemented in the instant messaging application prototype and what type of experiences and feelings the test created for test users. Testers were recruited by text message and phone calls. The recruit messages and calls consist of a brief explanation of what

type of application it is and that the user had to participate in the interview after the test. All the tests and interviews were conducted within one month, and the test setting or interview questions did not change during the tests. Only the place where the test was made changed between test users according to where they live. However, it was ensured that the test place was a familiar place that does not affect how the test user is expressing affects, for example, no tests were performed in public areas. All test users have already had experience with instant messaging applications, such as WhatsApp, Skype, or Teams. The participants were not videotaped or audio recorded during the test, and discussion texts inside the instant messaging application were not saved. Only the interview answers were documented for the results and analysis. That information was also communicated to participants before the test began.

The test users were the facilitator's friends and family members. Testers who do not know the facilitator were not recruited because it was thought that testers who know the facilitator beforehand could then freely express their affect to the facilitator and test users are more transparent on expressing emotions. No particular restriction was set on the recruitment of participants. However, the test users had to have previous experience with some instant messaging applications and know the facilitator well. All of the ten testers use instant messaging applications almost daily in their free time or at work. The age of the testers ranged between 21 to 60, seven men and three women, and test users mainly come from the age groups between 20 and 39. The average age of testers is 34,9 and the median age of 29. The testers' age distribution is shown at Table 5.1. Nine of the test users are in working life, and one is a student. Because the test application has two emotions conveying variants, real and randomized, the variant was randomly picked from a list of five real and five randomized variants. The tests were conducted in a double-blind test fashion. Both the test user and the facilitator did not know which one of the tests

was chosen. The facilitator, who is also the researcher and the developer of the prototype, could guess during the test which variant was most likely chosen by the randomization. Whenever one test has conducted, the variant which was picked from the list was removed. The ten testers are then equally distributed between the variants to ensure we have five testers for both variants. The probability of what test variant is randomly picked by the software changes whenever a new test is conducted.

Table 5.1: Age distributon of test users.

Age	Testers
20-29	5
30-39	3
40-49	-
50-59	1
60+	1

Before the test started, the facilitator gave a brief introduction to the application and what is likely to be seen in the test. For example, there are emojis sent automatically with the message, and the web camera is on during the test duration to analyze the test users' facial expressions. The tests were conducted with two laptops where the application was ensured to work. The facilitator had an Apple Macbook Pro and the test users used Apple Macbook Air. Test users and the facilitator sat down on the different tables where both testers had good natural lighting on the face. Tests were also performed during daylight hours to maximize the performance of the prototype application. The laptops needed to be as close to an eye-level that the laptops' web camera and the emotion analyzing software would work well. While the laptops had to be as close to the eye level as possible, it was also ensured that it was not taken to extremes and that the tester is comfortable writing with the laptop. The application was launched on both of the computers before the test started.

When the facilitator first wrote to the application, the test started, and the timer for a minimum of 15 minutes began. All tests took 15 to 25 minutes, depending on how much discussion of the topic was still ongoing when 15 minutes were over. The testing ended when the facilitator wrote to the prototype applications chat that the testing has ended. After the initial testing was over, users had to answer nine interview questions. The interview questions are presented in Appendice A. When all of the questions were answered and the facilitator had no other questions, the test was over.

It was required from the testers and test setting that:

- Test user and the facilitator should know each other.
- Test user have previous experience with an instant messaging application.
- Nether the test user or facilitator does not know which variant is chosen.
- The place where the tests are carried out is not public.

5.3 Interviews and Data Collection

Interviews are an effective method to solve peoples' experiences [41]. Therefore, data gathered from the tests are from the semi-structured qualitative interviews conducted after the tests. A semi-structured interview was chosen as the interview method because the facilitator can then ask follow-up questions to better understand the testers' answers and thoughts. Furthermore, the situation is more relaxed when the interview is more like a conversation. A semi-structured interview also has the good side that the questions have been made and carefully thought out beforehand. Questions are more like direction indicators in the interview and help the interviewer remember the things to ask. The interviews lasted approximately 20 to 30 minutes, but the interviewees usually stayed longer to discuss the topic after

all the questions were asked. The facilitator documented answers on his computer during the interview. The interviewer transcribed the documented text right after the interview to minimize the chance of forgetting something from the interviewees' answers because interviews were not filmed or recorded.

The common problem with an interview conducted by students and inexperienced researcher is that they have too many questions and ask too many questions during the interview [41]. It can sometimes strain and overload the interviewees. Strain in a way that they do not want to answer thoroughly and overload because there are too many questions. In the interviews conducted in the tests, test users were asked nine questions. The questions were kept open to induce better conversations. During a specific interview question, the interviewer had to give the interviewee examples to raise thoughts about the subject of that question. For example, in the interview question: "Are there any emotional states that you would not want the affect conveying application to transmit automatically?" the facilitator had to give examples of what emotional states there are. Three of these nine interview questions were questions where test users had to give a Likert scale grade from 1 to 5 according to how satisfied they were with the particular functioning of the prototype. The grading scale has a middle option of 3, which is considered a neutral answer. The grades were labeled before the interview as shown in Table 5.2. From now on, these three questions are referred to as interview questions 1, 2, and 3. The research method used in this thesis is a mix of two types of research methods, qualitative and quantitative. Interview as a qualitative research method and questionnaire conducted concurrently as a quantitative research method. The numeral grades were asked verbally. There were originally four questions for which the interviewees should give a grade, but the first tester noticed that one question could not be graded the way other questions can be.

After the testers gave the grade, they were asked to explain their numeral answer

Table 5.2: Grade and the corresponding label.

Grade	Label
1	Very dissatisfied
2	Dissatisfied
3	Neutral
4	Satisfied
5	Very satisfied

and give a more precise verbal answer. These three questions answered with a grade were chosen to get comparable data between the real and the random variant of the affect conveying prototype application. The three questions were asked first before the other questions. That way, the facilitator could get the tester's initial feelings and the feelings they felt in the test regarding the implementation before the test's experience disappears from the user's mind. These nine interview questions help us answer the research questions formed in this thesis, especially to the RQ2, which asks the possible use cases of this application. The interview questions were formed after the research questions were formed. That way, we ensure the interview questions are helpful for solving the research questions.

The topics covering in the interviews were: how well the application analyzed the test users affective state during the test, how well the method of showing the affects worked, did the tester experience any pressure or privacy concerns because of the emotion analysis, and are there any use cases or people that the tester would see using this kind of application.

Before the interview, test users were asked to answer as honestly as possible. It was found necessary because the test users and the facilitator already knew each other, and the test users may answer according to what the interviewer might want. During the interview, it did not appear that the interviewees were significantly afraid to answer honestly. Testers did not know about the testing mechanic where real

and randomized variant is randomly picked when the application launches. They only knew the information given briefly before the test (see Section 5.2), and the information did not consist of a randomization mechanic hidden in the prototype.

5.4 Data Analysis

The data gathered from the tests are from the open-ended interview questions and from three interview questions to which the interviewee had to give a numerical grade from 1 to 5 according to how satisfied the tester was with the subject of the interview question. The three interview questions give measurable data for answering research question 1. With this type of mixed methods research, we analyze the data with qualitative and quantitative methods. The test user sample size of N=10 for the quantitative data is too small for comprehensive statistical analysis. There is no reason to use these statistical methods if we consider the research questions or research problems. However, the numeral answers obtained from the interviewees can be analyzed and described with an informal interpretation of descriptive statistics, such as averages and percentages. Also, the numbers can be placed in the graphs and tables to visually present the answers, which gives an excellent addition to the interview answer analysis. The sample size of N=10 is so small that we cannot make generalizations based on the results to any broader population.

Interviews are often analyzed with different general methods of analysis [42]. There is no exact way of doing the analysis based on the data gathered from the interviews. The most common way is to start forming codes and themes from the data. That helps the researcher understand the interview answers and the test user's opinions, views, and experiences during the test and interview. The analysis of interviews started during the conducted interviews. The interviewer wrote down the thoughts and ideas that came to mind during the interviews in addition to the answers collected from the interviewee. The further the interviews continued and

more test users were interviewed, the same themes began to recur between the test users. That helped the researcher organize the data roughly into themes, form initial findings of the prototype, and test users' experiences. Coding and theming of the data were also done after the interviews were completed. The interview data was first coded with short descriptive codes to describe the sections in the data. That was done for all of the interviews individually. When all of the data had been gone through and coded, we formed themes from similar interview codes. All formed themes were collected, and the most relevant themes were selected for results and analysis. The themes are described in Chapter 7.

If the numeral measurable data obtained from the interviewees are crucial for answering RQ1, the user's verbal interview answers are analyzed to form the answer for RQ2. The analysis methods used in this thesis support each other in answering both of the research questions. Furthermore, they are made separately and used simultaneously along within Chapter 7.

Before going into the result and analysis section of this work, we have to know how the prototype application works, which we tested with test users. The next chapter goes through the prototype made for this thesis and the underlying mechanic and design of the application's affect conveying method.

6 Prototype

The main contribution of this thesis is the affect conveying instant messaging application prototype. We wanted to use ways within the affective computing topic to build and develop something that would help people to understand each other better. Computer-mediated communication (CMC) and instant messaging applications were picked to be the target of this work. CMC has been known for the problems of not being able to convey non-verbal communication in the discussion between two or more peoples [3]. The most popular mobile applications in the world right now are social media and instant messaging applications. Whether it is WhatsApp or Facebook Messenger, the affect information is not communicated to the other discussion participants the same way we humans communicate in face-to-face conversation, with such things as facial expressions, body posture, and the intonation of the vocal sound. That can lead to unpleasant misunderstandings between the discussion partners [43]. In this work, we have designed and developed an affect conveying instant messaging application which conveys the human affects with emojis. The application also has a test mechanic, which we go through later in this chapter.

The prototype aims to test how the test users feel about the affect conveying method implemented in the application, what kind of feelings and thoughts the affect analyzing gives to them, and whether they see any use cases for this kind of application in real life. Furthermore, how the interview answers differ between the two groups testing different variants (real and random).

6.1 Design

We had an idea to make text-based communication software with an affective module in it. It formed and refined to an instant messaging application as the design and fast prototyping phase continued. The original idea was to develop an affect conveying instant messaging application with a modern way of conveying the affect. There are application prototypes in the literature that convey emotions with colors, avatars, kinetic typography, or emojis [3]. There are just a few prototypes that convey the analyzed affect with emojis. For example, reaction-based affect conveying with using emojis or prototype where emojis are filtered regarding the affect analyzed by the analyzing software. In the filtering application, the user can choose the appropriate emoji conveyed to the other conversation participant from a set of emojis filtered by the affect analyzing software. The basic idea we had when we choose to use emojis was that people already use emojis to describe emotions in text-based conversations. So, could we do it then automatically? Users may filter the emotion conveying in messaging application to what they want to share or how they think they should react to the specific message. The real emotional state might not be the one that the user is conveying through emojis. Often users might give smiling emojis to be polite and keep the discussion in a positive state. People do the same in face-to-face conversation too. The most obvious example is in customer service. Customer service representatives may smile and seem excited during the situation without being overly happy or excited. That way, the customer gets a friendly feeling about the customer service.

With the previously presented idea in mind, we have developed a user interface, client-server connection, and affect conveying method for our prototype. The following subsections present their implementation.

6.1.1 User Interface

The instant messaging applications today influenced the user interface design of the test application. Like we mentioned in Section 4.1, user interfaces of instant messaging applications have not changed dramatically. The layout of the elements has stayed the same between applications. We are not going to break the design choices of those applications in this thesis because the layout of those applications has proven to work at least for 15 years. There has not been any indication of a change in the near future, at least when traditional user interfaces are considered. With that said, we have chosen the user interface elements in the test application from top to down to be the name box where the user writes her name, a text field where the messages appear, a text field for writing the messages, and the yellow send button. We can see the user interface in Figure 6.1.

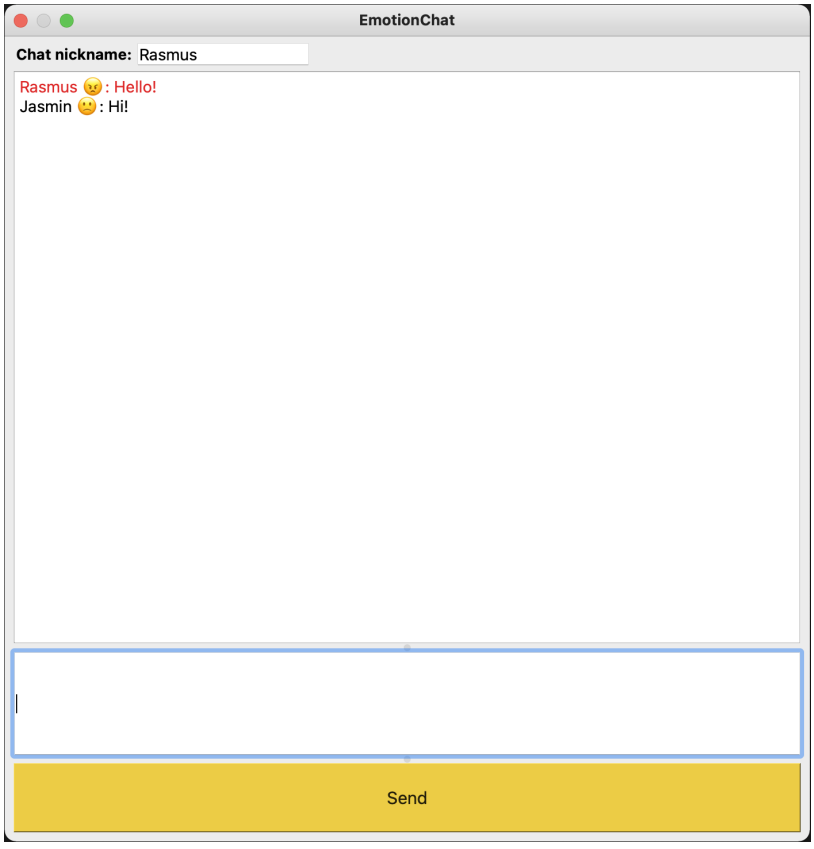

Figure 6.1: Affect conveying instant messaging application prototype

We also wanted the test application to be as clean and straightforward as possible not to distract the test users from the initial affect conveying test we wanted to make. That is why there are no sounds or other distracting things like animation implemented in it. The sent messages from test users appear in the largest text box located in the middle. Own sent messages are colored red, and the other participant's messages are in black. That helps the users to distinguish messages sent from the other conversation participant. The better method to distinguish the messages in the text box could be to separate the messages with individual boxes or show the messages in the left and right parts of the box regarding who is sending the message. That turned out to be too much work to change because it needed the whole user interface implementation to change, and that is why the text separation idea was omitted. We also find that comma and point do not work correctly in the prototype application when used in computers with macOS operating systems. Users have to press the spacebar twice to get space after comma and point. The problem does not exist in the Windows operating system. As with the name box where the test user writes their name, we could have made it the way it is not visible the whole time the discussion continues, but it is not so distracting that it would create any harm.

The test application's user interface is developed with the PyQt version 5 GUI (Graphical User Interface) framework. The PyQt5 is a Python binding for the Qt cross-platform C++ framework, and it was chosen for this project for its modern computing capabilities [44]. The primary reason the PyQt5 was used in this project as the GUI framework was that it supports the colored emojis by default. Tkinter, which is the de facto standard user interface in Python and comes default with the Python install, does not support colored emojis. Tkinter uses the Unicode's Basic Multilingual Plane (BMP), where all the most commonly used characters in computational devices are, like alphabets and basic special characters [45]. The emojis are in the Unicode's Supplementary Multilingual Plane (SMP), which Tkinter

does not support. The test application was first developed with the Tkinter, but it was not suitable for this application after noticing the emoji support problem. Overall, switching to the PyQt5 from Tkinter was a good decision because it is modern and works better within multiple platforms. Moreover, the same chat server implementation works with the new GUI without any changes.

6.1.2 Client-Server Connection

The prototype test application made for the experiment has two scripts for it to work, client and server scripts. It uses basic TCP sockets and threads to communicate between the clients. The server's responsibility is to connect all the clients to the server, receive messages from clients and broadcast the client's messages to every client connected to the server. When the server script is run, it also randomly chooses the test variant (real or random) the application uses for its affect conveying method. That information is sent to the client when it connects to the server. The client-side then know what variation of affect conveying should be used. Each client instance has its TCP socket connection to the server, and the server manages threads for multiple client sockets. The clients also have their threads to handle the server's communication and the user interfaces simultaneously. There is a thread for handling the socket and another thread for handing the user interface. The client also has a thread for the affect analyzing software PyEmotion. PyEmotion facial emotion detection is run in its own thread on the client-side to avoid user interface freezing because there are two processes, emotion detection, and the user interface, to handle at the same time. We go through the affect analysis part of the application in Section 6.1.3 in detail.

The client-server implementation is as simple as it can be so that we had more time to refine the affect analyzing and conveying, the main part of the test application. Also, the application is used in the local area network, so there will be no

problems with information security whatsoever.

6.1.3 Affect Analysing

If we consider the subject of this thesis and the experiment, the central part of the test application is the affect analyzing, conveying, and representation of affects inside the application. The experiment and the interviews both produce data which is an essential part of solving and answering the research questions presented in Chapter 5.

In the study of the ReactionBot application, researchers concluded that users feel like the affect or reaction does not come from the user but the ReactionBot [4]. The researchers felt like this could be fixed with more frequent affect sampling, analyzing, and better feedback-loop, which could perform quicker. Although the sampling and analysis were more frequent in the research conducted by Poguntke et al. [3], we wanted to test with this application, what if the conveying method is emojis, and not the colors, font, or profile picture changes. Emojis tend to feel more intuitive than, for example, colors. However, both have a high possibility of misinterpretation [32]. In the study conducted by Poguntke et al., the researchers had to give users the meanings of the colors to ensure that both conversation partners are on the same page regarding the colors [3].

For the facial emotion recognition software, we used the PyEmotion package acquired from the Python Package Index (PyPI) [46]. PyPI is a repository where developers and the Python community can share their Python packages with other developers. The requirement for facial emotion recognition software was that it should be easy to implement. It should have a pre-trained model in it, and it should also recognize emotions from web cameras' real-time video. In retrospect, emotion recognition software should have been more known, for example, Microsoft's Face API or the FER (Facial Expression Recognition) package. Moreover, the software

should have been more information on what type of training data has been used and how accurate it is to recognize emotions. However, FER, for example, does not recognize the emotions from the web camera's real-time video but the images taken from the video image or video file given to the analyzer [47]. Although taking images from the web camera's video image and then analyzing them would not have been a bad option, however, the PyEmotion package was easier to implement. The recognition accuracy of the PyEmotion package was not a concern when the prototype was made because when the emotion recognition accuracy of the PyEmotion package was tested empirically in different lights and positions, it turned out to be quite effective. However, that does not tell the truth of how well does it work without proper testing. The downside of the PyEmotion package is that it needs good natural lighting and angle of the image to perform well, but it can be solved with testing made in daylight and by positioning the testers so that the web camera is high enough compared to the face.

The PyEmotion uses and is built on top of the open-source machine learning library PyTorch to detect facial emotions. PyTorch convolution layers are used in PyEmotion to predict emotions from images or videos. Face recognition is done with pretrained Pytorch face detection and facial recognition models, and OpenCV library handles the laptop's camera. The technical information was not given by the developer, but the technology used in the package can be seen from the source code of the project. PyEmotion facial emotion recognition has a pre-trained model by the developer for recognizing emotions.

The PyEmotion software can recognize seven different emotions: happy, sad, neutral, angry, fear, surprise, and disgust. Neutral was also used as a distinct emotion. All the emotions were given an emoji describing them by the researcher. Emojis were selected from the Emojipedia website by searching with keywords such as happy, sad, and angry, and from what emojis are used in the affect conveying

😀	Happy
😕	Sad
😐	Neutral
😠	Angry
😨	Fear
😮	Surprise
😖	Disgust

Figure 6.2: Emotions and the corresponding emojis used in the prototype

instant messaging literature. The emojis used in this prototype are presented in Figure 6.2. Emojis are placed in the message after the user's chat nickname whenever the user sends a message. The place where the emojis appear changed few times during the design phase of the application. There was an idea to place them after the message to make it feel like the user was using them. It was then changed to appear after the user's name because if the user wants to use emojis by themself, it will mix with the automatically appearing emoji. There was also an idea that the emojis would appear in the top right corner of the message box surrounding the text sent in chat. As we mentioned in this chapter, the box surrounding the text was too hard to implement. Thus, this idea was also forgotten.

6.2 Affect Coveying Variants

As presented in Chapter 5, the prototype has a testing mechanic which randomly picks between two emotion conveying variants, real and random variant. The idea here is to compare the two variants and how the real variant performs compared to

the random variant. Next, we go through the variants more closely, starting with the real affect conveying variant.

6.2.1 Real variant

The real affect conveying variant is the variant that analyzes the test user's affective state in between the messages sent by the user. The PyEmotion module loop begins to work when the prototype starts. The loop continuously analyzes and predicts the emotional state from the user's face, and every two seconds adds the user's actual emotion predicted by the PyEmotion module to the list of emotions. The emotions added to the list are keywords such as happy, sad, angry, or disgust. When the user is ready to send a message in the chat and presses send button or enter, the prototype splits the list of emotions in half and chooses the end part. The application chooses the most common emotion from that list and inserts it into the message. That procedure repeats every time the user sends a message. The beginning of the list is deleted and no longer used at all. The idea behind selecting the end part of the list of emotions is that the user's reaction to the chat partner's message is at the end of the user-sent message interval. By selecting the end part of the list, we try to get the user's response to the chat partner's message instead of the user's messages. The decision of using just the end part of the list was chosen after the researcher tested it before the actual tests.

The operation logic of the real variant is described here step by step:

1. Loop starts to analyze facial expressions continuously when the prototype starts.

2. Every two seconds loop adds the word of what facial expression was expressed in the list of emotions.

3. When the user presses send, the list of emotions is split in half.

4. End part of the list is chosen.

5. Lastly, the most common emotion from that list gets picked to be conveyed and sent to the conversation partner.

6. The loop starts over with a new empty list when emoji and sent message appear in chat.

6.2.2 Random variant

The random affect conveying variant randomly picks emotion from the list of seven emotions (happy, sad, neutral, angry, fear, surprise, and disgust) used in this application and adds it to the message when a message is sent. The PyEmotion module launches and opens the web camera in the same way in real and random affect conveying variants. However, no emotions are analyzed or predicted in the random variant from the user's face to make it look like that the test user's facial expressions are analyzed by the prototype. There are no visual differences in the test application when the application uses a real or random affect conveying variant, except when sending messages. The variation of emojis that appears in the chat is larger because of the randomization. It is highly unlikely that the test user could understand or figure out that the test application has two affect conveying variants.

We have now taken a good look at the affect conveying instant messaging application prototype developed for this thesis. In the next chapter, we go through the results of the tests made for the prototype application.

7 Results

In this chapter, we go through the results formed from the interview data. The results consist of the tester's thoughts about how well the emotion analysis worked, what they thought about the emotion conveying method, the possible usage of this application, and the affect conveying feature. The results shown in this chapter are everything that is considered relevant for answering the research questions. We have gathered two types of data: test users' verbal answers from the semi-structured interviews and numeral grades asked from the first three interview questions. The test users were not recorded in the interview, so there are no exact and precise quotations of the test user's answers. In this chapter, the answers are written and referred to as close to the original answers as possible. The interviews were conducted in Finnish, and the answers are translated here into English. We start by looking at RQ1 and then continue with RQ2. The section which describes the results relevant to research question 1 is divided into three parts, separate section for each interview question 1, 2, and 3. To better understand whether the automatic affect conveying has any benefits when comparing real and random variants, we have to know how the software and its design worked in general. How users perceived the aspects of, how well the emotion analyses worked, and how well designed was the conveying method. Then we can more meaningfully try to understand, do the prototype has any benefits. All the test users quotes are named with P (participant) followed by the number, for example, P1. That way we ensure that the participants cannot be

identified from the answers. Before we go to the sections where all the interview answers results are presented more specifically, let us start by describing the initial findings and themes found from the interview data.

7.1 Initial Findings and Themes

We start by presenting the initial findings of the tests and interviews made. These initial findings were written down after each of the tests made and after all the tests were conducted. It was done to form a general idea of what we found during the tests and help further in the analyzing part. Answers in the interviews began to repeat themselves in the last tests, which indicates that the major flaws and what can be said from the application already saw daylight. For example, the test users who tested, especially the random affect conveying variant, were uncertain whether the emotion prediction worked. In the tests, the random variant test users did not notice that the emotions conveyed were randomized. Users just thought that emotion analysis and prediction do not work. When conducting the tests, we also noticed that test users could not see any specific, clear usage of this kind of application. The use cases in real life that were seen would be non-professional, like friends, close ones, or family. The test users who said that there could be some possible uses in professional life, like in healthcare between the professional and the customer, it should be brought out carefully what the application does so that the customer knows what is happening within the application. The results are shown in Section 7.3.

Many test users said that the automatic conveying of strong or negative feelings, such as angry, fear, and disgust, may be problematic without users being able to filter those emotions from not showing in the application. That is also noted in the literature of affect conveying instant messaging applications [3]. The outdated user interface design of the test application did get some comments during the interview.

The users generally thought that the text messages which appear in the chatbox had an outdated feeling and did not feel modern. In addition, some users were bothered by the outdated design of the chatbox and would have liked the texts would appear separately from each other.

7.1.1 Themes

Interview data was themed after all the interviews were conducted. That helps to form a general understanding of what the interviewees answered and whether the themes correlate with the interviews' initial findings. The theming was also a helpful way of finding out if there is something that has the researcher is forgotten during the analysis. Major themes found from the interview data are shown in the list below.

- Uncertainty of performance
- Untrust of experts
- Fun and interesting
- Mixed feelings on benefits across testers
- Mostly for non-professional use (half saw some benefits in professional use)
- No strong emotions

The themes were primarily the same as the initial finding made during the tests. Most of the test users in the control group had a feeling of uncertainty about the performance of the prototype. Test users were not sure does the applications affect analyzing and conveying work at all. The application did not have any technical difficulties, which would have affected the results, but especially the random variant test group users were sure that, for example, the affect analyzing from the face does

not work. Untrusting the experts were also a major theme in the interviews. Test users from both of the test groups thought they were skeptical of what the companies would do with the data gathered with the application. They would not like to use this kind of feature from a big company. The test users generally felt mixed feelings about the benefits of automatically conveying emotions between chat partners, at least in the test setting, and when they tried to imagine use cases where this kind of feature would be beneficial in their own lives and the lives of others. Some test users believed that the application could be used in non-professional lives, and half saw benefits in people's professional lives.

When the test users were asked what emotions they would not like to be conveyed automatically, most answered that they do not want strong emotions to be conveyed by default. Therefore, it should be optional, or users should be able to filter the emotions they do not want to convey. The emotions used in the automatic conveying could then be filtered according to the situation in the conversation and to whom they are talking. During the interviews some users found the application and its functionality to be fun and interesting.

7.2 RQ1: Benefits of Automatic Affect Conveying

From this section on, we go through the interview questions which the test users were asked to give a grade according to how satisfied they were about the subject of the question. The grades given by the test users are shown in the tables and graphs. We start by showing the results related to the first research question, which asked:

> *RQ1: Does automatic affect conveying with emojis benefit the instant message application users compared to a control group where affect conveying is randomized?*

It does not ask for any specific benefit but more general benefits of the application.

We will show the results of each three interview questions here to get an idea of whether the results of the real variant compare with the random affect conveying variant. The benefits were measured by the first three interview questions where the users had to give a grade according to how satisfied they were with the functioning and benefits of the prototype application. We also use verbal interview answers from the test users, which clarify the experiences of the testers.

7.2.1 Interview Question 1: How Well Emojis Describe Emotions

Interview question 1 asked the test users: how well did the emojis describe your current state of emotion? In Table 7.1, we can see the averages of the two groups real and random. The mean is 1.2 higher in the real group versus the random group. Also, the mode and median are one grade higher in the real group. The users were generally slightly more satisfied in the real variant group than in the random group, or at least the random group is in the dissatisfied area, and the real group leans more towards the neutral and satisfied values. That can also be seen in the explanation interview answers after the test users gave the grades. For example, one of the real variant test users P10 explained the grade given by saying that it gave suprisingly accure emojis during the test.

Table 7.1: Interview question 1: How well did the emojis describe your current state of emotion?

Test variant	N	Mean	Mode	Median
Real	5	3.4	3	3
Random	5	2.2	2	2

> "The application gave surprisingly accurate emojis during the test conversation. Whenever I pressed the wrong buttons while writing a mes-

> sage and got frustrated, the emoji placed in the message was as it should be, angry or sad. Or when I thought I made a mistake when writing a message, I laughed to myself and the emoji added to the message was expressing happiness.", P10

The answers for the interview question 1 were not that positive in general. The following answer from real variant test users describes the majority of answers and experiences given in a real variant group. The users communicated that they had some doubts about the performance of emotion analysis.

> "It gave angry or sad emoji without me being angry or sad. It gave more right emojis regarding the emotional state than the wrong emojis. I had little confusion during the test conversation on why there were wrong emojis appearing in the chat which, in my opinion, should not come at that very point.", P8

The test users grades for interview question 1 are more dissatisfied in the random group than in the real variant group, which can be seen in Table 7.2. 80 % of random variant testers were dissatisfied with how well the emojis described their current emotional state, and 20 % of them were neutral in their answer. On the other hand, 40 % of real variant testers were satisfied, and 60 % of them were neutral in their answer.

Table 7.2: Satisfaction percentages for interview question 1 (%).

Test variant	Satisfied	Neutral	Dissatisfied
Real	40	60	-
Random	-	20	80

In the random variant group, test users were generally dissatisfied with the performance of emotion analyzing, and some of the test users said that they were

confused that do they look like that what the emoji is telling. For example, one of the test users stated that they tried during the test how well the analysing performs and begin to doubt does it work at all.

> "I tested the performance of the emotion analyzing during the test conversation, and I began to doubt, do the software work at all.", P2

> "The emotion analysis did not go well. I began to feel that some of the emojis were correct, and on the other hand, I was confused that do I look like that. I was so focused on writing that I thought it was the reason why the emojis were not right.", P1

In Table 7.3 and in Figure 7.1, we can see how the grades were distributed between all the test users. It can be stated from the answers, tables, and figures that the real variant generally performed better than the random variant, and the test users experienced the emotion analyzing more positively on the real group than in the random group. However, the experience was not significantly better in the real group compared to the random group.

Table 7.3: Answers distribution to the interview question 1 (pcs).

	Real	Random
Very satisfied	-	-
Satisfied	2	-
Neutral	3	1
Dissatisfied	-	4
Very Dissatisfied	-	-

7.2.2 Interview Question 2: Ease of Reading Emojis

Interview question 2 asked: was it easy to follow the emojis or emotions of another participant in the conversation? The average values from the grades given to the

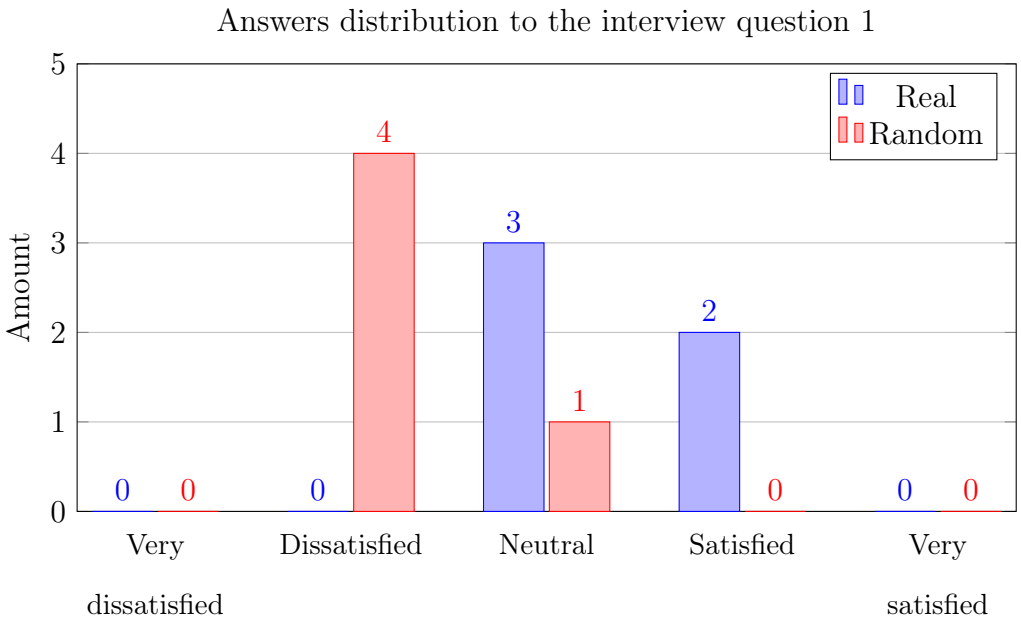

Figure 7.1: Answer distribution to the interview question 1

interview question 2 can be seen in Table 7.4. The calculated mean from the second interview question answers is explained by the fact that the amount of different emojis was not that wide compared to the random variant of the test than in the real emotion variant. It is easier to note and process the emojis when there is not so much to process. When there are too many different emojis, users have to look at them more closely and note the emojis' differences. For example, some test users said that the angry and disgust emojis were visually too close to each other. The disgust emoji never appeared in the application during the tests with the real affect conveying variant. One test user gave grade 1 to this question, which is different from other random group's test users answers. Therefore, the answer lowers the mean value quite much.

Test users were more satisfied with the affect conveying method in the real group than in the random group. 100 % of the testers in the real variant group were more satisfied with how easy it was to follow the emotions during the conversation. However, in the random variant test group, 40 % of the testers were satisfied, 40

Table 7.4: Interview question 2: Was it easy to follow the emojis/emotions coming from another participant in the conversation?

Test variant	N	Mean	Mode	Median
Real	5	4.4	4	4
Random	5	3	3, 4	3

% of testers were neutral about the subject, and 20 % were dissatisfied. See the results from Table 7.5. The answers exact distribution can be seen in Table 7.6 and in Figure 7.2

Table 7.5: Satisfaction percentages for interview question 2 (%).

Test variant	Satisfied	Dissatisfied	Neutral
Real	100	-	-
Random	40	20	40

Generally, all the test users in the real variant group were quite satisfied with emojis' visual presentation and the emotions presented by emojis. As one test user said in the real group:

> "The way the chat nickname and emojis order was made, it was easy to follow emojis. When the emojis always came to the same place after the name, the emojis were easy to follow. There were no problems with this.", P6

The test users in the random group generally said that emojis were somewhat easy to follow but gave lower grades. Many of them used the word "ok" to describe the feeling that the prototype gave from the affect conveying method. The comment from one random group test user described well what most of the users were thinking about the conveying method:

> "When there are many messages in the chat, the emojis are harder to see and separate from each other. Otherwise, it was ok", P3

Although, one user from the random group thinks that they were not easy to follow at all. The user described that the emojis were too close to each other, and they did not follow the emojis of the conversation partner because they were too busy to follow their own emojis appearing after their chat name.

Table 7.6: Answers distribution to the interview question 2 (pcs).

	Real	Random
Very satisfied	2	-
Satisfied	3	2
Neutral	-	2
Dissatisfied	-	-
Very Dissatisfied	-	1

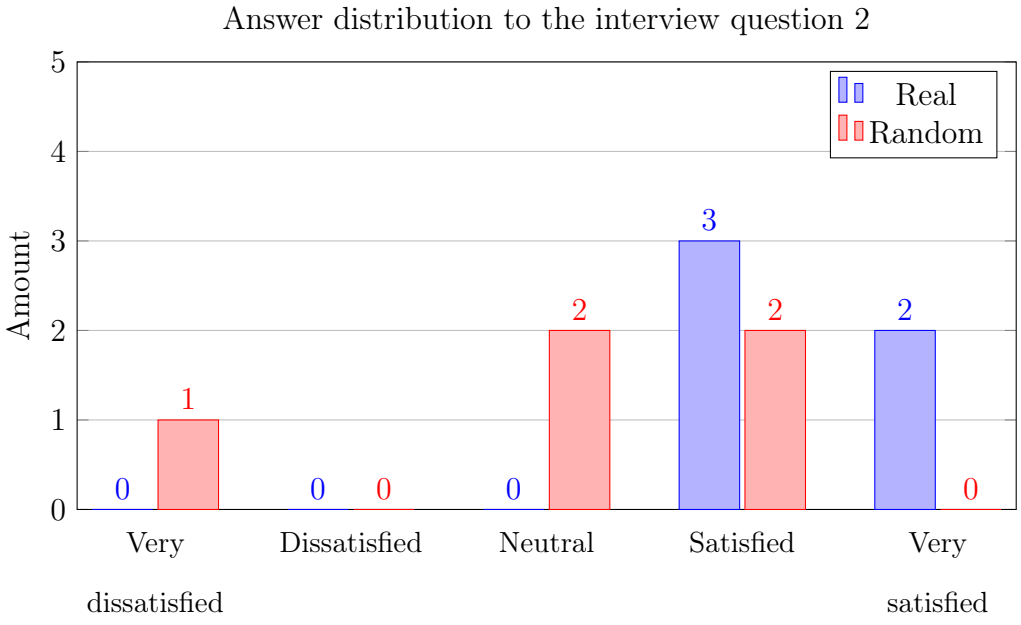

Figure 7.2: Answer distribution to the interview question 2

7.2.3 Interview Question 3: Understanding Messages Better

The last question the test users had to give a grade is the most relevant question related to research question 1. It asks whether the emojis from another conversation participant help to understand the messages better. The affect conveying instant messaging application made here was done to mitigate face-to-face conversations in computer-mediated communication. The average values do not differ significantly between test groups. We can see from Table 7.7 that the average grades are very close to each other between the real and random groups.

Table 7.7: Interview question 3: Did the emojis from another participant help to understand the messages better?

Test variant	N	Mean	Mode	Median
Real	5	2.8	3	3
Random	5	2.2	2, 3	2

The satisfaction ratings for interview question 3 were on the neutral and negative sides. 80 % of the testers in the real variant group were neutral in their answers, and 20 % were dissatisfied. 60 % were dissatisfied in the random variant group, and 40 % were neutral. What we can see from Table 7.8 and Table 7.9 is that from all of the 10 test users, no one thought that the application was not improving the understanding of the conversation partner's messages. From the real variant group, one test user said that:

> "It gave some benefit for the conversation. At least it was fun to follow them. However, did it give any significant help? I do not think so.", P6

The test users in the real group generally thought that the automatic affect conveying creates unpleasant misunderstanding situations. For example, one said:

> "It felt like if the emoji coming from me is angry, and I am not angry, then it will confuse another's thoughts. So if the application adds angry emoji

> in my message, I feel like I cannot explain that it was not describing my real feelings.", P7

The real variant produced many neutral emojis to the conversations. During the tests, the conversations were mainly full of neutral emojis and gave happy emojis from time to time. Furthermore, the application did give some emojis that were not expressed during the tests, such as sad or angry emoji.

> "I did not pay that much attention to it because the application gave so much neutral emoji. I also knew that the other whom I was messaging with did not go through terribly big emotions. A couple of happy emojis came during the test, which caught my eye.", P9

Table 7.8: Satisfaction percentages for interview question 3 (%).

Test variant	Satisfied	Dissatisfied	Neutral
Real	-	20	80
Random	-	60	40

Table 7.9: Answers distribution to the interview question 3 (pcs).

	Real	Random
Very satisfied	-	-
Satisfied	-	-
Neutral	4	2
Dissatisfied	1	2
Very dissatisfied	-	1

Test users in the random group were generally not satisfied with the functioning of the emotion recognition. Some of the test users gave a neutral grade, but the verbal answers were on the dissatisfied side. That lead to answers to interview question 3, which were quite the same in tone. For example, test users stated that:

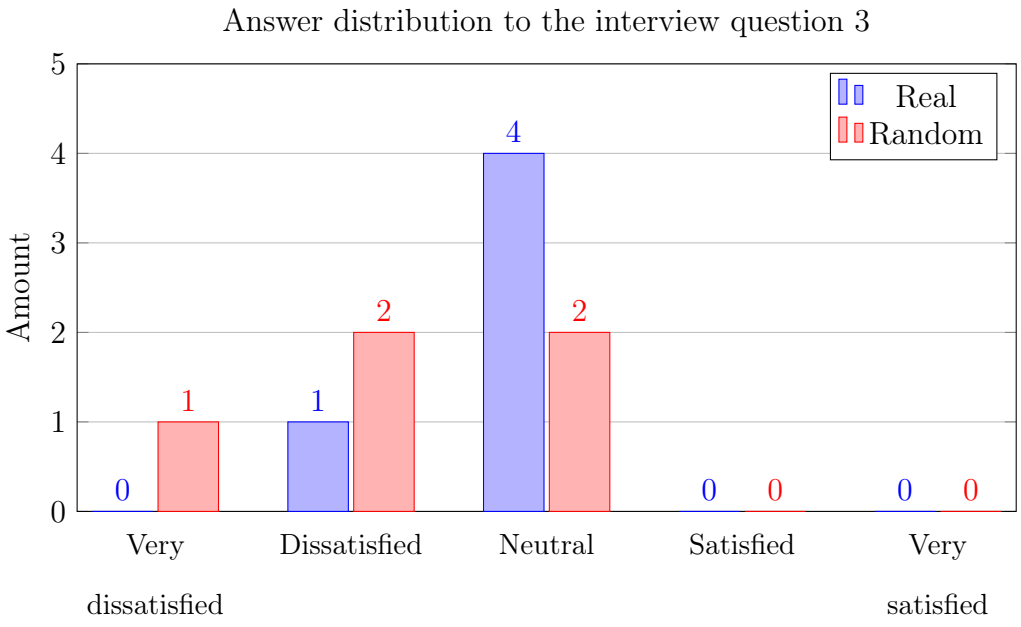

Figure 7.3: Answer distribution to the interview question 3

> "It did not help because I did not trust the application. If the application would work so maybe then.", P2

We can see the test users answer distribution for interview question 3 in Figure 7.3. Random variant test users had trust issues with the application. Moreover, because they had trust issues, they did not see any help with understanding messages better with emojis. The random variant test users also felt that it was fun to follow the emojis.

> "It did not help understand the messages better in this application. More for the sake of interest, it was fun to follow those emoticons.", P3

Overall, when pulling together the results from all of the interview questions 1, 2, and 3, which the testers had to give a numeral grade, the test users in the real group were more satisfied with the application's automatic affect conveying feature and its performance than the test users from the group which tested the application with the random variant affect conveying method. We can see the results from all

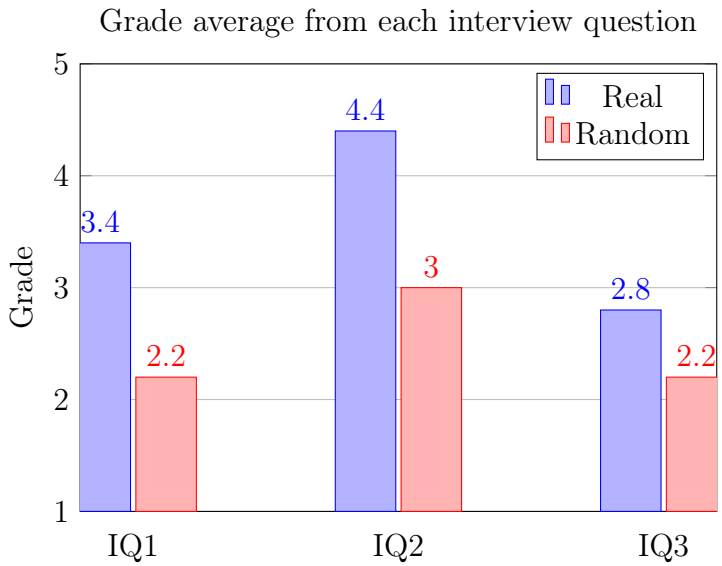

Figure 7.4: Grade average from each interview question

of the interview answers grade averages in Figure 7.4. These results are discussed more specifically in the next chapter, as well as the results gathered for research question 2, which we will bring out in the next section.

7.3 RQ2: Real World Use of Automatic Affect Conveying

To find out whether the automatic affect conveying in the instant messaging application has any use cases in free time, we have gathered related interview data from the test users interviews, which we use to answer the RQ2. The question is as follows:

> *RQ2: Does automatic emotion conveying have a place in an instant messaging application when the application is used in the free time?*

The RQ2 includes only the free time aspect because the tests were conducted with the free time setting. So that the conversation partners both know each other, and

the conversation that took place in the test application was informal.

The test users in both of the test groups were a bit unsure of where the application could be used, and also, the test users inside the same test group had a different feeling on the application usage. In the real variant group, four test users said this type of application and functionality could be used in free time, and two of the four mentioned possible usages in a professional context. Moreover, the one remaining test users also thought that this kind of functionality or application could be used in healthcare. That user did not see the need for it in their own life.

> "Maybe with the customers in the social work or workmates who know better and not with no so well known workmates. Could also be used with friends in free time.", P9

Most of the test users started to think of real-life use cases and how the application could add value to the work-life or communication between friends and family. The following quote from the random group test user gives another aspect to the discussion. The user saw possible use cases in the entertainment area, for example, in the virtual world and games.

> "Cannot see any possible use cases in a real-life context. It can be maybe used in the virtual world for not-so-serious functionality. For example, in the virtual world, your characters facial expression could be changing the same way the player's facial expressions change.", P6

Also, one user thought that this could be helpful in the online dating applications to enhance intimacy in them. The user also strongly pointed out that both users have to have a clear understanding of what the application analyzing does, and the functionality could not be used if one of the users does not want to use it.

Three out of five random group test users could see this application used with friends and family, and the remaining two could see some usages in professional life.

For example, test users from a random group said that:

> "Suitable for more informal communication, for example, with friends, family, and close ones. Does not suit for professional use.", P1

One of the test users thought that, for example, the instant messaging application Snapchat could have an automatic affect conveying functionality which could be used if both of the conversation partners agree to use it. It would be used as a filter that, for example, the Snapchat or Instagram applications have. With the filters, the users can, for example, change how they look in selfies, or the filters can add augmented reality objects to the picture or video. The automatic affect conveying filter could then be used as an element of entertainment and not for any serious use. One of the test users brought an idea of using this kind of application with people who have difficulties expressing and understanding one's emotions.

> "It may be helpful for those who have difficulty expressing emotions or understand them.", P4

Although more test users from the random group saw the application and its functionality to be used in more informal communication and in the free time, one of them thought that it could also be used in a professional context, for example, in healthcare or a job interview. That is not far-fetched if we consider that the job applicants, for example, are sometimes interviewed and analyzed by the psychologist in job interviews. However, like stated in the initial findings, if the emotion analysis is used, it has to have clear permission given from the patient or applicant.

> "Could be used in healthcare or job interview. If used in a professional context, the users should be carefully informed and explained that the application is analyzing their emotions.", P5

This ends the results chapter of this thesis. In the following discussion chapter, we will analyze the results and form a answers for the research questions according to the results presented in this chapter.

8 Discussion

The automatic affect conveying instant messaging application was made to determine what the everyday instant messaging application users thought about automatic affect conveying in text-based communication between two participants. Furthermore, do they feel they could understand the messages better than in discussion where affects are not automatically conveyed with emojis? With affective computing, we could make computer-mediated communication more natural by integrating automatic non-verbal communication into the conversation without users doing anything other than using the application like before.

The results indicate that the application's real variant performed better overall compared to the random variant. Test users thought that the emotion analyzing worked better in real variant compared to random, and users on the real variant group thought that the emojis were easier to follow. However, the real variant did not perform significantly better than the random variant, at least when comparing the grades given to the question, which asked how much it improved the understanding of the messages. Interview question 3, which found out was the messages more understandable when affect conveying was used, showed us that the automatic affect conveying did not help test users understand the messages better, at least in this test. The test users thought it gave "something" but could not describe what the "something" was. The grades and averages are very close to each other between real and random groups when considering the grades regarding interview question 3. It is

widely acknowledged that humans tend to understand better the information coming from other humans if the information is coupled with non-verbal communication like facial expressions, which was used in this research, or pose, and intonation of voice. However, that was not the case in this test. Test users in this research did not see significant help from automatic affect conveying when used in free time compared to today's instant messaging applications. For example, the test users in the real variant group described the affect conveying functionality as a fun feature to follow. As found in the research conducted by Liu et al. [4], we could say that the affect conveying functionality in the instant messaging application could increase the test user's emotional awareness of the chat partner and their own awareness of the experienced emotions during conversations. It allows or forces them to follow and interpret emotions from their own and from the conversation partners.

8.1 Use Cases

If we consider the use cases of automatic affect conveying instant messaging, most test users see more value in free time usage than professional usage. So, we can say that there are possible use-cases, but the test users do not see it significantly improve the instant messaging applications today when used in their free time. Answers were relatively the same in the real and random variant groups. The test users in both groups saw more use cases in the free time than in the professional life, and also, the distribution of responses was close to each other. Nothing significant can be said from the differences in the answers according to use cases of the application between the two groups. The test users also did not feel that the application could have clear uses in people's lives and significantly improve today's messaging applications. The interviews revealed that all of the test users understand to some degree what is the point of automatic affect conveying. Which then ensures that the testers could think of all the possible uses in the time of that short interview.

Interestingly, the interview answers concerning the possible use-cases are that a couple of the test users saw that the application could be used with people who have problems understanding emotions or expressing them. If the application is used in the healthcare context, we must ensure that the application is working as perfectly as possible. Using the application or automatic affect conveying, in general, creates worry if the emotion conveyed is wrong. It could create more harm than improvement to the training or treatment.

8.2 Comparing Existing Literature to Results

The results in this work somewhat agree with the results of previous research made on this topic. The random group users generally did not see any help from affect conveying, which was expected. The users who tested the random variant of the prototype generally thought that the conveying was confusing because it felt like the emotion analysis did not work. However, the real variant group users thought it was fun to follow emojis appearing in the application. The previous research regarding automatic affect conveying instant messaging has also found it to be a fun feature, according to the interviewed test users as it was found in this tests. Also, some test users have found affect conveying in instant messaging to be a clear improvement compared to traditional instant messaging. However, the test users in this research did not see that clear improvement in the communication with the developed prototype. What differs significantly from other studies compared to this research is that the tests were longer lasting than the tests conducted in this work. We could have seen more positive comments from the real variant test users, for example, on the question, how much did the emotion conveying help with understanding one's messages. Conducting tests, for example, few times a week, would have made it possible for test users to experience various discussion topics and form a broader spectrum of emotions during tests. That could then give the test user

a better understanding of the benefits of the application. The emotions analyzing method also varies between existing research. Some have made the analysis interval longer, and others have made the emotion conveying reaction-based. However, the affect conveying comes with a cost of privacy issues which the test users are also afraid of. For example, users are concerned what the company which created the application could use the emotional data gathered from the users. Privacy concerns are also discussed in-depth in the existing literature.

8.3 Suggestions for Improvement

During the tests and interviews, we saw a couple of subjects to improve in this research and the tests conducted. Some of the test users saw the test set to be too artificial. It created too much tension to test users. One of the testers even said that he feels tension like going into a school test. The test environment and setting should have been more natural and spontaneous so that the users could more freely express emotions, and the test could have worked better. Also, the test should be lasting longer, for example, days, to give a better feeling to the test users how much the automatic affect conveying help free-time conversations with friends and family. The feeling of artificial test setting and too short tests could have fixed by developing the chat application work within a web browser. Then, the test users could test the application, for example, in their homes whenever it suits both the tester and the facilitator. Then the tests would have been more natural compared to how it was conducted now that the facilitator had two computers and both sat in separate rooms to make this test. What was also considered to influence the results is that when the testers know each other and have had a conversation before in instant messaging applications with and without emojis, we already know what the other probably means. Like one of the test users in the interview said that they know that the other conversation partner does not experience strong feelings when

discussing the topic that was then discussed. That, too, could have been fixed with more spontaneous tests.

Neutral emotion or emoji was too dominant on the real variant tests, resulting in neutral emotion over shadowing other emojis in the chatbox. Although many emojis that appear in the chat represent neutral emotion, the other emojis came out visually better because the user saw something else popping at the chat. A better idea would be that if the list which gathers the emotion at one-second intervals when the application is used, and there is some other emotion than neutral, then the other emotion has more weight than neutral. For example, the list has [neutral, happy, neutral, neutral], then the emotion conveyed is happy, although a more neutral face is analyzed than a happy face. Also, if the emotion conveying happens every time the user presses enter or send the message, the text should be in a bubble, and emoji should appear, for example, in the corner of the chat bubble, clearly separated from the text and other emojis. The same type of design that the WhatsApp Web chat box has with the text bubbles (See figure 4.2). During the ideation phase of how the analysis of the emotions would work during the ongoing conversation in the prototype, it was seen as a difficult task to solve. When discussing with the test users after the tests about when the facial expression analysis should read the user's face and which emotions should be taken into account, different opinions were brought up on when the emotion is expressed during the conversation. For example, do we need to weigh the emotions expressed right after the user sends the message or while writing the message so that the emotions in some end would be more important?

Lastly, it could have been better if the question for which a grade was asked first, the interviewee would have given the questions, for example, in the form of paper, and asked the user to give and write the grading before the interview starts. There was a fear that test users do not dare to give an honest opinion, although it was made clear that users do not need to be afraid of giving the most honest answer.

The test users could have then had more time to think about the grades. It felt that the users were in a hurry to give the grades when they were asked.

8.4 Future Research

An interesting future development subject would be making the affect conveying instant messaging application work in the mobile platform. The emotion recognition from the face has to be better for analyzing emotion from various angles, and facial expression analysis could be backed up by having other emotion analyzing methods like analyzing emotions from written text. It is also considered that the analyzed emotions are more accurate when multiple modalities are analyzed and do not rely on one modality like facial expression. Generally, affects are better analyzed with multiple modalities, especially physiological like heartbeat and skin conductivity coupled with other affect analyzing methods. Those physiological signs are also easier to measure because of today's wearable technology. The everyday applications with affect analysis, people do not like to couple with physiological sensors because it usually needs some sensors taped to the human body and wires connected from the sensors to the computer. However, it changes as technologies and sensors evolve over the years, making them unnoticeable and wireless. Even today, smartwatches which can measure the heartbeat from the wrist are quite popular.

The other subject which came out during the research was to develop automatic affect conveying for games. We could improve the immersion of games by sensing players' emotions and automatically conveying them during players having a conversation or other activities that need two or more players to interact. Generally, making other software emotionally interactive through automatic affect conveying creates interesting possibilities in today's computer-mediated communication or even for human-computer interaction.

9 Conclusion

This thesis aimed to identify whether automatic affect conveying with emojis in instant messaging applications make the conversation feel more natural and closer to conversations that occur face-to-face. Also, does the affect conveying reduce the possibility of misunderstanding between interlocutors. Furthermore, does the application's real variant benefit the users compared to the random variant of the application. Moreover, are there any use cases for this kind of application in one's own life. To test the idea, we developed an instant messaging application that had two types of automatic affect conveying developed in it. The other variant was the real variant that tried to convey emotions as close as the users express them, and the other variant conveyed random emotion without truly analyzing any emotions from the user's face. The random emotion conveying variant, therefore, acted as a control group.

Based on the tests and interview results, users from the real variant group thought it was fun functionality. However, they did not feel it could improve the understanding of other's messages when both users know each other well. Random group test users felt that the emotion conveying functionality did not work, which negatively affected the interview answers. Based on the results, the real variant performed better overall than the random variant, and the emojis as an affect conveying method was intuitive and easy to follow. Real variant test users were more satisfied with the affect conveying functionality than random group users. Research question

1 asked: *Does automatic affect conveying with emojis benefit the instant message application users compared to a control group where affect conveying is randomized?* We can then conclude that the real affect conveying variant did give some benefits compared to the random variant, like increased the fun of using an instant messaging application. Although it increased the fun of using the application, it also raised the concerns of what if the emotion conveyed is something that the user does not want to convey and where the analyzed emotion data is going. The benefits were not so significant that the users would see them using the application or other application with affect conveying functionality.

Research question 2 asked: *Does automatic emotion conveying have a place in an instant messaging application when the application is used in the free time?* No clear use cases were found which could enhance the instant messaging application used today in people's free time. The test users thought it could be used with friends and family, but the users did not see why they would use it other than for the sake of fun. One test user thought that the affect conveying could work in other social applications like dating applications or games, which could make them more fun to use. On the other hand, there were seen the possible benefit for people who cannot understand other's emotions or express them. Moreover, users who thought that could be one use case were not experts in that field, and the thought came from intuition when they tried to figure out the uses cases in a professional field like healthcare.

References

[1] G. Gao, S. Y. Hwang, G. Culbertson, S. R. Fussell, and M. F. Jung, "Beyond information content: The effects of culture on affective grounding in instant messaging conversations", *Proc. ACM Hum.-Comput. Interact.*, vol. 1, no. CSCW, Dec. 2017. DOI: 10.1145/3134683.

[2] M. Jeon, *Emotions and Affect in Human Factors and Human-Computer Interaction*, 1st. USA: Academic Press, Inc., 2017.

[3] R. Poguntke, T. Mantz, M. Hassib, A. Schmidt, and S. Schneegass, "Smile to me: Investigating emotions and their representation in text-based messaging in the wild", in *Proceedings of Mensch Und Computer 2019*, ser. MuC'19, Hamburg, Germany: Association for Computing Machinery, 2019, pp. 373–385. DOI: 10.1145/3340764.3340795.

[4] M. Liu, A. Wong, R. Pudipeddi, B. Hou, D. Wang, and G. Hsieh, "Reactionbot: Exploring the effects of expression-triggered emoji in text messages", *Proc. ACM Hum.-Comput. Interact.*, vol. 2, no. CSCW, Nov. 2018. DOI: 10.1145/3274379.

[5] R. W. Picard, *Affective Computing*. Cambridge, MA, USA: MIT Press, 1997.

[6] R. A. Calvo, S. D'Mello, J. Gratch, and A. Kappas, *The Oxford Handbook of Affective Computing*, 1st. USA: Oxford University Press, Inc., 2014.

[7] B. Guthier, R. Dörner, and H. Martinez, "Affective computing in games", in. Oct. 2016, vol. 9970. DOI: 10.1007/978-3-319-46152-6_16.

[8] R. E. Jack, O. G. B. Garrod, H. Yu, R. Caldara, and P. G. Schyns, "Facial expressions of emotion are not culturally universal", *Proceedings of the National Academy of Sciences*, vol. 109, no. 19, pp. 7241–7244, 2012. DOI: 10.1073/pnas.1200155109.

[9] S. Weinschenk, *100 Things Every Designer Needs to Know About People*, 2st. New Riders, 2020.

[10] D. A. Sauter and F. Eisner, "Commonalities outweigh differences in the communication of emotions across human cultures", *Proceedings of the National Academy of Sciences*, vol. 110, no. 3, E180–E180, 2013. DOI: 10.1073/pnas.1209522110.

[11] R. Picard, "Affective computing", *M.I.T Media Laboratory Perceptual Computing Section Technical Report No. 321*, 1995.

[12] A. Sloman, "Review of: Affective computing", *AI Magazine*, vol. 20, pp. 1449–54, 1999.

[13] S. Alghowinem, M. AlShehri, R. Goecke, and M. Wagner, "Exploring eye activity as an indication of emotional states using an eye-tracking sensor", in *Intelligent Systems for Science and Information: Extended and Selected Results from the Science and Information Conference 2013*. Springer International Publishing, 2014, pp. 261–276. DOI: 10.1007/978-3-319-04702-7_15".

[14] J. Lim, J. Mountstephens, and J. Teo, "Emotion recognition using eye-tracking: Taxonomy, review and current challenges", *Sensors*, vol. 20, p. 2384, Apr. 2020. DOI: 10.3390/s20082384.

[15] K. Makantasis, A. Liapis, and G. N. Yannakakis, "From pixels to affect: A study on games and player experience", in *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2019, pp. 1–7.

[16] J. M. Garcia-Garcia, V. M. R. Penichet, M. D. Lozano, J. E. Garrido, and E. L.-C. Law, "Multimodal affective computing to enhance the user experience of educational software applications", *Mob. Inf. Syst.*, vol. 2018, 8751426:1–8751426:10, 2018. DOI: 10.1155/2018/8751426.

[17] *Ageing*, accessed 2021-01-12. [Online]. Available: https://www.un.org/en/sections/issues-depth/ageing/.

[18] H. Silfvernagel, *Ai talks with håkan silfvernagel - affective computing*, Nov. 2020. [Online]. Available: https://www.youtube.com/watch?v=Xu1_O62LMN8 (visited on 03/02/2021).

[19] E. Hudlicka, "Affective computing for game design", *North American Conference on Intelligent Games and Simulation (GAMEON-NA)*, pp. 5–12, 2008.

[20] R. V. Aranha, C. G. Corrêa, and F. L. S. Nunes, "Adapting software with affective computing: A systematic review", *IEEE Transactions on Affective Computing*, pp. 1–1, 2019. DOI: 10.1109/TAFFC.2019.2902379.

[21] T. Henriques, S. Silva, S. Brás, S. C. Soares, N. Almeida, and A. Teixeira, "Emotionally-aware multimodal interfaces: Preliminary work on a generic affective modality", in *Proceedings of the 8th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-Exclusion*, ser. DSAI 2018, Thessaloniki, Greece: Association for Computing Machinery, 2018, pp. 80–87. DOI: 10.1145/3218585.3218589.

[22] J. A. Rajendran, H. Baharin, and F. Mohmad Kamal, "Understanding instant messaging in the workplace", in *Advances in Visual Informatics*, H. Badioze Zaman, A. F. Smeaton, T. K. Shih, S. Velastin, T. Terutoshi, N. Mohamad

Ali, and M. N. Ahmad, Eds., Cham: Springer International Publishing, 2019, pp. 640–652.

[23] H. Tankovska, *Whatsapp - statistics and facts*, Feb. 2021. [Online]. Available: https://www.statista.com/topics/2018/whatsapp/ (visited on 03/12/2021).

[24] E. Protalinski, *Windows live messenger has 330 million active users*, Jun. 2009. [Online]. Available: https://arstechnica.com/information-technology/2009/06/windows-live-messenger-has-330-million-active-users/ (visited on 03/12/2021).

[25] A. Chowdhry, *Msn messenger is completely shutting down on october 31st*, Aug. 2014. [Online]. Available: https://www.forbes.com/sites/amitchowdhry/2014/08/31/msn-messenger-is-completely-shutting-down-on-october-31st/ (visited on 03/19/2021).

[26] B. Petrovan, *Whatsapp now has windows and mac desktop apps*, May 2016. [Online]. Available: https://www.androidauthority.com/whatsapp-windows-mac-desktop-apps-691901/ (visited on 03/19/2021).

[27] *Want some facebook with that skype? new windows client adds im support*, Jun. 2011. [Online]. Available: https://techcrunch.com/2011/06/20/want-some-facebook-with-that-skype-new-windows-client-adds-im-support/?guccounter=1 (visited on 03/19/2021).

[28] *Emoji and pictographs*, 2021. [Online]. Available: https://www.unicode.org/faq/emoji_dingbats.html (visited on 03/10/2021).

[29] S. H. Hsieh and T. H. Tseng, "Playfulness in mobile instant messaging: Examining the influence of emoticons and text messaging on social interaction", *Computers in Human Behavior*, vol. 69, pp. 405–414, 2017. DOI: https://doi.org/10.1016/j.chb.2016.12.052.

[30] D. J. Skiba, "Face with tears of joy is word of the year: Are emoji a sign of things to come in health care?", *Nursing Education Perspectives*, vol. 37, no. 1, pp. 56–57, 2016. DOI: 10.1097/01.NEP.0000476112.24899.a1.

[31] G. Troiano and N. Nante, "Emoji: What does the scientific literature say about them?-a new way to communicate in the 21th century", *Journal of Human Behavior in the Social Environment*, vol. 28, no. 4, pp. 528–533, 2018. DOI: 10.1080/10911359.2018.1437103.

[32] G. W. Tigwell and D. R. Flatla, "Oh that's what you meant! reducing emoji misunderstanding", in *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, ser. MobileHCI '16, Florence, Italy: Association for Computing Machinery, 2016, pp. 859–866. DOI: 10.1145/2957265.2961844.

[33] *Word of the year*, 2021. [Online]. Available: https://languages.oup.com/word-of-the-year/ (visited on 03/10/2021).

[34] *Emojipedia*, 2021. [Online]. Available: https://emojipedia.org (visited on 03/10/2021).

[35] A. R. Lebbon and J. G. Sigurjónsson, "Debunking the instant messaging myth?", *International Journal of Information Management*, vol. 36, no. 3, pp. 433–440, 2016. DOI: https://doi.org/10.1016/j.ijinfomgt.2016.02.003.

[36] J. A. Sánchez, N. P. Hernández, J. C. Penagos, and Y. Ostróvskaya, "Conveying mood and emotion in instant messaging by using a two-dimensional model for affective states", in *Proceedings of VII Brazilian Symposium on Human Factors in Computing Systems*, ser. IHC '06, Natal, RN, Brazil: Association for Computing Machinery, 2006, pp. 66–72. DOI: 10.1145/1298023.1298033.

[37] Z. Yeo, "Emotional instant messaging with kim", in *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '08, Florence, Italy: Association for Computing Machinery, 2008, pp. 3729–3734. DOI: 10.1145/1358628.1358921.

[38] R. El Kaliouby and P. Robinson, "Faim: Integrating automated facial affect analysis in instant messaging", in *Proceedings of the 9th International Conference on Intelligent User Interfaces*, ser. IUI '04, Funchal, Madeira, Portugal: Association for Computing Machinery, 2004, pp. 244–246. DOI: 10.1145/964442.964493.

[39] J. T. Hancock, K. Gee, K. Ciaccio, and J. M.-H. Lin, "I'm sad you're sad: Emotional contagion in cmc", in *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, ser. CSCW '08, San Diego, CA, USA: Association for Computing Machinery, 2008, pp. 295–298. DOI: 10.1145/1460563.1460611.

[40] J. T. Hancock, C. Landrigan, and C. Silver, "Expressing emotion in text-based communication", in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '07, San Jose, California, USA: Association for Computing Machinery, 2007, pp. 929–932. DOI: 10.1145/1240624.1240764.

[41] M. Hyvärinen, E. Suoninen, and J. Vuori, *Haastattelut*, 2021. [Online]. Available: https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/laadullisen-tutkimuksen-aineistot/haastattelut/ (visited on 03/29/2021).

[42] K. Günther, K. Hasanen, and K. Juhila, *Johdanto: Analyysi ja tulkinta*, 2021. [Online]. Available: https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/analyysitavan-valinta-ja-yleiset-analyysitavat/analyysi-ja-tulkinta/ (visited on 04/08/2021).

[43] G. Gao, S. Y. Hwang, G. Culbertson, S. R. Fussell, and M. F. Jung, "Beyond information content: The effects of culture on affective grounding in instant messaging conversations", *Proc. ACM Hum.-Comput. Interact.*, vol. 1, no. CSCW, Dec. 2017. DOI: 10.1145/3134683.

[44] *What is pyqt?*, Jul. 2021. [Online]. Available: https://riverbankcomputing.com/software/pyqt/intro (visited on 07/21/2021).

[45] *Tkinter — python interface to tcl/tk*, Jul. 2021. [Online]. Available: https://docs.python.org/3/library/tkinter.html (visited on 07/21/2021).

[46] K. Nagarajan, *Pyemotion 1.0.0*, Aug. 2020. [Online]. Available: https://pypi.org/project/PyEmotion/ (visited on 04/01/2021).

[47] S. Justin, *Fer*, Jul. 2021. [Online]. Available: https://github.com/justinshenk/fer (visited on 07/21/2021).

Appendix A Interview questions

Nine different questions are asked during the interviews conducted after the tests. The first three questions are questions in which the interviewee had to give a grading from 1 to 5 and then verbally explain the grading. The interviewees are also free to say anything that comes to their mind from the subject during the interview. Questions with 1-5 grading:

- How well did the emojis describe your current state of emotion?
- Was it easy to follow the emojis or emotions of another participant in the conversation?
- Did emojis help you understand your conversation partner's messages better?

Original questions with 1-5 grading in Finnish:

- Kuinka hyvin mielestäsi emojit kuvasivat sen hetkistä tunnettasi?
- Oliko toisen osapuolen tunteita/emojeja helppo seurata keskustelun aikana?
- Auttoiko emojit ymmärtämään vastapuolen viestejä paremmin?

Questions without grading:

- During the conversation, did you feel pressure because of facial recognition OR when the camera was on all the time?

- Did you feel that the app violated your privacy because of the filming? For example, worrying about where the data is going.
- Do you think you recognized all the emojis and what emotions they described?
- Are there any emotional states that you would not want the application to transmit automatically?
- What purpose do you think the application would be used for? For whom do you think the application might be best suited?
- What do you think could be improved in the application? UI/UX, emotion transmission.

Original questions without grading in Finnish:

- Tunsitko keskustelun aikana paineita kasvojentunnistuksen takia TAI kun kamera oli päällä koko ajan?
- Koitko että sovellus rikkoi sinun yksityisyyttäsi kuvaamisen takia? Esim. huoli siitä mihin tiedot menevät.
- Tunnistitko mielestäsi kaikki emojit ja mitä tunteita ne vastasivat?
- Onko jotain tunnetiloja, joita et haluaisi, että sovellus välittäisi automaattisesti?
- Mihin käyttötarkoitukseen sovellus mielestäsi kävisi? Kenelle sovellus sopisi parhaiten?
- Mitä parannettavaa sovelluksessa mielestäsi olisi? UI/UX, tunteiden välitys.