Title	Empathy enhancement based on emotion recognition and presentation in written communication
Sub Title	
Author	韩, 嘉雯(Han, Jiawen) Kunze, Kai
Publisher	慶應義塾大学大学院メディアデザイン研究科
Publication year	2019
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2019年度メディアデザイン学 第775号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40001001-00002019-0775

Master's Thesis
Academic Year 2019

Empathy Enhancement Based on Emotion Recognition and Presentation in Written Communication

Keio University
Graduate School of Media Design

HAN JIAWEN

A Master's Thesis
submitted to Keio University Graduate School of Media Design
in partial fulfillment of the requirements for the degree of
Master of Media Design

HAN JIAWEN

Master's Thesis Advisory Committee:

Professor Kai Kunze	(Main Research Supervisor)
Senior Assistant Professor	
Marcos Sadao Maekawa	(Sub Research Supervisor)

Master's Thesis Review Committee:

Professor Kai Kunze	(Chair)
Senior Assistant Professor	
Marcos Sadao Maekawa	(Co-Reviewer)
Project Senior Assistant Professor	
Professor Masato Yamanouchi	(Co-Reviewer)

Abstract of Master's Thesis of Academic Year 2019

Empathy Enhancement Based on Emotion Recognition and Presentation in Written Communication

Category: Science / Engineering

Summary

In my master's research, one type of possible interactive way for enhancing empathy was explored. According to the mechanism of empathic interaction, the method based on providing people with more accurate and real emotional information behind the words in written communication. To verify the viability of the approach, the feasibility of both the in-put stage (emotion recognition via handwriting features) and for the out-put stage (emotion presentation) was tested in two separate experiments. The user-independent Support Vector Classifier (SVC) for emotion recognition shows up to 66% accuracy for certain types of writing tasks for 1 in 4 classification (1. High Valence, High Arousal; 2. High Valence, Low Arousal; 3. Low Valence, High Arousal; 4. Low Valence, Low Arousal). The Effects of four types of Emotion Presentation which are Color, Word, Shape, and Graph have been tested. The out-put stage's results show the possibility of changing the perception of the emotional coloring of texts in terms of both, valence and arousal. While future work is required to improve the classification rate and effectiveness of Emotion Presentation, this work should be seen as a proof-of-concept for empathy enhancement in written communication.

Keywords:

Human Computer Interaction, Affective Computing, Computational Empathy,Emotion Recognition, Emotion Presentation

Keio University Graduate School of Media Design

HAN JIAWEN

Contents

List of Figures

List of Tables

Acknowledgements

I would first express my sincere gratitude to my supervisor Professor Kai Kunze who is always supporting and encouraging my conducting own research. Without his guidance and valuable advice, I would not find my enthusiasm in academia and conduct my research. Professor Kai Kunze not only provides me with professional remarks but also opportunities to join the community where I can work with top researchers as well. This can always motivate me in my future research work. I would love to thank the graduate school of Media Design for the great platform for collaborations and innovations and my sub supervisor, Senior Assistant Professor Marcos Sadao Maekawa, for the useful advice and meaningful questions.

In addition, I would love to thank Professor Katrin Wolf and Professor Takuji Narumi for their professional advice about emotion recognition and paper writing. Also, I want to say thank you to my fellow teammates and those who were willing to spend their valuable time in the experiments.

Last but not least, I would like to thank my family and friends who are always supporting me even when I lost confidence myself. I will be grateful for all your help, kindness, and love.

Chapter 1
Introduction

1.1. Background and Motivation

Empathy, as the "glue" of the social world, plays a pivotal role in communications since we live in company of others. Only when people take others' feelings into consideration, can the interaction go smoothly. During the past few decades, computers have also become ours companions with the advent of technology. Likewise, the empathic interactions between human and computers have been been gaining increasingly more attention in the field of affective computing and the HCI community [1–5]. Recent studies have explored possibilities of computers empathic interactions with human both as empathizee and empathizer [6–9]. Therefore, I assume that computer can play as a mediator to enhance the empathy between human and human during communication.

Among various types of communication, it is typical that written communication has more risk in losing emotional hints without supplementary information such as mimic, voice tone and gestures in other communications. Consequently, the emotions related to messages, and even the entire message, is every so often misunderstood. Since the essence of empathic interaction is accurate understanding of another's feelings [10], helping people understand emotions behind the words can bring noticeable empathy enhancement in written communication. Moreover, according to the perception-action theory, it is suggested that perception and even imagination of someone else's emotional state automatically activate the shared presentation of perception and action, thus leading to an empathic response [2, 11].Therefore, my master's project is to test the feasibility of improving people' empathic understanding by presenting emotions which recognized via handwriting features.

Different from traditional methods of emotion recognition such as physiological sensors and facial expressions, I am trying to recognize and predict emotional states by extracting features from handwriting. Handwriting can be seen as a fine motor task highly dependant on procedural memory (responsible for "muscle memory"). The reliance on procedural memory rather than deliberate conscious control of every hand muscle movement presumes the low degree of conscious control over the handwriting process, which would make it possible for unconscious manifestation of certain emotion-related features. This could be explained by the change in neurotransmitter quantities associated with emotional experiences, particularly the monoamine family: serotonin, dopamine, and noradrenaline [12]. In addition to this, there are several studies demonstrating the effects of Autonomic Nervous System (ANS) activity and neurotransmitter balance on motor activity in humans as well as in birds and other mammals [13–15]. So, handwriting is supposed to have the sensitivity of an emotional barometer and can give an idea about the excitement, fear, anxiety, irritability, depression, anger, and other such emotions [16].Although the decline of handwriting is undeniable, yet handwriting still has its place when people are conveying sincere, emotional and personal messages. Moreover, there are many recently released tablet products that encourage people in handwriting and drawing by hands such as Apple Pencil and Microsoft Surface Pen. Those products are not only commercially successful but also scientifically convenient since they can provide handwriting related data in a more detailed and objective way.

The test was divided into two separate experiments consisting of emotion recognition as the input stage and emotion presentation as the output stage. Both of the stage were based on the Russell's Circumplex Emotion Model and the Self-Assessment Manikin (SAM) for assessing and quantifying emotional feelings. For the input stage, whether emotional states can be recognized and predicted from handwriting features was tested. Handwriting data with emotion labels were used for training and prediction. A user-independent Support Vector Classifier with up to 66% classification accuracy for certain types of handwriting for 1 in 4 classification has been built. The accuracy is much more accurate than random which is 25% and this can prove the potential of recognizing emotions from handwriting features. For the output stage, whether visualisation of specific emotional meanings can affect people's perceiving of emotions was tested. In this experiment, Emotion Presentation represents the visualizations added to provide supplementary emotion cues. There are four types of Emotion Presentations selected from four typical types of emotion stimuli which were Color, Word, Shape and Graph. Every Emotion Presentation possesses its own emotional meanings based on its SAM scores. By comparing values of the Original Text, Emotion Presentation and Self-report, the effectiveness of four types of Emotion Presentation was tested. Also, subjects' empathy levels were measured by the short forms of Empathy Quotient (EQ-Short) to evaluate different performances of people with different empathy levels. The results proved the capability of using Emotion Presentation to change reader's emotion understanding towards texts written. It was especially effective in lowering the perceived intensity of emotions in both valence and arousal scales. In addition, people with lower empathy level was found to be more successfully led to emotion perception changes compared with those with high empathy level.

Although, both the emotions' relative levels, Emotion Presentation types and handwriting tasks were restricted. The results of two separate experiments together can prove the feasibility of improving people' empathic understanding by presenting emotions which recognized via handwriting features. In the future, the classifier algorithm will be polished and the system will be built for continuing test.

1.2. Contributions

Contributions of my master thesis are the following:

- Describing the physiological mechanisms for emotion assessment and empathy enhancement.
- Presenting the feature set extracted from the digital handwriting recording which is enable to recognize the emotional condition of the writer.
- Presenting a user independent classifier with up to 66% classification accuracy for certain types of handwriting. And the results were organized into a full paper published in the 10th Augmented Human International Conference 2019.

- Presenting the modified version of Affective Norms of English Texts in first person tone.
- Verifying the ability of Emotion Presentation types of Shape and Graph in promoting empathic understanding.

1.3. Structure of the Thesis

The thesis is consisted of 6 Chapters in total. Following this Chapter, Related Works about Theories of Emotion, Theories of Empathy, and Affective Computing are presented in Chapter 2. Later in Chapter 3,the Concept and Hypothesis of the whole system and two experiments are explained. Chapter 4 and Chapter 5 describe two Experiment Designs and their corresponding Evaluations and Results. The Conclusion with possible Future works are explored in Chapter 6. The questions used to measure Empathy Level are attached in the Appendix as well.

Chapter 2
Related Works

2.1. Theories of Emotion

2.1.1 Neural Substrate of Emotion

Skin Conductance (SC also known as EDA), Heart Rate Variability (HRV), respiratory rate, and pupil dilation, are all controlled by the Autonomic Nervous System (ANS). ANS is the part of the peripheral nervous system that controls involuntary functions that are critical for survival. Within the brain, the ANS is located in the medulla oblongata in the lower brainstem. During emotional stimulation, projections from the central nucleus or the bed nucleus of the stria terminalis (BNST) to the ventral tegmental area appear to mediate increases in dopamine metabolites in the prefrontal cortex [17].

Most of the methods of implicit assessment of the emotional state of the subject without self-assessment rely on recording physiological signals related to the ANS activity. These approaches seem very reasonable, as there are multiple works demonstrating how emotions influence the ANS [18]. In many evolutionary researchers, emotion is even considered to organize the activity of ANS [19]. This, combined with the lack of voluntary conscious control over the ANS, makes ANS a very interesting source of information regarding the cognitive or emotional state of the subject.

With significant advances in the ways that ANS activity can be detected and quantified, there are more possible methods to detect emotions based on the ANS-mediated changes. There are some visible ANS-mediated changes in emotional situation which are more obvious to see. For example, basing on the ANS-mediated change of sweat glands, fear will result in sweating and clamminess. Basing on the ANS-mediated change of salivary glands, disgust will result in salivating and drooling [20]. Besides those physiological changes, motor activity in humans as well as in birds and other mammals has been demonstrated that its connection with ANS activity and neurotransmitter balance [13–15]. Handwriting, which can be seen as a fine motor task highly dependant on procedural memory (responsible for "muscle memory"), thus is possible for unconscious manifestation of certain emotion-related features. So in my research, handwriting features related to motor activity and muscle memory will be extracted and later used for emotion recognition and interpretation.

2.1.2 Models of Emotion

The assessment and classification of emotion has received considerable attention in the field of psychology. Emotions can affect people's life in many aspects ranging from decision making to both mental and physical health. Though everyone seems to know what emotion is indistinctly, the fundamental concept of emotion is hard to define.

For decades, researchers have been continually developing alternative emotional measurement methods and technologies. There are several model of emotions that are widely accepted and validated. Mehrabian and Russel developed the PAD emotional state model to represent all emotions. the PAD model uses three dimensional scales: Pleasure/Valence, Arousal and Dominance [21]. Russell developed the Circumplex Model based on the PA part of PAD. In this model the emotions are distributed in a two-dimensional circular space, containing arousal and valence dimensions [22]. In this model, emotional states can be represented at any level of valence and arousal, or at a neutral level of one or both of these factors. PAD is also used by Lang and colleagues to develop a non-verbal pictorial self-assessment also known as SAM [23].

In the early 60s, Tomkins' proposed the eight basic emotion theory in which he defined eight basic emotions as follows. Two positive: Interest/ excitement and enjoyment/ joy, one neutral: Surprise/ startle, and five negative: distress/ anguish, fear/ terror, shame/ humiliation, contempt/ disgust and anger/ rage [24]. Based on Tomkin's affect theory, Lovheim proposed a new three-dimensional model for emotions and monoamine neurotransmitters also known as Lovheim cube of emotion [12]. In this model, the eight basic emotions located in each corner of an orthogonal coordinate system with Serotonin (5-HT, 5-hydroxytryptamine) is represented on the x-axis, noradrenaline (NE) on the y-axis and dopamine (DA) on the z-axis.

In scientific studies, emotion is best defined as a process that involves multiple responses rather than a single reaction [25]. Emotional processing involves both attentional engagement and behaviors preparatory to motivated action. Lang indicated there were at least three different systems in emotional response measuring: affective reports, physiological reactivity, and overt behavioral acts [26]. He also noted both physiological reactivity and overt behavioral acts often dictate clear preference for the scientific recordings [27]. As matter of fact, many prior studies demonstrated the possibility of emotion detection based on physiological signals such as SC, HRV, pupil dilation, respiration, temperature, etc. For instance, SC is considered as one of the most sensitive markers and is frequently used to assess emotional arousal. Many studies found Skin Conductance Level (SCL) increases when emotional arousal increases [28–30].

2.1.3 Emotional Meaning of Visual Presentation

A large number of researches have been conducted to understand the emotional context of different types of visual presentation. Word is considered as a direct and initial way for expressing and triggering emotions. The words that we used to express the character of objects can also be simplified into their related emotions [31]. In Plutchik's emotion theory, 8 primary emotions are described as 8 words which are joy,surprise, trust, anger, fear, disgust, sadness and anticipation [32]. Plutchik's three-dimensional Circumplex model not only describes the relations among emotion concepts, but also provides an analogy to color wheel, which shows the relevance between emotion and colors as well [33].In the Russell's Circumplex Model,the emotions are distributed gradually in a two-dimensional circular space, containing arousal and valence dimensions [22], which shows the connections between subtle emotions and emotion related words.

In addition, certain shapes are able to evoke emotional perceptions. Kohler reported the bias when people match "maluma" and "takete" to rounded and angular shapes [34] showing that common perception exits in simple lines and shapes.Collier and Poffenberger's research are in pivotal roles confirming that subjects can map emotions onto simple stimuli reliably [35].Experiments conducted by Poffenberger and Barrows indicates line characters such as direction,rhythm and form can represent different feelings and the distribution was mapped into valence-arousal dimensions by Collier later.Besides lines, shapes also convey some emotion information. Shapes with sharp transition in contour might elicit negative affect and threat while rounded shapes induced a more positive mood [36] [37].Moreover, 3D models also have the ability to present emotional feelings. A CAD tool developed by Mothersill,mainly built on Plutchik's model of primary emotions and Russell's circumplex model as well as shape-perception theory, shows a system can convey emotion contexts to 3D shapes [38].

2.2. Theories of Empathy

2.2.1 Definition and Concept

Studies pointed out the discussion of empathy can date back to "the beginnings of philosophical thought" [39], while the term "empathy" was coined over 100 years ago by Titchener, an adaptation of the German word "Einfuhlung". However, empathy has an inherent ambiguity and is often merged with associated terms such as sympathy. This may result in empathy is locating in the midrange for all three dimensions which are the degree of cognitive representations of the target' emotional state, the degree of emotion sharing and the degree to which a self/other distinction is maintained according to Ickes et al. [40].

In common understanding, empathy is understanding another's emotions through perspective taking while sympathy is intentionally reacting emotionally especially involving sorrow and concern [41]. From the research perspective, there are there are three components and several key attributes to help understand empathy. Rogers described empathy's three components as affective (sensitivity), cognitive (observation and mental processing), and communicative (help's response) [42]. Wiseman has summarized four defining attributes in the concept analysis of empathy [43]:

- See the world as others see it
- Non-Judgemental

- Understanding another's feelings
- Communicate the understanding

In this research, both affective and cognitive sides will be taken into consideration and I mainly worked on interactions helping people present and understand emotional feelings at this stage.

2.2.2 Neural Substrate and Mechanism of Empathy

Since autonomic nervous system (ANS) is fundamental to affective experience, emotional expression, facial gestures, vocal communication and contingent social behavior [44], it is also closely related with empathy experience.The ANS is the part of the peripheral nervous system that controls involuntary functions that are critical for survival. The ANS and the bidirectional neural pathways communicating between it and brain influences the range of emotional expression and affect awareness which are important components in humans' empathy feelings [45, 46]. Besides,the mechanism of mirror neuron is also one of the most pivotal neural substrate of empathy. Mirror neurons are thought to represent goal-directed actions, serving the fundamental role of action understanding, imitation, intention understanding, and empathy [47, 48].

In line with mirror neurons' theory , Perception-Action Model (PAM) has been used by researchers to explain the basic framework and mechanism of empathy. The hypothesis of PAM is perception of a behavior in another automatically activates one ' s own representations for the behavior thus leading to an empathic response. And this model activates not only during prompt interaction but also imagination of a movement and emotional feelings [2, 11]. Referring to the PAM, an empathic interaction mainly involves two roles, one is the empathizer expressing the empathy and the other is empathizee receiving the empathy expressed. The two roles have different behaviors during a dyadic interaction and the empathic interaction can be activated and then continue in a cycle. Firstly, it start with empathizee's self-expression. Then empathizer should recognize the other's emotions with perspective-sharing and without judgemental thinking. After that, the empathizer should express the empathy to empathizee followed by empathizee's self-reflection to support the empathizer back as an understanding of themselves leads to enhancements in the ability to empathize with others [49, 50].

2.3. Affective Computing

2.3.1 Emotion Recognition

There are some attempts that researchers have made to improve the interaction between human and computer by helping computer to understand human's emotional context better. These attempts are based on different data sources. Shugrina presented an interactive emotional estimating and visualizing algorithms [51]. This approach is based on recognizing users' facial expressions through the detection of facial action units. Users' emotional ambiance can be estimated and later be mapped to the digital canvas after the rendering algorithm was performed.

Kirsch analyzed the results of three experiments and conducted a new experiment and developed the Sentic Mouse inspired by the results of emotional prediction theories [52]. The Sentic Mouse was an ordinary mouse augmented with a sensor to collect sentic data and proved to be able to capture valence information. Besides using mouse as the source of biometric characteristics, keyboard is also considered having the potential to help recognizing users' emotions as an input device. Clayton Epp conducted a research on determining user emotion by analyzing the rhythm of their typing patterns [53] and used the supervised machine learning algorithms for the classification. And the results proved that the keystroke dynamics is able to classify at least two levels of seven emotional states.

Kedar has reviewed the processing steps in his paper which were Image Preprocessing, Feature Extraction and Classification [54]. Handwriting characteristics,such as baseline, slant, pen-pressure, were also mentioned as analyzing features for the recognition of emotional outlays and other personality traits. Janet Fisher focused on the ability of automatic handwriting analysis in identifying traits in violent behavior [55]. Multiple samples from incarcerated violent offenders were compared with non-violent offenders. The results were found to be helpful in violent behavior prediction.

Michael Fairhurst worked on the enhancement of handwriting's forensic value to predict users' specific emotional state, which were the "happy" and the "stressed" [56]. For data processing, classification is performed by using KNN, Jrip and SVM classifier. SVM classifier performed best for the prediction in both cases with accuracy close to 80% in the "happy" emotion state and 70% in the "stressed" emotion state.

2.3.2 Computational Empathy

Reeves and Nass found that people tend to relate to computers in a social manner and treat them as real people [1]. As a consequence, empathy can happen not only between human and human but also between human and computers.Computational empathy means making computer empathize with humans and has now become an important field in affective computing. Like interactions between human and human, computers can be either empathizer as observers or empathizee as targets.

Developing computers as empathizers usually involves recognizing emotions and simulate empathic processes. There are many attempts in builiding empathic virtual robots as empathizers. Several researchers have developed empathy models through the implementation of specific empathic behaviors or empirical models and indicated that virtual robots with more empathy will be seen as being more caring, likeable,trustworthy, intelligent, dominant and less submissive [6, 7]. Besides scenarios in virtual world, the importance of increasing computers' understandings of emotions as a way to increase computer's empathy with human to enhance user experience is also recognized by the HCI community [3–5].

There has been research toward developing computers as empathizees in this field as well with the aim of evoking human's empathic responses to bring behavior change.Marsella etal. have developed an agent-based interactive pedagogical drama to improve problem solving skills of mothers of pediatric cancer patients. They adopted a presentational approach which is supposed to enhance users' empathy with characters but do not have to assume the burdens on themselves [8]. Paiva etal. have also developed a system called FearNot! to solve bullying problems in school by letting children from 8 to 12 witness bullying scenes played by empathic synthetic characters [9].

Since computers can act as both empathizer and empathizee during human computer interactions, it is also possible that computers can play as a mediator to evoke empathy between human and human.

Chapter 3
Concept and Hypothesis

3.1. Concept of the System

In the process of empathy, there are several essential attributes when defining empathy according to the summary of Wiseman [43], which are "see the world as others see it", "non-judgemental", "understanding another's feelings", and "communicate the understanding". He also mentions the possibility to be empathic and acknowledge the feelings of the past. Moreover, in the communication, empathic interactions involve both the empathizer and the empathizee [50, 57] when considering the empathy framework. For the empathizee, self-expression and self-reflection are necessary. While for the empathizer, perspective sharing, non-judgemental , recognizing emotion, and communicating understanding are essential.

Therefore, the system that I am going to build is trying to improve both the experiences of the empathizee and the empathizer, which helps empathizee record and express their emotional feelings based on the automatic emotion recognition and provide empathizer with more emotional hints behind written words with visualization. That is, the system is capable of recognizing writers' emotional situations from handwriting features instead of the contents of texts. Based on the recognition and prediction, the system is supposed to add emotional components back with visualization. In this way, the system can help people record and later interpret the real emotional meaning behind the words, which could enhance empathic understanding and interaction even during written communication.

As a result, the system must have two stages, one is emotion recognition as an input stage and the other is emotion presentation as an output stage. During my master project, I conducted two experiments respectively to test the feasibility of this system accordingly.

3.2. Emotion Recognition

3.2.1 Concept

Recently, people can type and write on the keyboard besides handwriting. However, handwriting still has its place when people are writing for authentication, creation and personal emotion-related communication such as letter and diary. The emotion recognition of my system is based on features related to movements of the hand instead of actual shapes of handwriting which is quite different from previous works. There are several reasons.

Firstly, the system to be developed aims to explore emotional feelings behind the words' meaning instead of what can be shown on the paper. In fact, even meanings of same words can change due to the different voices, tones and gestures during face-to-face communication. However, lack of those non-verbal information, it is hard to infer the real meaning and emotional feeling behind the words in written communication. Therefore, I presume data about hand movements during writing process can provide more prompt and relevant information to reveal emotions behind the words. Additionally,as a fine motor task highly dependant on procedural memory rather than deliberate conscious control of every hand muscle movement, handwriting process has low degree of control which would make it possible for unconscious manifestation of certain emotion-related features. As a result, data about how hand moves without the conscious control is qualified to prove the emotion recognition. Also, today's technology enables collecting such data in real time in an unobtrusive way, which makes the system more workable and accepted in future applications.

3.2.2 Research Question

I assume that the experienced emotions can manifest through the motor tasks, such as handwriting in our case. The research question for the first study is whether there is a correlation between these manifestations and the emotional condition of a given subject. Furthermore, if the correlation exists can it be generalized to a wider population or is it user-dependant.

3.3. Emotion Presentation

3.3.1 Concept

In previous work, presentation of another person's biosignals such as heart rate can help facilitate empathy and social awareness since biosignals naturally change with fluctuations in cognitive and emotional states [58, 59]. That might because of the aware and accuracy of the empathizee's emotional situation. Therefore, it can be assumed that other expressive ways of presenting emotional states which detected via handwriting features may also be effective to increase empathy in written communication. So, I conducted my second experiment to test whether the system can change empathizer's emotional perception after exposure to the original texts by adding certain emotion representing content to it, such as emotion-related word, color, shape and graph. Since empathy is an innate skill which is different from person to person, the empathy level was also taken into consideration.

3.3.2 Research Question

There are two conditions when emotion representing content is attached to the texts. One is the content represents similar emotional meanings with the original text. The other is the content represents opposite emotional meaning with the original texts. The research question for the second study is whether the presence of emotional state information is able to change empathizer's emotional perception in both the two conditions. If so, which types of Emotion Presentation is more effective. And whether there is a difference between people with different empathy levels.

Chapter 4
Experiment Design

4.1. Emotion Recognition

4.1.1 Method Overview

In order to induce certain emotions a series of emotionally-charged video clips was shown to participants. Due to short length of the clips, film clips were grouped into sets of 4 clips for each possible combination of Valence and Arousal: High Valence with High Arousal (HV-HA), High Valence with Low Arousal (HV-LA), Low Valence with High Arousal (LV-HA), and Low Valence with Low Arousal (LV-LA). The length of each set is 160 seconds. After watching each series of clips participants were asked to fill in a self-assessment form and to perform a series of writing tasks. Dominance was not included in the test because the high dominance videos were overlapping with high arousal and/or valence. Another reason for leaving out dominance is the test session length, adding high and low dominance to the existing 4 sets would require 8 testing sessions that would result in increased fatigue of the participants, which could possibly affect the data.

4.1.2 Materials

This section describes the materials used as emotion stimulus, emotion measurement and parts of the handwriting tasks.

The Emotional Movie Database

Film clips from the Emotional Movie Database (EMDB) [60] was used as emotion stimulus to elicit emotional states of the participants. EMDB is a database of 52 non-auditory film clips with different ratings of valence, arousal and dominance based on the self-assessment manikin (SAM) [27]. The duration of each film clip is 40 seconds. In this study, 16 film clips are selected and classified into 4 groups based on the arousal levels and valence levels due to their rating scores provided. Since the SAM scores for the film clips are split by gender, we prepared separate sets for male and female participants.

The Self-Assessment Manikin

The Self-Assessment Manikin (SAM) [23] was used to measure emotion states of participants after watching each film clip. SAM is a non-verbal pictorial self-report assessment associate with the reporter's affective reaction to a stimuli. It was designed to measure three aspects- pleasure, arousal and dominance.

The Affective Norms for English Words

The Affective Norms for English Words(ANEW) [61] was used as written materials to refer in the handwriting task. ANEW is a database of approximately 600 hundreds English words with different ratings of valence, arousal and dominance based on the self-assessment manikin(SAM). The most neutral words were selected separately for male and female participants based on the SAM scores.

4.1.3 Experiment

Participants

Thirteen volunteers aged 19 to 40 years (mean:24, SD:5 ;female: 8, male: 5) were recruited for this study. Every participant received a 1000 JPY gift-card as a compensation for their time, since the whole experiment takes each participant 30 to 40 minutes.

Procedure

The structure of the study was the following.

1. Written informed consent.
2. Baseline SAM test and recording session without film clips.

3. Four recording sessions, one session for each film clip category (HV-HA, HV-LA, LV-HA, LV-LA). Order of the sessions is counterbalanced to avoid ordering effects.

Each session, except the baseline recording that did not include film clips, is structured as follows. The orders of writing tasks in each session remains the same for all the participants.

1. 4 EMDB film clips, 160 seconds in total. Order of the clips is randomized each time.

2. SAM test for perceived Valence, Arousal and Dominance levels.

3. Task 1. Rewrite the words from the screen with time limit of 30 seconds. After the time limit the words disappear. The test is intended to force the participants to write fast which would increase the reliance on the procedural memory and weaken the conscious control over the fine motor performance. This task is aimed to gather normal handwriting data.

4. Task 2. Free doodling while watching the same video clips as in P.1 of this list. Order of clips is random. During this task we expect to record data related to drawing.

5. Task 3. Write a line of the following characters: o, /, —, -, :, +. the test is added to record very basic and repetitive strokes which could simplify the analysis of the stroke features.

The total duration of the test with each participant was between 30 and 45 minutes. The experiments were conducted in a sound-proof audio recording studio, with TV screen covering most of the participant's field of view. The only objects on the desk in front of the participant were SAM test sheets, tablet with a stylus and a charging cable(See Figure.4.1).

Figure 4.1 Participant during the experiment

4.2. Emotion Presentation

4.2.1 Method Overview

The experiment was conducted to evaluate the feasibility of using Emotion Presentation to affect readers' emotion understanding. The SAM scores of Original Texts were referred to the Affective Norms for English Text (ANET) [62]. These texts and SAM scores were used as a baseline. Emotion Presentation is the expressive visualization of emotions. In this experiment, 4 types of emotion stimuli were used as Emotion Presentation. The SAM scores of those emotion stimuli were also known in advance. Self-report measures reader's emotion perception via SAM as well. The later analysis based on calculating and comparing those three groups of SAM scores. The general structure of this experiment was as following.

Participants were asked to read one text from the ANET modified to first person tone. Each text was presented together with an additional Emotion Presentation containing emotional meaning. There were four types of Emotion Presentation which were Color, Word, Shape, and Graph in Valence-Arousal coordinate space.

Modifications to ANET and the Emotion Presentation used will be discussed below in details.

Since each of the ANET Text has it's valence and arousal scores already, the experiment's goal was to controllably change the self reported scores with Emotion Presentation in conjunction with the text. Valence and arousal scores for each Emotion Presentation were also known beforehand. This setup is mean to present both, the Original Text with it's semantic meanings and Emotion Presentation about the emotional state of the author on the same page. So far, Emotion Presentation was matched to the Original Text randomly instead of presenting the real emotional state of the author. In this way, more conditions were supposed to be tested even when the verbal meaning of the text and the real emotional meaning of the author were conflicted.

There are 3 sessions in all, and for each session there are 120 sentences respectively. After being exposed to the set of reading materials, participants were asked to fill the Self-Assessment Manikin (SAM) to assess their emotion understanding. With the SAM scores of Original Text, Emotion Presentation and Self report, the effect of the Emotion Presentation can be evaluated by calculation and comparison. Thus, whether the intervention by adding Emotion Presentation is useful in leading people have more insight into emotional meanings behind the words can be tested.

4.2.2 Materials

This section describes the materials used as Emotion Presentation, emotion measurement, empathy measure and reading tasks.

The Empathy Quotient

The Empathy Quotient is a self-report questionnaire to measure empathy developed by Baron-Cohen and Wheelwright. It contains 40 empathy items and 20 filler/control items. On each empathy item a person can score 2, 1, or 0 and the total score ranges from 80 to 0 [63]. Later the short forms of the Empathy Quotient (EQ-Short) was developed with 22 essential items sufficient to measure empathizing due to principal component analyses were selected from the Empathy

Quotient. [64]. The EQ-Short (See.Appendix.A) was used in this experiment to group participants into high empathy level and low empathy level for evaluating the difference effectiveness of Emotion Presentation.

The Self-Assessment Manikin

The Self-Assessment Manikin (SAM) [23] was used to measure emotion states of participants after watching each film clip. SAM is a non-verbal pictorial self-report assessment associate with the reporter's affective reaction to a stimuli. It was designed to measure three aspects- pleasure, arousal and dominance. Similar to the first experiment, only SAM scores of Valence and Arousal will be used for later analysis.

I-version of the Affective norms for English Text

The Affective norms for English Text (ANET) [62] is a set of normative emotional ratings for a large number of brief texts in the English language. Each sentence in this database is rated in terms of pleasure, arousal, and dominance with SAM scores. In the original ANET, all of the sentences used second person (e.g. You lie lazily in the hammock as a gentle summer breeze rocks you.) to refer the subject is the participant who was asked to rate these sentences with the SAM scores.

This study aimed to test how people will empathise events written in text by someone else, therefore subject was changed into first person in each sentence.Since while people are reading the sentences in second person tone, they usually tend to substitute themselves into the situation. So there should not lead to too much difference in SAM scores.

Also,the SAM scores in first person tone were collected both from experiments and online survey strictly following the instruction manual of ANET. And later the SAM scores were compared with those in second person tone. According to the instruction manual of ANET [62], there are two separate sets of 60 texts which are Set A and Set B. Texts in Set B are slightly longer in length than those in Set A. Thus, the online survey also consisted of two Sets. Both the Valence (r= 0.81,p¡.001) and Arousal(r= 0.60,p¡.001) scores in first person tone version have been proved to have significant correlation with those of the original database which means the SAM scores from ANET can still be used as a general proof for later analysis (See Table.4.1).

Text Number	Text(Second Person Tone)	Text(First Person Tone)	Pleasure	Arousal
1300 (A)	The dog strains forward, snarling, and suddenly leaps out at you.	The dog strains forward, snarling, and suddenly leaps out at me.	2.86 (1.85)	7.71 (1.77)
1460 (A)	Your new kitten nestles comfortably in your lap as you stroke her fur.	My new kitten nestles comfortably in my lap as I stroke her fur.	8.06 (1.24)	2.81 (2.14)
1500 (A)	The snake darts forward, jaws open, and sinks its fangs into your leg.	The snake darts forward, jaws open, and sinks its fangs into my leg.	1.38 (0.89)	8.50 (1.26)
1350 (B)	You jump back, muscles tense, as the large dog strains against the bared, leaping and snarling in a crazy rage.	I jump back, muscles tense, as the large dog strains against the bared, leaping and snarling in a crazy rage.	2.88 (1.46)	7.70 (1.77)
1940 (B)	You awake, heart pounding, startled in the darkness. A snake has crawled into your sleeping bag. It slides along your leg. You struggle frantically trying to get out.	I awake, heart pounding, startled in the darkness. A snake has crawled into my sleeping bag. It slides along my leg. I struggle frantically trying to get out.	1.13 (0.35)	8.88 (0.35)
1950 (B)	You watch a giant snake coiled in a display case. You freeze, as the snake's eyes move in your direction, and a red forked-tongue darts out.	I watch a giant snake coiled in a display case. I freeze, as the snake's eyes move in my direction, and a red forked-tongue darts out.	2.00 (1.07)	8.50 (0.76)

Table 4.1 Examples of Set A and B

the Affective Norms of English Words

The Affective Norms for English Words(ANEW) [61] is a database of approximately 600 hundreds English words with different ratings of Valence, Arousal and Dominance based on the Self-Assessment Manikin(SAM). Mainly Valence score and Arousal score were referred during this experiment. Word Emotion Presentation were selected based on 8 primary emotions and their SAM scores were defined referring to the ANEW database (See Table.4.2). Besides, the ANEW database was used as a dictionary which can connect other types of Emotion Presentation with their corresponding SAM scores. For example, the Color Emotion Presentations' SAM scores can be searched in the ANEW database basing on their corresponding verbal descriptions; the Shape Emotion Presentation were generated after emotion-related word were inputted to the system.

The theory of Plutchik's Wheel of Emotions

Color's corresponding emotion related word	Valence Score	Arousal Score
Joy	8.6 (0.71)	7.22 (2.13)
Surprise	7.47 (1.56)	7.47 (2.09)
Anticipation	7.06 (1.96)	5.44 (2.47)
Trust	6.68 (2.71)	5.3 (2.66)
Fear	2.76 (2.12)	6.96 (2.17)
Disgust	2.45 (1.41)	5.42 (2.59)
Anger	2.34 (1.32)	7.63 (1.91)
Sadness	1.61 (0.95)	4.13 (2.38)

Table 4.2 SAM scores of 8 Color Emotion Presentations

The theory of Plutchik's wheel of emotions [33] (See Figure.4.2shows that there are several important characters when we define emotions, which are primary, opposites, combination and intensity. This description are analogous to the colors on a color wheel indicating emotions' color meanings. In this experiment, 8 primary emotions and their corresponding colors were used as one type of Emotion Presentation. By searching the 8 emotion related words in the Affective Norms of English Words (ANEW), the corresponding colors' SAM scores can be indicated and mapped on Valence-Arousal dimensions(See Figure. 4.3.) The specific SAM scores of Color Emotion Presentation were listed in Table.4.2.

The EmotiveModeler

The EmotiveModeler is a CAD tool that can visualize emotion-related words into shapes in Rhinoceros3D modeling softeware [38, 65, 66]. To generate the Shape Emotion Presentation materials, the emotion related words should be inputted

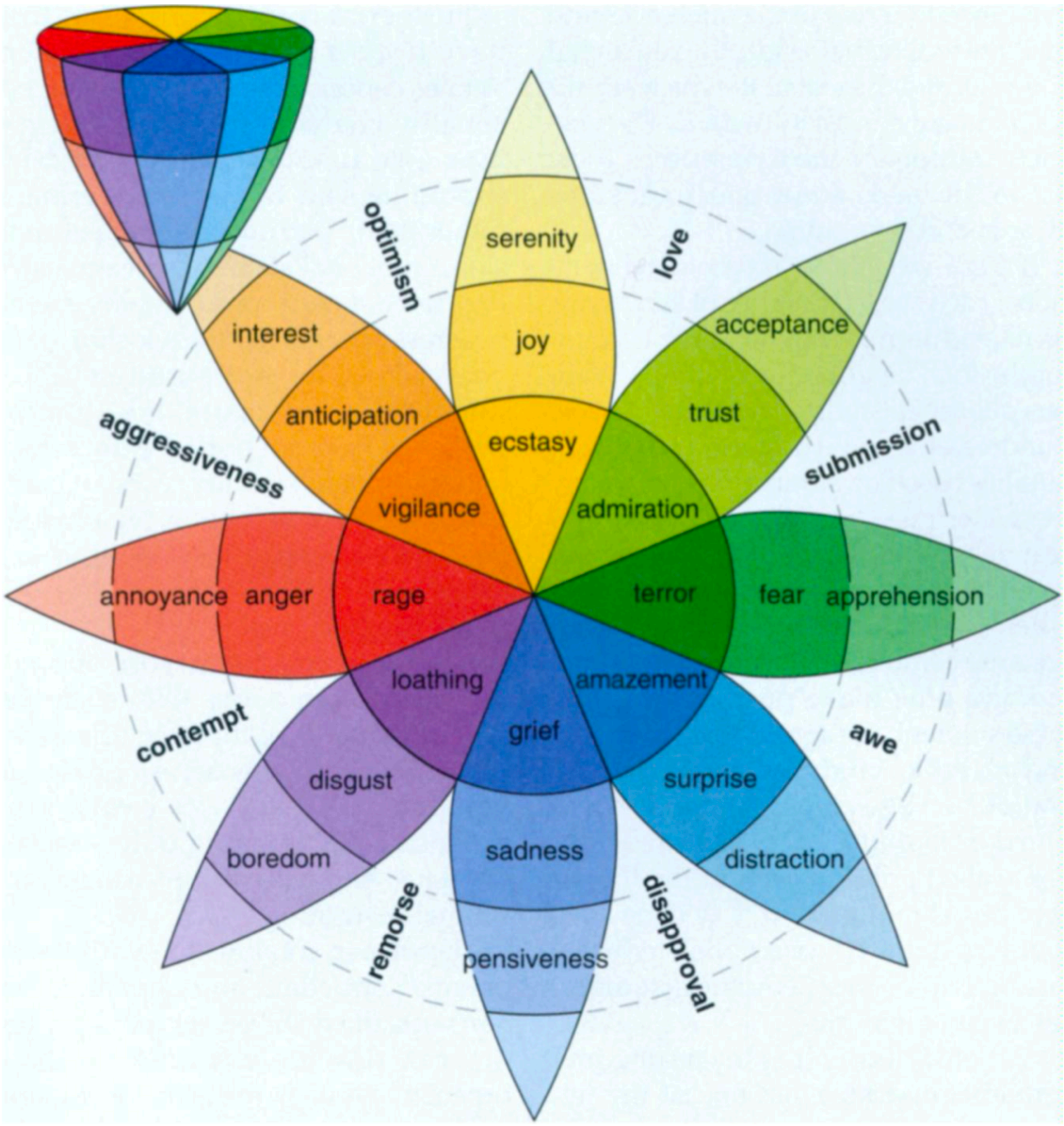

Figure 4.2 The Plutchik's circumplex model describes the relations among emotion concepts, which are analogous to the colors on a color wheel. The eight sectors are designed to indicate that there are eight primary emotion dimensions.

to the system. The emotion related words were selected referring to the SAM scores (mainly Valence and Arousal scores) in the ANEW database. The selection followed the standard that words should cover four emotion groups: High Valence

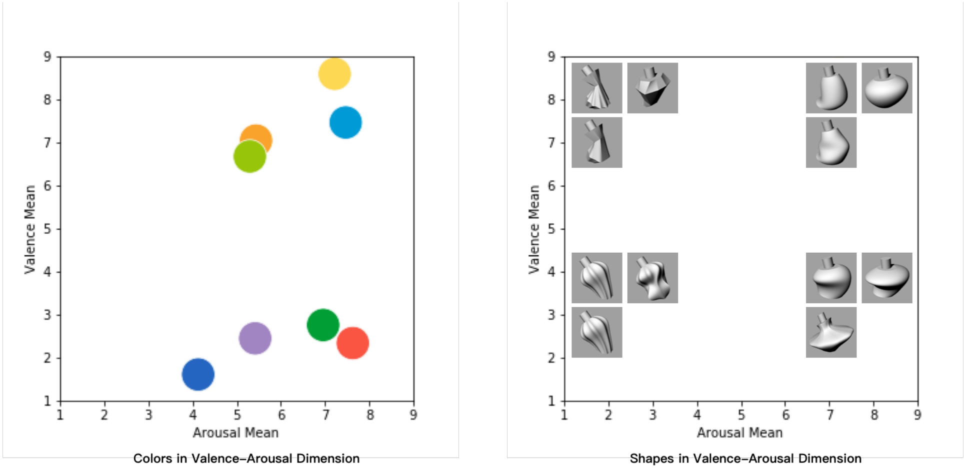

Figure 4.3 Color Emotion Presentation and Shapes Emotion Presentation mapped in Valence-Arousal coordinate space. Left: Color Emotion Presentation. Right: Shape Emotion Presentation.

with High Arousal (HV-HA), High Valence with Low Arousal (HV-LA), Low Valence with High Arousal (LV-HA), and Low Valence with Low Arousal (LV-LA). The Shapes generated were mapped on Valence-Arousal dimensions (See Figure. 4.3). The specific SAM scores of Shape Emotion Presentation were listed in Table.4.3.

4.2.3 Experiment

Participants

Twenty volunteers aged 18 to 44 years (female: 10, male: 10) were recruited for this study. Every participant received an 1000 JPY Amazon gift-card as a compensation for their time, since the whole experiment takes each participant 40 to 60 minutes.

Procedure

The structure of the study was the following.

1. Written informed consent.

Shape's corresponding emotion related word	**Emotion Group**	**Valence Score**	**Arousal Score**
Excitement	HV-HA	7.5 (2.20)	7.67 (1.91)
Miracle	HV-HA	8.6 (0.71)	7.65 (1.67)
Orgasm	HV-HA	8.32 (1.31)	8.1 (1.45)
Comfort	HV-LA	7.07 (2.14)	3.93 (2.85)
Gentle	HV-LA	7.31 (1.30)	3.21 (2.57)
Peace	HV-LA	7.72 (1.75)	2.95 (2.55)
Anger	LV-HA	2.34 (1.32)	7.63 (1.91)
Assault	LV-HA	2.03 (1.55)	7.51 (2.28)
Nightmare	LV-HA	1.91 (1.54)	7.59 (2.23)
Fatigued	LV-LA	3.28 (1.43)	2.64 (2.19)
Gloom	LV-LA	1.88 (1.23)	3.83 (2.33)
Obesity	LV-LA	2.73 (1.85)	3.87 (2.82)

Table 4.3 SAM scores of Shape Emotion Presentations

2. The Empathy Quotient survey [63], a 22-item version of Empathy Quotient (EQ-Short) [64].

3. Reading and Self-assessment.

There are approximately 120 sentences randomly shown on the tablet screen without repetition. For each of them, one of 5 different designs is be presented below the page.

1. Blank without any feature

2. Color with specific emotional meaning according to SAM scores

3. Shape with specific emotional meaning according to SAM scores

4. Word with specific emotional meaning according to SAM scores

5. Graph with specific emotional meaning in Valence-Arousal coordinate space

The total duration of the test with each participant was between 40 and 60 minutes. The experiments were conducted in a sound-proof audio recording studio, the experiment software was running on a 12.3 inch Microsoft Surface Pro (See the interface in Figure.4.4).

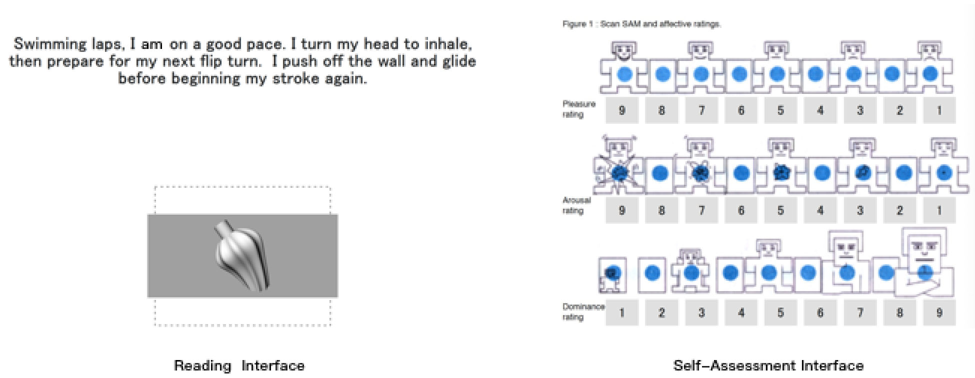

Figure 4.4 Test software interface. Left: An Original Text with an additional Emotion Presentation (Shape); Right: The SAM survey interface for self-assessment

Chapter 5
Evaluation

5.1. Emotion Recognition

5.1.1 Feature Extraction

In total throughout 5 session with 13 participants, 1442789 samples of the pen position were recorded. And later the samples were split into 14282 strokes based on the touch phase data provided by the tablet. Recordings for a few test sessions were lost or discarded due to technical or network reasons. For further analysis only the sessions reliably recorded from start to finish were used.

For each stroke was described by the following parameters, 44 in total. 16 most usable features were selected based on Gini importance(See Table 5.1).

- Stroke start and end time from experimenter's PC and tablet.
- Stroke length in pixels (total and on X and Y axis separately) and in millimeters.
- Stroke duration.
- Stroke SD of the distances between each sample in the stroke.
- Number of samples in the stroke.
- Average force and its SD.
- Average altitude and azimuth angles, their SD and total angular path.
- Time from the last stroke.
- Features extracted by splitting the stroke into 25ms windows with 50% overlap:

Table 5.1 16 stroke features selected as most suitable for classification

Name	**Description**
ALT_ANGL	Altitude(in radians) of the stylus
AZ_ANGL	Azimuth(in radians) of the stylus
PATH_AZ_ANGL	Total angular path of azimuth angle
SPEED_AL_MIN	Minimal speed of the altitude angle (from 25ms windows)
SPEED_AZ_MIN	minimal speed of the azimuth angle (from 25ms windows)
SPEED_MAX	Maximum of stroke speed
SPEED_STD	Standard deviation of stroke speed
SPEED_X_MIN	Minimal speed on the X axis (from 25ms windows)
SPEED_Y_MIN	Minimal speed on the Y axis (from 25ms windows)
FORCE	Average force of the stylus touch
FORCE_SD	Standard deviation of stroke force
PATH_X	Total stroke path on the X axis
PATH_Y	Total stroke path on the Y axis
TIME	Duration of the stroke
TIME_FROM_LAST	Time from the last stroke
SAMPLES	Number of samples in the stroke

  - Top window speed on X and Y axis, on both axes combined, angular speed of azimuth and altitude angles and maximal force.
  - Lowest window speed on X and Y axis, on both axes combined, angular speed of azimuth and altitude angles and minimal force.
  - Maximal speed increase between neighboring windows on X and Y axis, on both axes combined, maximal angular speed increase of azimuth and altitude angles and maximal force increase.
  - Maximal speed decrease between neighboring windows on X and Y axis, on both axes combined, maximal angular speed decrease of azimuth and altitude angles and maximal force decrease.

Excluding stroke start and end time I analyzed the remaining 40 features in order to test whether there is any relation between the handwriting features and the self-reported emotional state of the participant recorded using the SAM. In order to do this I used the C-Support Vector Classification (SVC) from scikit-learn library [67] with valence, arousal and dominance ratings as labels.

In order to select the most promising features I have calculated the Gini Importance for each feature for each of the three tasks separately. Then selected and combined top 10 features from each of the tasks. The resulting list contains 16 features and is shown in the Table 5.1. The details on the SVC and its performance is discussed in the following section.

5.1.2 Classifier Training

To assert the effect of the chosen emotional stimuli, I calculated the correlation between the SAM scores reported in the EMDB and those self-reported. For valence scores the correlation was 0.905. For arousal scores the correlation was 0.959. This confirms that the effect of emotional stimulation was present and as expected.

The classifier I used is SVC from scikit-learn library. Before training, I split all the data into train data, and test data (train to test ratio is $9:1$). In order to make sure that the classifier does not overfit on the train data, I used a 5-fold cross-validation for the train data and calculated the average of the 5 accuracy scores as the training score of the classifier (Fig.5.2.A (CV)). When training the classifier, I tuned it for the best training score by shifting the hyperparameter C and gamma to optimize each classifier. I then test the classifier on the test data and calculate the test score. (Fig.5.2.A).

User Dependent

The user-dependent models were made using the type of the video content as the label (HV-HA, HV-LA, LV-HA, LV-LA). Due to data loss for participants 9 and 11 results for task 1 of participant 9 are irrelevant, and participant 11 was completely removed from the analysis.

The classifier scores for all strokes combined, short and long separately and cross validation scores are presented on Fig.5.2. The results represent data for all 12 participants individually and the last chart for all participants combined. For each participant results were presented split into 3 groups: for all strokes, for short and long strokes. In each group there are classifiers for tasks 1, 2, 3, and all tasks combined.

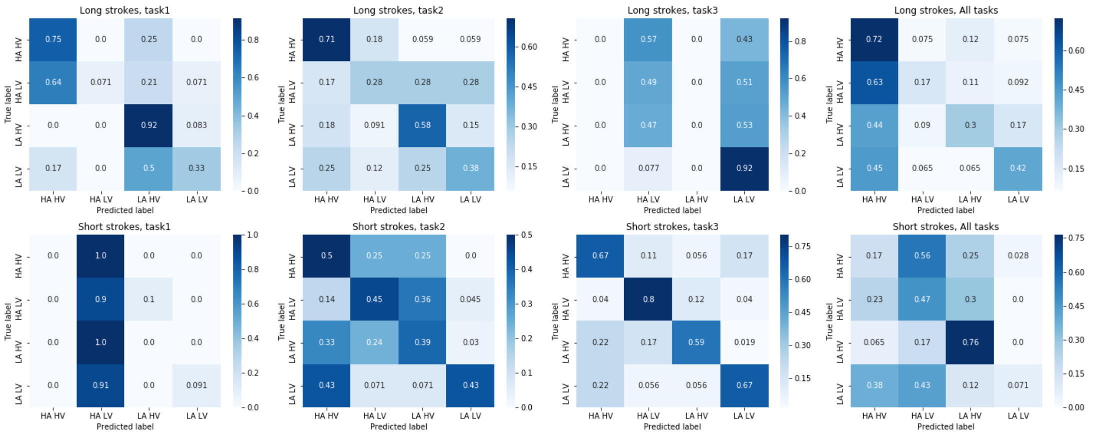

Figure 5.1 Confusion matrices for user-independent model. The stroke data was split into short (bottom) and long (top) strokes before classification.

For all strokes combined, the accuracy for 1 out of 4 classification averaged between all tasks for all 12 users is 58.4% (Score Task 1,2,3 on Fig.5.2A), 43% for the task-independent classifier (Score All Tasks on Fig.5.2A). Data for short and long strokes separately is presented on Fig.5.2B. For short strokes the accuracy averaged between all tasks for all 12 users is 62%, task-independent is 57%. For tasks 1, 2 and 3 it is 54%, 61% and 69% respectively. For long strokes the accuracy averaged between all tasks for all 12 users is 65%, task-independent is 50%. For tasks 1, 2 and 3 it is 64%, 60% and 70% respectively. The accuracy of user independent classifier is 28%, 57% and 47% for the tasks 1, 2, and 3 respectively. Accuracy of task-independent user-independent classifier was 35%.

User Independent

In order to build the classifier the valence and arousal scores of the SAM were used as labels. The scores were grouped into "high" (6-9 points) and "low" (1-4 points) categories. The data scored at 5 points of valence or arousal was omitted. This assures that the actual arousal and valence levels of the participant correspond to the handwriting, since not all the participants had intended emotional response to the video stimuli.

The confusion matrices of each condition are presented on Fig.5.1, and the score results of this classification are presented on Fig.5.3. I ran classification on the

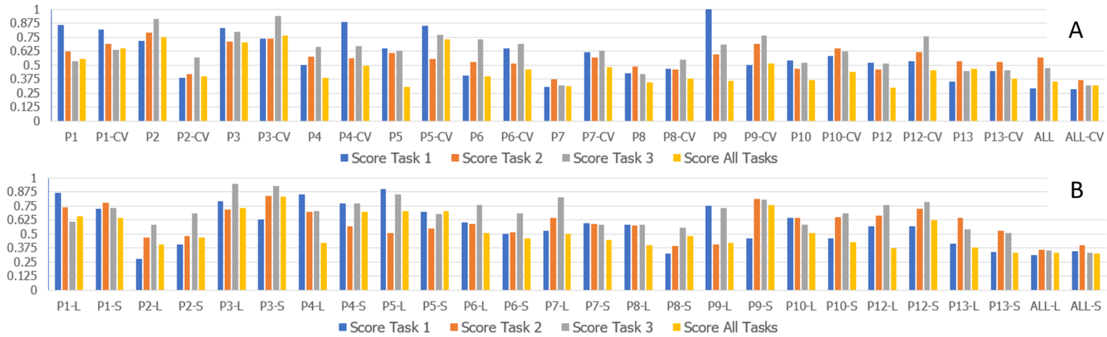

Figure 5.2 Classification scores for user-dependent model. A. Test scores sing all strokes, CV - cross validation scores for each participant's model. B. Models using only short (S) and long (L) strokes. "Score All tasks" is the task-independent classifier, "ALL" represents data for user-independent classifier.

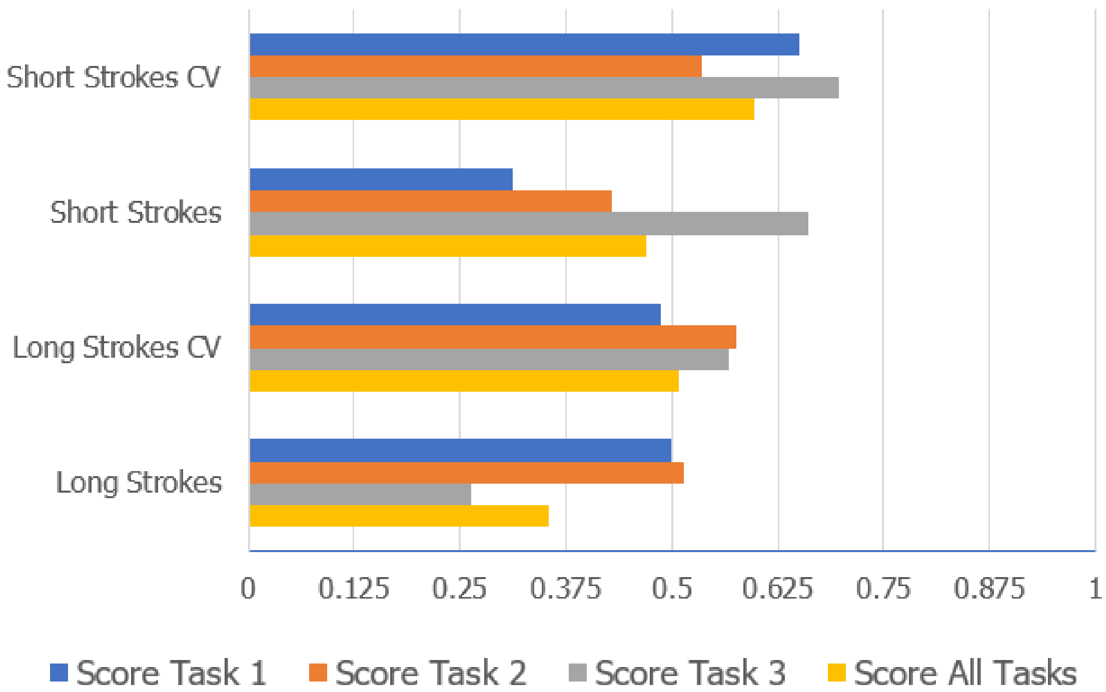

Figure 5.3 Classification scores for user-independent model. CV - cross validation results. The stroke data was split into short and long strokes before classification.

data sets from each task and on all three tasks combined. It was found that the classification precision changes greatly if only short or only long strokes were used. The stroke data was split into short and long strokes in relation to the median for each data set used. Short strokes gave particularly good results for task 3, as it required participants to draw mostly short lines. Accuracy for this test is reaching 66% for 1 in 4 groups classification. Using long strokes showed better results for tasks 1 and 2 with accuracy of 50 and 51% respectively. Surprisingly, the accuracy of classification of the short stroke data set for all 3 tasks was higher than for long stroke data set (47% for short and 35% for long).

On the downside, although this approach provides a reliable labeling,all the data have to be excluded if theydid not fit into one of the four labeled groups: 1. High Arousal and Valence. 2. High Arousal low Valence. 3. Low Arousal and high Valence. 4. Low Arousal and Valence. After applying this labeling to the stroke data, the data for some of the participants was not sufficient to build a reliable classifier. Thus this approach was used only for user-independent model that was trained on the data from all the users.

These results were organized into a full paper which was published in the 10th Augmented Human International Conference 2019 [68].

5.2. Emotion Presentation

I have evaluated the effectiveness of the 4 types of Emotion Presentation by calculating and comparing the SAM scores of orginal texts, emotional stimuli added to the text, and the user score report. And I used the same mean values and standard variations of valence and arousal scores of texts as the ANET [62] database rated. These scores are used as the baselines for analyzing the effect of the added stimuli. I define the stimulus modification direction by comparing valence and arousal scores of the added emotional stimuli and the original text's. If the added emotional stimuli have a higher score than the original text, the expected effect on the perception would be positive or increased. If supplementary emotional stimuli have a lower score than the original text, then it is expected to decrease the score on a given axis.

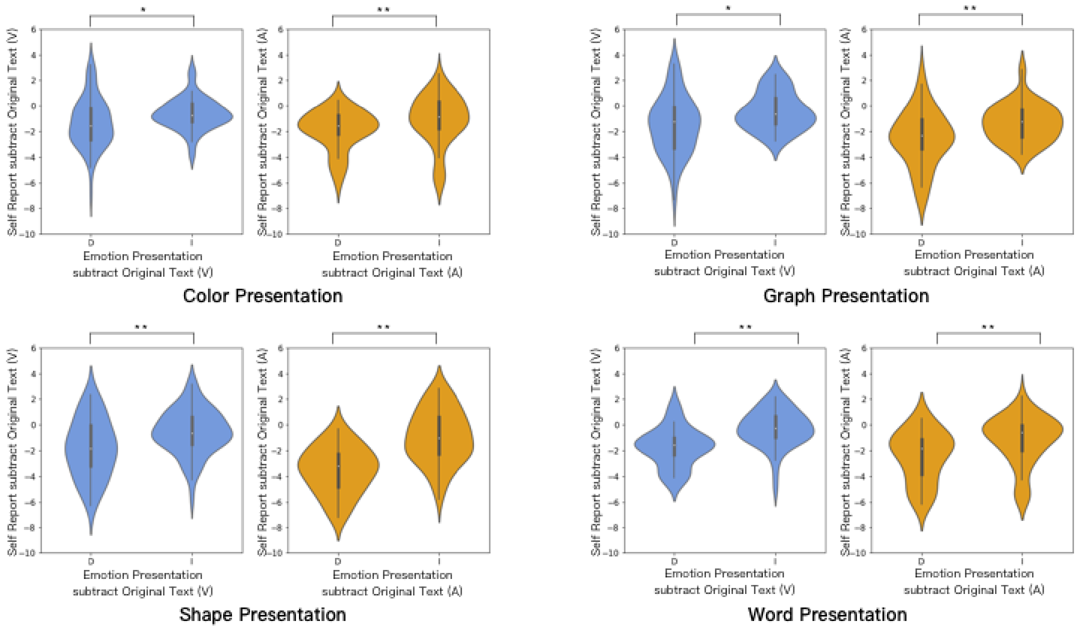

Figure 5.4 Different effects on Self-report when applying Emotion Presentation with different modification directions (Down/Up)

5.2.1 Emotion Modification

emotion stimuli Type t value (p value)	**Emotion Dimension**	**Modification Direction**	**Emotion Perception Change**
Color	Valence	Down	-1.78 (1.91)
t= 2.30 (p¡.05)	Valence	Up	-0.93 (1.19)
Color	Arousal	Down	-3.44(2.06)
t= 3.89 (p¡.001)	Arousal	Up	-1.34 (2.29)
Graph	Valence	Down	-1.47 (2.26)
t= 2.55 (p¡.05)	Valence	Up	-0.42 (1.45)
Graph	Arousal	Down	-2.53(2.20)
t= 2.98 (p¡.01)	Arousal	Up	-1.32 (1.46)
Shape	Valence	Down	-1.78(1.91)
t= 2.94 (p¡.01)	Valence	Up	-1.81 (1.19)
Shape	Arousal	Down	-3.54(1.80)
t= 6.62 (p¡.001)	Arousal	Up	-0.91(2.07)
Word	Valence	Down	-2.00(1.52)
t= 4.06 (p¡.001)	Valence	Up	-0.52(1.49)
Word	Arousal	Down	-3.45 (1.97)
t= 5.91 (p¡.001)	Arousal	Up	-0.61 (2.06)

Table 5.2 Means of emotion perception's change in two emotion dimensions after applying emotion stimuli with different modification directions

I quantified the changes of emotion understanding by subtracting the SAM score of texts from that of self-report. Then I applied Welch's t-test on the changes of emotion understanding between two modification direction for each stimuli type. According to the results, there was significant difference in the emotion perception's change when the added stimulus was expected to increase or decrease the perceived emotional connotation of the text. In valence dimension, color (t(71) = 2.30,p¡.05), graph (t(72) = 2.55,p¡.05), shape(t(100)=2.94,p¡.01) and word(t(71) = 4.06, p ¡.001) were all shown to have substantial effects. In arousal dimension, color (t(71) = 3.89,p¡.001), graph (t(72) = 2.98,p¡.01), shape(t(100)=6.62,p¡.001) and word(t(71) = 5.91, p ¡.001) were all having substantial effect as well. Means across conditions are reported in Table.5.2.

Besides testing the effectiveness of 4 types of emotion stimuli, I also found interesting results by comparing the emotion modifications of people with different empathy levels. The empathy levels were divided into two groups referring to the mean score measured in the EQ-short test. The mean score of empathy is 12, so there are 11 subjects in the low empathy group while 9 subjects in the high empathy group. Data on subject number one and subject number fifteen was excluded due to the data loss. Therefore, there were 10 subjects in the low empathy group while 8 subjects in high empathy group. When I attempt to modify the arousal perceptions, people who have low empathy level (M = 66.70, SD = 10.86) compared to those with high empathy level (M = 60.13, SD = 11.31) the effects were significantly higher then anticipated , t(18) = 2.19, p ¡.05.This result may suggest that people who may struggle in developing accurate or prompt emotional understandings can be affected more easily or make more use out of the provided supplementary emotional clues. One possible explanation of this may be that people with high empathy levels can easily form a relatively thorough understanding of the emotional meanings even just by reading the texts themselves. Once they have their own impression of the emotional coloring, it might be harder to change their opinions.

5.2.2 Effects of Emotion Stimuli

I conducted an independent t-test to see whether the difference between SAM scores of self-report and SAM scores of original text is significant or not when various emotional stimuli are added. Firstly, I separated the original texts, emotion stimuli, and self-report into two groups respectively basing on the levels of valence or arousal.

- Original text: High/Low Arousal Level,High/Low Valence Level
- Emotional presentation: High/Low Arousal Level,High/Low Valence Level
- Self-report: High/Low Arousal Level,High/Low Valence Level

Color stimuli

In the condition of the original text with high valence level, there was significant difference between valence score of original text (M = 7.34, SD = 1.13) and that of self-report (M = 7.04, SD = 1.79) caused by adding a stimulus with high valence level (t(139) = 2.06, p¡ .05). Moreover, there was also significant difference between valence score of original text (M = 7.17, SD = 1.20) and that of self-report (M = 6.84, SD = 2.03) caused by stimuli with low valence level (t(142) = 2.26, p¡ .05). In the condition of original text with low valence levels, there was no significant difference found with neither high (t(85) = 0.32, p= 0.74) nor low valence level stimuli(t(101) = -0.15, p= 0.87) attached.

In the condition of original text with high arousal level, there was significant difference between arousal score of original text (M = 5.07, SD = 2.78) and that of self-report (M = 4.82, SD = 2.99) when stimuli with high arousal level was used (t(337) = 6.74, p¡ .005). However, there was no significant effect of stimuli with low arousal level (t(46) = 2.30, p= 0.87). In the condition of original text with low arousal level, there was no significant difference found with neither high (t(76) = 0.89, p= 0.17) nor low arousal level stimuli(t(10) = 0.38, p= 0.41) attached.

Graph Stimuli

In the condition of original having a high valence level, there was significant difference between score of original text (M = 7.35, SD = 1.18) and that of self-report (M = 6.91, SD = 2.05) with application of stimuli with high level (t(98) = 2.35, p¡ .05). Moreover, there was also significant difference between score of original text (M = 7.31, SD = 1.09) and that of self-report (M = 6.80, SD = 2.07) with low level additional stimuli (t(113) = 2.86, p¡ .05). In the condition of original text with low level, there was significant difference between score of original text (M = 2.22, SD = 0.54) and that of self-report (M = 2.65, SD = 1.81) triggered by stimuli with low level (t(78) = -2.12, p¡ .05). However,there was no significant difference found with high level stimuli attached(t(77) = -0.03, p= 0.97).

In the condition of original text with high arousal level, there was significant difference between arousal score of original text (M = 5.36, SD = 2.87) and that of self-report (M = 5.00, SD = 3.05) triggered by stimuli with low arousal level (t(159) = 4.10, p¡ .05). However, there was no significant effect of stimuli with high arousal level (t(139) = 4.19, p= 0.68). In the condition of original text with low arousal level, there was no significant difference found with neither high (t(37) = 1.26, p= 0.42) nor low arousal level stimuli(t(31) = 0.38, p= 0.41) attached.

Shape Stimuli

In the condition of original text with high level, there was significant difference between score of original text (M = 7.37, SD = 1.09) and that of self-report (M = 6.50, SD = 2.18) triggered by stimuli with high level (t(115) = 4.62, p¡ .001). Moreover, there was also significant difference between score of original text (M = 7.33, SD = 1.14) and that of self-report (M = 6.75, SD = 1.76) triggered by stimuli with low level (t(128) = 4.02, p¡ .001). In the condition of original text with low level, there was no significant difference found with neither high (t(108) = -0.43, p= 0.66) nor low level stimuli(t(94) = 0.83, p= 0.40) attached.

In the condition of original text with high arousal level, there was significant difference between arousal score of original text (M = 4.81, SD = 2.80) and that of self-report (M = 4.53, SD = 2.95) triggered by stimuli with high arousal level (t(180) = -0.14, p¡ .05). Moreover, there was also significant difference between arousal score of original text (M = 4.91, SD = 2.81) and that of self-report (M = 4.48, SD = 2.74) triggered by stimuli with low arousal level (t(195) = 5.92, p¡ .001). In the condition of original text with low arousal level, there was no significant difference found with neither high (t(37) = -0.68, p= 0.06) nor low arousal level stimuli(t(35) = -0.57, p= 0.54) attached.

Word Stimuli

In the condition of high level original text, there was significant difference between score of original text (M = 7.16, SD = 1.12) and that of self-report (M = 6.78, SD = 1.84) triggered by stimuli with high valence level (t(127) = 2.54, p¡ .05). Moreover, there was also significant difference between valence score of original text (M = 7.16, SD = 1.16) and that of self-report (M = 6.73, SD = 2.05) triggered by stimuli with low valence level (t(126) = 3.03, p¡ .01). In the condition of original text with low valence level, there was significant difference between valence score of original text (M = 2.36, SD = 0.66) and that of self-report (M = 2.68, SD = 1.37) triggered by stimuli with high valence level (t(82) = -2.26, p¡ .05).However,there was no significant difference found with low valence level stimuli attached(t(104) = 0.86, p= 0.41).

In the condition of original text with high arousal level, there was no significant difference found with neither high (t(310) = 5.83, p= 0.31) nor low arousal level stimuli(t(46) = 2.87, p= 0.16) attached. In the condition of original text with low arousal level, there was significant difference between arousal score of original text (M = 6.22, SD = 1.12) and that of self-report (M = 5.68, SD = 1.94) triggered by stimuli with high arousal level (t(77) = -1.92, p¡ .01).However,there was no significant difference found with low arousal level stimuli attached(t(7) = -0.73, p= 0.81).

From Table.**??**, only when Original Text's score level and Emotion Presentation's score level are different, does the significance exist. In other words, effective emotion triggering only happens when Original Text and Emotion Presentation have different levels of Valence or Arousal. Moreover, only Positive Effect Type occurs and mostly the pattern type is Decrease-Positive, which means the SAM score Self Report can be leaded lower when the SAM score of Emotion Presentation is lower than that of Original Text. This result enlightens the possible condition when people need to have empathic feeling with things related to sadness or calm down.

Furthermore, since empathy is considered as an innate skill, different performances of people with High or Low level of empathy have been explored. So the t-test was conducted in three groups of samples which are All subjects regardless of empathy level, subjects with high empathy level and subjects with low empathy level. According to Table.**??**, Emotion Presentation of Shape is effective to subjects with low empathy level only in triggering changes in Valence. Emotion Presentation of Graph is effective to subjects with low empathy level only in triggering changes both in Valence and Arousal.

5.3. Discussion

For emotion recognition, in both user-dependent and user-independent models length-based separation of the strokes clearly resulted in increased accuracy. The possible explanation for this may be that certain features manifest themselves differently depending on the type of the strokes and tasks. Meaning that check marks or crosses may convey information related to the emotional states different from long strokes used for drawing or doodling. Especially for task3 (simple and repetitive characters), it showed better performance on short strokes, as it required participants to write short lines. For different tasks, it may be necessary to train different models for the classification. Additionally, there are differences in performance of user-dependent and user-independent model suggesting that the differences between individual manifestations of the emotional states in fine motor performance are present and have to be considered for such applications.

For emotion presentation, the hypothesis that adding Emotion Presentation with different trigger directions (Down/Up) can lead to different change of emotion perception. According to these findings, when the original text has high valence level, color stimuli, graph stimuli, shape stimuli, and word stimuli can reduce the valence perception of the text to a lower level, regardless of stimulus's valence levels. When the original text has high arousal level, color stimuli in high arousal level, graph stimuli in low arousal level, and shape stimuli regardless of its arousal level have been proved to be effective in lowering the arousal perception. However, against the intuition, when a high arousal word stimuli presents with a low arousal original text can lead to a even lower arousal perception.

In all, only graph stimuli and word stimuli indicated promising effects on enhancing valence perception of low valence texts. And so far no supplementary emotion stimuli type was shown to increase arousal perception. However, I did prove the capability of using supplementary emotion stimuli to lower the perceived intensity of emotions in both valence and arousal scales, unless the original text already scores low on these scales. Our results show that system with emotion stimuli supplementation to the text that can influence the perception of emotional meanings of the text. Especially the effectiveness of alleviating the level of perceived arousal. Thus this may suggest the possible solutions for aggressive emotion regulations especially during the communication in online scenarios such as on Twitter or Facebook. In addition, with the various approaches to the emotion recognition in text communication, the emotional stimuli can be added to the text to represent more precisely the writers' emotional state, thus promote empathic understanding between writers and readers.

Chapter 6
Conclusion

Both the in-put stage (emotion recognition) and the out-put stage (emotion presentation) have been tested in two separate experiments in order to prove the feasibility of the system.

The in-put stage's experiment was conducted to collect and analyze emotionally-labeled handwriting data. Out of 44 extracted features, 16 most promising features have been selected according to their Gini Importance. Based on this feature set, classifiers was built using the type of video content as the label for the user-dependent model and the SAM scores as the label for the user-independent model. For user-dependent 4-class classifier, the classification accuracy is up to 70% for certain tasks. For user-independent 4-class classifier, classification accuracy is up to 66% for a specific handwriting task.

The out-put stage's experiment was conducted to find out the most effective presentation types among Color, Word, Shape, and Graph. And all the Emotion Presentation types have been proved to be effective in lowering down valence and arousal perception. This may suggest that it is possible to help people understand the emotions in lower emotion levels, especially arousal level. By lowering down arousal level's perception, readers can feel less offensive and calmer, which can promote emotion understanding with less judgemental thinking. Among the four types of Emotion Presentation, Graph has been proved both effective in Valence and Arousal dimensions. Since in the in-put stage (emotion recognition), users' emotional states will be recognized based on Russell's Circumplex Model and mapped into Valence-Arousal coordinate space, the presentation type of Graph will also be able to match with the emotion recognition correspondingly. The results also indicate people with low empathy levels can be affected more easily as anticipated. Compared with those already with adequate empathic understanding ability, it is people that with low empathy need more empathic communication's enhancement. And this result proves the potential to help with people with low empathy level by using this system.

In conclusion, even though the results are promising restricting to specific handwriting tasks and Emotion Presentation Types, they can still prove the feasibility and the potential of recognizing emotions and adding emotion presentation to help people understand real emotional feelings behind words in written communication. Since understand emotions properly plays a fundamental and essential part in an empathic communication, this improvement is supposed to lead empathy's enhancement.

Based on the results from two experiments, the system can be developed in the future. The classification algorithms will be polished to increase its prediction accuracy and other approaches to data analysis, such as deep learning will also be tested. In the first stage, the Emotion Presentation Type of Graph will be applied to the system to record and present the emotions recognized. With the prototype, user tests will be conducted based on specific scenarios such as artistic creation, personal letter communication, and working condition where handwriting is still needed.

References

[1] Byron Reeves and Clifford Ivar Nass. *The media equation: How people treat computers, television, and new media like real people and places.* Cambridge university press, 1996.

[2] Ana Paiva, Iolanda Leite, Hana Boukricha, and Ipke Wachsmuth. Empathy in virtual agents and robots: A survey. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 7(3):11, 2017.

[3] Effie Lai-Chong Law, Virpi Roto, Marc Hassenzahl, Arnold P.O.S. Vermeeren, and Joke Kort. Understanding, scoping and defining user experience: A survey approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 719–728, New York, NY, USA, 2009. ACM. URL: `http://doi.acm.org/10.1145/1518701.1518813`, `doi:10.1145/1518701.1518813`.

[4] Jodi Forlizzi and Katja Battarbee. Understanding experience in interactive systems. In *Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, DIS '04, pages 261–268, New York, NY, USA, 2004. ACM. URL: `http://doi.acm.org/10.1145/1013115.1013152`, `doi:10.1145/1013115.1013152`.

[5] Rosalind W. Picard. Affective computing for hci. In *Proceedings of HCI International (the 8th International Conference on Human-Computer Interaction) on Human-Computer Interaction: Ergonomics and User Interfaces-Volume I - Volume I*, pages 829–833, Hillsdale, NJ, USA, 1999. L. Erlbaum Associates Inc. URL: `http://dl.acm.org/citation.cfm?id=647943.742338`.

[6] Scott Brave, Clifford Nass, and Kevin Hutchinson. Computers that care: investigating the effects of orientation of emotion exhibited by an em-

bodied computer agent. *International journal of human-computer studies*, 62(2):161–178, 2005.

[7] Sérgio Hortas Rodrigues, Samuel Mascarenhas, João Dias, and Ana Paiva. A process model of empathy for virtual agents. *Interacting with Computers*, 27(4):371–391, 2015.

[8] Stacy C Marsella, W Lewis Johnson, and Catherine LaBore. Interactive pedagogical drama. In *International Conference on Autonomous Agents: Proceedings of the fourth international conference on Autonomous agents*, volume 3, pages 301–308, 2000.

[9] Ana Paiva, Joao Dias, Daniel Sobral, Ruth Aylett, Polly Sobreperez, Sarah Woods, Carsten Zoll, and Lynne Hall. Caring for agents and agents that care: Building empathic relations with synthetic agents. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 194–201. IEEE Computer Society, 2004.

[10] Paul S Bellet and Michael J Maloney. The importance of empathy as an interviewing skill in medicine. *Jama*, 266(13):1831–1832, 1991.

[11] Stephanie D Preston and Frans BM De Waal. Empathy: Its ultimate and proximate bases. *Behavioral and brain sciences*, 25(1):1–20, 2002.

[12] Hugo Lövheim. A new three-dimensional model for emotions and monoamine neurotransmitters. *Medical Hypotheses*, 78(2):341 – 348, 2012. URL: `http://www.sciencedirect.com/science/article/pii/S0306987711005883`, `doi:https://doi.org/10.1016/j.mehy.2011.11.016`.

[13] Christian Collet, Franck Di Rienzo, Nady Hoyek, and Aymeric Guillot. Autonomic nervous system correlates in movement observation and motor imagery. *Frontiers in Human Neuroscience*, 7:415, 2013. URL: `https://www.frontiersin.org/article/10.3389/fnhum.2013.00415`, `doi:10.3389/fnhum.2013.00415`.

[14] Deanna M Sasaki-Adams and Ann E Kelley. Serotonin-dopamine interactions in the control of conditioned reinforcement and motor behavior. *Neuropsychopharmacology*, 2001. `doi:10.1016/S0893-133X(01)00240-8`.

[15] Kiyoshi Ando, Chris E. Johanson, Lewis S. Seiden, and Charles R. Schuster. Sensitivity changes to dopaminergic agents in fine motor control of rhesus monkeys after repeated methamphetamine administration. *Pharmacology Biochemistry and Behavior*, 22(5):737 – 743, 1985. URL: http://www.sciencedirect.com/science/article/pii/0091305785905222, doi:https://doi.org/10.1016/0091-3057(85)90522-2.

[16] SV Kedar, DS Bormane, Aaditi Dhadwal, Shiwali Alone, and Rashi Agarwal. Automatic emotion recognition through handwriting analysis: a review. In *Computing Communication Control and Automation (ICCUBEA), 2015 International Conference on*, pages 811–816. IEEE, 2015.

[17] Lee E Goldstein, Ann M Rasmusson, B Steve Bunney, and Robert H Roth. Role of the amygdala in the coordination of behavioral, neuroendocrine, and prefrontal cortical monoamine responses to psychological stress in the rat. *Journal of Neuroscience*, 16(15):4787–4798, 1996.

[18] P Ekman, RW Levenson, and WV Friesen. Autonomic nervous system activity distinguishes among emotions. *Science*, 221(4616):1208–1210, 1983. URL: http://science.sciencemag.org/content/221/4616/1208, arXiv:http://science.sciencemag.org/content/221/4616/1208.full.pdf, doi:10.1126/science.6612338.

[19] Robert W Levenson. The autonomic nervous system and emotion. *Emotion Review*, 6(2):100–112, 2014.

[20] Robert W Levenson. Blood, sweat, and fears: The autonomic architecture of emotion. *Annals of the New York Academy of Sciences*, 1000(1):348–366, 2003.

[21] Albert Mehrabian. Basic dimensions for a general psychological theory implications for personality, social, environmental, and developmental studies. 1980.

[22] James A Russell. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161, 1980.

[23] P.Lang M.Bradley. Measuring emotion: The self-assessment semantic differential manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25:49–59, 1994.

[24] Silvan Tomkins. *Affect imagery consciousness: Volume I: The positive affects*. Springer publishing company, 1962.

[25] Lang Peter J Davis, Michael. *Handbook of psychology, Volume 3: Biological psychology.* John Wiley & Sons, Inc, 2003.

[26] Peter J Lang. The mechanics of desensitization and the laboratory study of human fear. *Behavior therapy: Appraisal and status. New York: McGraw-Hill*, pages 160–191, 1969.

[27] Margaret M Bradley and Peter J Lang. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59, 1994.

[28] John L Andreassi. *Psychophysiology: Human behavior and physiological response*. Psychology Press, 2010.

[29] Wolfram Boucsein. *Electrodermal activity.* Springer Science & Business Media, 2012.

[30] Marieke van Dooren, Joris H Janssen, et al. Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology & behavior*, 106(2):298–304, 2012.

[31] Philippa Philippa Jane Mothersill. *The form of emotive design.* PhD thesis, Massachusetts Institute of Technology, 2014.

[32] Robert Plutchik. *The emotions.* University Press of America, 1991.

[33] Robert Plutchik. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American scientist*, 89(4):344–350, 2001.

[34] Köhler. Gestalt psychology. 1929.

[35] Geoffrey L Collier. Affective synesthesia: Extracting emotion space from simple perceptual stimuli. *Motivation and emotion*, 20(1):1–32, 1996.

[36] Moshe Bar and Maital Neta. Humans prefer curved visual objects. *Psychological science*, 17(8):645–648, 2006.

[37] Tamirat Abegaz, Edward Dillon, and Juan E Gilbert. Exploring affective reaction during user interaction with colors and shapes. *Procedia Manufacturing*, 3:5253–5260, 2015.

[38] Philippa Mothersill and V Michael Bove Jr. The emotivemodeler: An emotive form design cad tool. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pages 339–342. ACM, 2015.

[39] E Stotland, KE Matthews Jr, SE Sherman, RO Hansson, and BZ Richardson. Fantasy, empathy, and helping, 1978.

[40] William Ickes. *Everyday mind reading: Understanding what other people think and feel*. Prometheus Books, 2010.

[41] Benjamin MP Cuff, Sarah J Brown, Laura Taylor, and Douglas J Howat. Empathy: a review of the concept. *Emotion Review*, 8(2):144–153, 2016.

[42] Carl R Rogers. Becoming a person. In *Symposium on Emotional Development, Oberlin College, Oberlin, OH, US; This chapter represents a lecture by Dr. Rogers given at the aforementioned symposium.* Association Press, 1957.

[43] Theresa Wiseman. A concept analysis of empathy. *Journal of advanced nursing*, 23(6):1162–1167, 1996.

[44] C Sue Carter, James Harris, and Stephen W Porges. 13 neural and evolutionary perspectives on empathy. *The social neuroscience of empathy*, page 169, 2011.

[45] Jean Ed Decety and William Ed Ickes. *The social neuroscience of empathy.* MIT Press, 2009.

[46] Stephen W Porges. The polyvagal perspective. *Biological psychology*, 74(2):116–143, 2007.

[47] Giacomo Rizzolatti and Michael A Arbib. Language within our grasp. *Trends in neurosciences*, 21(5):188–194, 1998.

[48] Giacomo Rizzolatti. The mirror neuron system and its function in humans. *Anatomy and embryology*, 210(5-6):419–421, 2005.

[49] Anne Böckler, Lukas Herrmann, Fynn-Mathis Trautwein, Tom Holmes, and Tania Singer. Know thy selves: Learning to understand oneself increases the ability to understand others. *Journal of Cognitive Enhancement*, 1(2):197–209, 2017.

[50] Sang Won Lee. Computer-mediated empathy, 2019. `arXiv:1910.02368`.

[51] Maria Shugrina, Margrit Betke, and John Collomosse. Empathic painting: Interactive stylization through observed emotional state. In *Proceedings of the 4th International Symposium on Non-photorealistic Animation and Rendering*, NPAR '06, pages 87–96, New York, NY, USA, 2006. ACM. URL: `http://doi.acm.org/10.1145/1124728.1124744`, `doi:10.1145/1124728.1124744`.

[52] Dana Kirsch. The sentic mouse: Developing a tool for measuring emotional valence. Technical report, MIT Media Laboratory Perceptual Computing Section, 1997.

[53] Clayton Epp, Michael Lippold, and Regan L Mandryk. Identifying emotional states using keystroke dynamics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 715–724. ACM, 2011.

[54] SV Kedar, DS Bormane, Aaditi Dhadwal, Shiwali Alone, and Rashi Agarwal. Automatic emotion recognition through handwriting analysis: a review. In *Computing Communication Control and Automation (ICCUBEA), 2015 International Conference on*, pages 811–816. IEEE, 2015.

[55] Janet Fisher, Anish Maredia, Anita Nixon, Nerissa Williams, and Jonathan Leet. Identifying personality traits, and especially traits resulting in violent

behavior through automatic handwriting analysis. *Proceedings of Student-Faculty Research Day, CSIS, Pace University*, 2012.

[56] Michael Fairhurst, Meryem Erbilek, and Cheng Li. Enhancing the forensic value of handwriting using emotion prediction. In *Biometrics and Forensics (IWBF), 2014 International Workshop on*, pages 1–6. IEEE, 2014.

[57] Gerald A Gladstein. What it all means. In *Empathy and Counseling*, pages 173–189. Springer, 1987.

[58] Mariam Hassib, Daniel Buschek, Paweł W Wozniak, and Florian Alt. Heartchat: Heart rate augmented mobile chat to support empathy and awareness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 2239–2251. ACM, 2017.

[59] Fannie Liu, Geoff Kaufman, and Laura Dabbish. The effect of expressive biosignals on empathy and closeness for a stigmatized group member. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):201, 2019.

[60] Sandra Carvalho, Jorge Leite, Santiago Galdo-Álvarez, and Óscar F Gonçalves. The emotional movie database (emdb): A self-report and psychophysiological study. *Applied psychophysiology and biofeedback*, 37(4):279–294, 2012.

[61] Margaret M Bradley and Peter J Lang. Affective norms for english words (anew): Instruction manual and affective ratings. Technical report, Citeseer, 1999.

[62] Margaret M Bradley and Peter J Lang. Affective norms for english text (anet): affective ratings of text and instruction manual. Technical report, NIMH Center for Emotion and Attention University of Florida, 2007.

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

[64] Akio Wakabayashi, Simon Baron-Cohen, Sally Wheelwright, Nigel Goldenfeld, Joe Delaney, Debra Fine, Richard Smith, and Leonora Weil. Development of short forms of the empathy quotient (eq-short) and the systemizing quotient (sq-short). *Personality and individual differences*, 41(5):929–940, 2006.

[65] Philippa Mothersill and V Michael Bove Jr. Emotivemodeler. *interactions*, 22(6):14–15, 2015.

[66] Wikipedia. Rhinoceros 3d.

[67] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[68] Jiawen Han, George Chernyshov, Dingding Zheng, Peizhong Gao, Takuji Narumi, Katrin Wolf, and Kai Kunze. Sentiment pen: recognizing emotional context based on handwriting features. In *Proceedings of the 10th Augmented Human International Conference 2019*, page 24. ACM, 2019.

Appendices

A. Questions in the short version of Empathy Quotient

Below is a list of statements. Subjects will read each statement carefully and rate how strongly you agree or disagree with it by choosing your answer from *Strongly Agree*,*Slightly Agree*, *Slightly Disagree*, and *Strongly Disagree*.

1. I can easily tell if someone else wants to enter a conversation.
2. I really enjoy caring for other people.
3. I find it hard to know what to do in a social situation.
4. I often find it difficult to judge if something is rude or polite.
5. In a conversation, I tend to focus on my own thoughts rather than on what my listener might be thinking.
6. I can pick up quickly if someone says one thing but means another.
7. It is hard for me to see why some things upset people so much.
8. I find it easy to put myself in somebody else ’ s shoes.
9. I am good at predicting how someone will feel.
10. I am quick to spot when someone in a group is feeling awkward or uncomfortable.
11. I can ’ t always see why someone should have felt offended by a remark.
12. I don ’ t tend to find social situations confusing.

13. Other people tell me I am good at understanding how they are feeling and what they are thinking.
14. I can easily tell if someone else is interested or bored with what I am saying.
15. Friends usually talk to me about their problems as they say that I am very understanding.
16. I can sense if I am intruding, even if the other person doesn ’ t tell me.
17. Other people often say that I am insensitive, though I don ’ t always see why.
18. I can tune into how someone else feels rapidly and intuitively.
19. I can easily work out what another person might want to talk about.
20. I can tell if someone is masking their true emotion.
21. I am good at predicting what someone will do.
22. I tend to get emotionally involved with a friend ’ s problems.